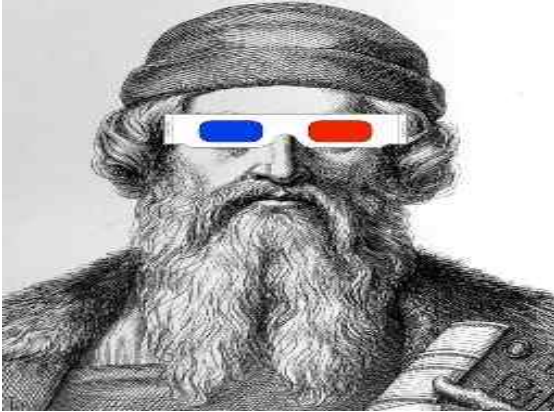# Project Gutenberg
4 July 1971
4 July 2011

4 July 1971 - 4 July 2011 >
An album to celebrate the 40ᵗʰ anniversary of Project Gutenberg

This album is dedicated to all Project Gutenberg and Distributed Proofreaders volunteers worldwide.
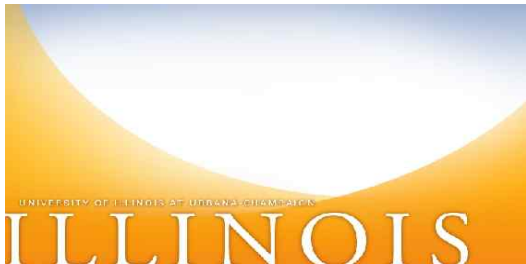

The first ebook was available on 4 July 1971, as eText #1 of Project Gutenberg, a visionary project launched by Michael Hart to create free electronic versions of literary works and disseminate them worldwide. In the 16th century, Gutenberg allowed anyone to have print books for a small cost. In the 21st century, Project Gutenberg would allow anyone to have a digital library at no cost. Project Gutenberg got its first boost with the invention of the web in 1990, and its second boost with the creation of Distributed Proofreaders in 2000.

4 July 1971 > eText #1 > *The United States Declaration of Independence*

On 4 July 1971, Michael Hart keyed in *The United States Declaration of Independence* (signed on 4 July 1776) to the mainframe he was using, in upper case, because there was no lower case yet. The file was 5 K. Michael mentioned to the 100 users of the pre-internet of the time where the etext was stored -- though without a hypertext link, because the web was still 20 years ahead. It was downloaded by six users.

4 July 1971 > As recalled by Michael Hart

As recalled by Michael Hart in January 2009: "On July 4, 1971, while still a freshman at the University of Illinois (UI), I decided to spend the night at the Xerox Sigma V mainframe at the UI Materials Research Lab, rather than walk miles home in the summer heat, only to come back hours later to start another day of school. I stopped on the way to do a little grocery shopping to get through the night, and day, and along

with the groceries they put in the faux parchment copy of *The U.S. Declaration of Independence* that became quite literally the cornerstone of Project Gutenberg. That night, as it turned out, I received my first computer account -- I had been hitchhiking on my brother's best friend's name, who ran the computer on the night shift. When I got a first look at the huge amount of computer money I was given, I decided I had to do something extremely worthwhile to do justice to what I had been given. (...) As I emptied out groceries, the faux parchment *Declaration of Independence* fell out, and the light literally went on over my head like in the cartoons and comics... I knew what the future of computing, and the internet, was going to be... 'The Information Age.' The rest, as they say, is history." (NEF Interview)

July 1971 > Project Gutenberg's mission

Michael decided to use the huge amount of computer time he had been given to search the literary works that were stored in libraries, and to digitize these works. A book would become a continuous text file instead of a set of pages. Project Gutenberg's mission would be to put at everyone's disposal, in electronic versions, as many literary works as possible for free.

1972 > *The United States Bill of Rights*

After keying in *The United States Declaration of Independence* in 1971, Michael typed in a longer text, *The United States Bill of Rights,* in 1972, i.e. the first ten amendments added in 1789 to the Constitution (dated 1787) and defining the individual rights of the citizens and the distinct powers of the Federal Government and the States.

1973 > *The United States Constitution*

A volunteer typed in *The United States Constitution* in 1973.

1974-1988 > *The Bible*

From one year to the next, disk space was getting larger, by the standards of the time -- there was no hard disk yet --, making it possible to store larger files. Volunteers began typing in *The Bible*, with one individual book at a time, and a file for each book.

1974-1988 > *The Collected Works of William Shakespeare*

Michael typed in *The Collected Works of William Shakespeare*, with the help of some volunteers, one play at a time, and a file for each play. This edition of Shakespeare was never released, unfortunately, due to changes in copyright law. Shakespeare's works belong to public domain, but comments and notes may be copyrighted, depending on the publication date. Other editions of Shakespeare from public domain were released a few years later.

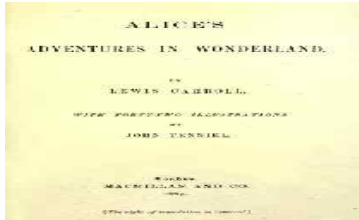## August 1989 > eBook #10 > *The King James Bible*

Its critics long considered Project Gutenberg as impossible on a large scale. But Michael went on keying book after book during many years, with the help of some volunteers. In August 1989, Project Gutenberg completed its 10th ebook, *The King James Bible* (1769), both testaments, and 5 M for all files.

1990 > The invention of the web

In 1990, there were 250,000 internet users. The standard was 360 K disks. The web was in its infancy, with Tim Berners-Lee dealing with HTTP (HyperText Transfer Protocol) and hyperlinks at CERN (European Center for Nuclear Research) in Geneva, Switzerland.

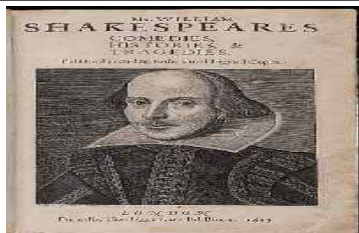| January 1991 > *Alice's Adventures in Wonderland*	
	In January 1991, Michael typed in *Alice's Adventures in Wonderland* (1865), by Lewis Carroll.

July 1991 > *Peter Pan*	
	In July 1991, Michael typed in *Peter Pan* (1904), by James M. Barrie. These two classics of childhood literature each fit on one disk.

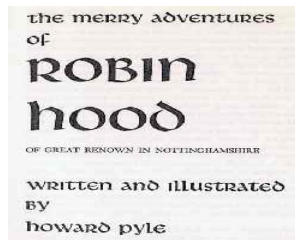| November 1993 > Mosaic, the first public browser	
	The first public browser, Mosaic, was released in November 1993. It became easier to circulate etexts and recruit volunteers. From 1991 to 1996, the number of ebooks doubled every year, with one ebook per month in 1991, two ebooks per month in 1992, four ebooks per month in 1993, and eight ebooks per month in 1994.

January 1994 > eBook #100 > *The Complete Works of William Shakespeare*	
	In January 1994, Project Gutenberg released *The Complete Works of William Shakespeare* as eBook #100. Shakespeare wrote most works between 1590 and 1613. The steady growth went on, with an average of eight ebooks per month in 1994, 16 ebooks per month in 1995, and 32 ebooks per month in 1996.

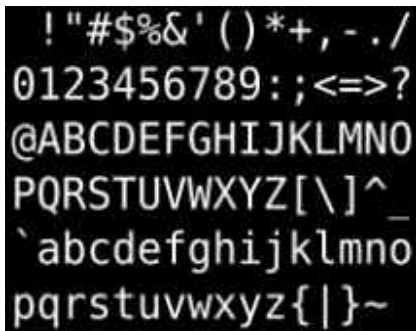|	
June 1997 > *The Merry Adventures of Robin Hood*	
	In June 1997, Project Gutenberg released *The Merry Adventures of Robin Hood* (1883), by Howard Pyle.
August 1997 > eBook #1000 > *La Divina Commedia*, by Dante	
	Project Gutenberg reached 1,000 ebooks in August 1997. eBook #1000 was *La Divina Commedia* (1321), by Dante Alighieri, in Italian, its original language.
1997 > Three main sections	
	With the number of ebooks on the rise, three main sections were set up: (1) "Light Literature", such as *Alice's Adventures in Wonderland*, *Through the Looking-Glass*, *Peter Pan* and *Aesop's Fables;* (2) "Heavy Literature", such as the Bible, Shakespeare's works, *Moby Dick* and *Paradise Lost;* (3) "Reference Literature", such as *Roget's Thesaurus*, almanacs, and a set of encyclopedias and dictionaries. A more detailed classification was released years later.
	Project Gutenberg's goal has been more about selecting books intended for the general public than providing authoritative editions. As explained on the website in 1997: "We do not write for the reader who cares whether a certain phrase in Shakespeare has a ':' or a ';' between its clauses. We put our sights on a goal to release etexts that are 99.9% accurate in the eyes of the general reader."

1998 > A collection in Plain Vanilla ASCII

The etexts, later called ebooks, were stored in the simplest way, using the low set of ASCII, called Plain Vanilla ASCII, for them to be read on any hardware and software. As a text file, a book could be easily copied, indexed, searched, analyzed, and compared with other books. Pietro di Miceli was the sole driving force behind Project Gutenberg's website for nearly a decade, after making the first pages in 1994.

August 1998 > "We consider etext to be a new medium."

As explained by Michael Hart in August 1998: "We consider etext to be a new medium, with no real relationship to paper, other than presenting the same material, but I don't see how paper can possibly compete once people each find their own comfortable way to etexts, especially in schools. (...) My own personal goal is to put 10,000 etexts on the Net *[a goal reached in October 2003]* and if I can get some major support, I would like to expand that to 1,000,000 and to also expand our potential audience for the average etext from 1.x% of the world population to over 10%, thus changing our goal from giving away 1,000,000,000,000 etexts to 1,000 times as many, a trillion and a quadrillion in U.S. terminology." (NEF Interview)

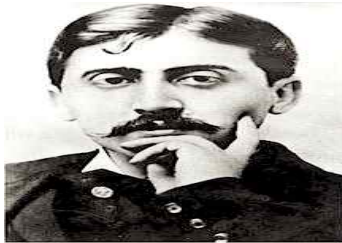## May 1999 > eBook #2000 > *Don Quijote*, by Cervantes

Project Gutenberg reached 2,000 ebooks in May 1999. eBook #2000 was *Don Quijote* (1605), by Cervantes, in Spanish, its original language.

October 2000 > Distributed Proofreaders

Distributed Proofreaders was founded in October 2000 by Charles Franks to share the proofreading of ebooks between many volunteers. Volunteers choose one of the digitized books available on the website and proofread a given page, or several pages, as they wish. It is recommended they do one page per day if possible. It may not seem much, but with thousands of volunteers it really adds up.

December 2000 > eBook #3000 > *À l'ombre des jeunes filles en fleurs*, by Marcel Proust

Project Gutenberg reached 3,000 ebooks in December 2000. eBook #3000 was *À l'ombre des jeunes filles en fleurs* (*In the Shadow of Young Girls in Flower*), vol. 3 (1919), by Marcel Proust, in French, its original language. From 1998 to 2000, there was an average of 36 new ebooks per month.

August 2001 > Project Gutenberg Australia

Project Gutenberg Australia was launched in August 2001. The collection included 500 ebooks in July 2005, and 1,500 ebooks in April 2007.

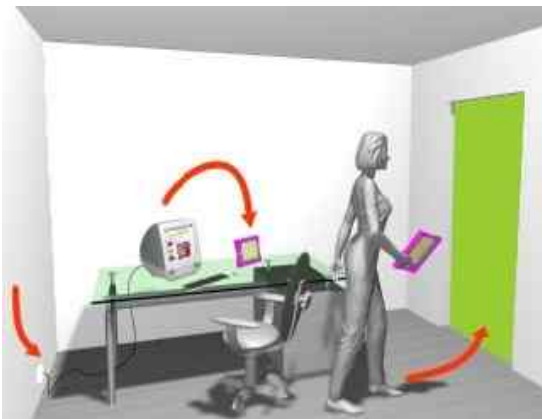## October 2001 > eBook #4000 > *The French Immortals Series*

Project Gutenberg reached 4,000 ebooks in October 2001. eBook #4000 was *The French Immortals Series* (1905), in English. This book is an anthology of short fictions by authors from the French Academy (*Académie Française*): Emile Souvestre, Pierre Loti, Hector Malot, Charles de Bernard, Alphonse Daudet, and others. In 2001, there was an average of 104 new ebooks per month.

April 2002 > eBook #5000 > *The Notebooks of Leonardo da Vinci*

Project Gutenberg reached 5,000 ebooks in April 2002. eBook #5000 was *The Notebooks of Leonardo da Vinci*, an English version of Leonardo's early 16th-century writings in Italian. Since its release, this ebook has constantly stayed in the Top 100 downloaded ebooks.

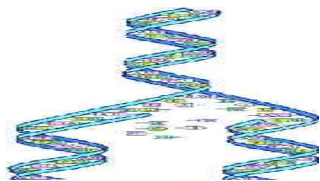## 2002 > 1.44 M standard disks and zipped files

In 1991, *Alice's Adventures in Wonderland* and *Peter Pan* each fit on one 360 K disk, the standard of the time. In 2002, the standard disk was 1.44 M and could be compressed as a zipped file. A practical file size is about 3 million characters, more than long enough for the average book. The ASCII version of a 300-page novel is 1 M. A bulky book can fit in two ASCII files, that can be downloaded as is or zipped. 50 hours on average are necessary to get an ebook selected, copyright-cleared, scanned, proofread, formatted, and assembled.

Spring 2002 > 25% of all public domain works

In spring 2002, the Project Gutenberg collection represented 25% of all public domain works freely available on the web and listed in the Internet Public Library (IPL). In 2002, there was an average of 203 new ebooks per month.

November 2002 > *The Human Genome Project*

In November 2002, Project Gutenberg released the 75 files of *The Human Genome Project*, with files of dozens or hundreds of megabytes, shortly after its initial release in February 2001 as a work from public domain.

September 2003 > Project Gutenberg Audio eBooks

In September 2003, Project Gutenberg Audio eBooks was launched as a collection of human-read ebooks, and the Sheet Music Subproject as a collection of digitized music sheet and music recordings. A collection of still and moving pictures was also available.

October 2003 > eBook #10000 > The Magna Carta

1,000 ebooks in August 1997, 2,000 ebooks in May 1999, 3,000 ebooks in December 2000, 4,000 ebooks in October 2001, 5,000 ebooks in April 2002, 10,000 ebooks in October 2003. eBook #10000 was *The Magna Carta*, the first English constitutional text, signed in 1215.

October 2003 > The collection doubled in 18 months

From April 2002 to October 2003, in 18 months, the collection of ebooks doubled, going from eBook #5000 to eBook #10000, with an average of 348 new ebooks per month in 2003. The fast growth was the hard work of Distributed Proofreaders, now going at full speed to digitize books and share the proofreading of ebooks between many volunteers.

December 2003 > The first Project Gutenberg DVD

eBooks were also copied on CDs and DVDs. As blank CDs and DVDs cost next to nothing, Project Gutenberg began burning and sending a free CD or DVD to anyone asking for it. People were encouraged to make copies for a friend, a library or a school. Released in August 2003, the Best of Gutenberg CD contained 600 ebooks. The first Project Gutenberg DVD was released in December 2003, with the burning of most titles (9,400 ebooks).

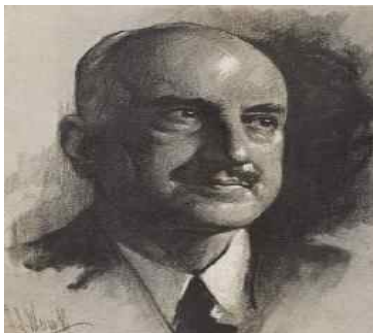## 2003 > The Project Gutenberg Consortia Center (PGCC)

The Project Gutenberg Consortia Center (PGCC) was affiliated with Project Gutenberg in 2003. Since 1997, PGCC had been gathering collections of existing ebooks, as a complement to Project Gutenberg working on producing ebooks. As explained by Michael Hart in January 2009: "The Project Gutenberg Consortia Center has over 75,000 ebooks rendered as PDF files (…). The difference? These files were prepared by other eLibraries, not Project Gutenberg, and are using our worldwide distribution network to be seen." (NEF Interview)

January 2004 > Project Gutenberg Europe

In Europe, Project Rastko, based in Belgrade, Serbia, launched Project Gutenberg Europe (PG Europe) in January 2004, as well as Distributed Proofreaders Europe (DP Europe), with a web interface in 12 languages. 100 ebooks were available in June 2005, and 500 ebooks in October 2008, in several languages, as a reflection of European linguistic diversity. eBooks were encoded in Unicode instead of ASCII, to take into account many different alphabets.

January 2005 > eBook #15000 > *The Life of Reason*, by George Santayana

In January 2005, Project Gutenberg reached 15,000 ebooks. eBook #15000 was *The Life of Reason* (1906), by George Santayana. The driving force behind Project Gutenberg has not only been great authors, but also outstanding volunteers like Greg Newby, director and CEO of the Project Gutenberg Literary Archive Foundation (PGLAF) since 2001, and Marcello Perathoner, the webmaster and main software developer of the www.gutenberg.org website.

July 2005 > eBooks in 42 languages

What about languages? There were ebooks in 25 languages in February 2004, and in 42 languages in July 2005, including Sanskrit and the Mayan languages. The seven main languages -- with more than 50 ebooks -- were English (with 14,548 ebooks on 27 July 2005), French (577 ebooks), German (349 ebooks), Finnish (218 ebooks), Dutch (130 ebooks), Spanish (103 ebooks), and Chinese (69 ebooks).

January 2006 > Project Gutenberg PrePrints

Project Gutenberg PrePrint Reading Room

Project Gutenberg PrePrints was launched in January 2006 to collect items submitted to Project Gutenberg which were interesting enough to be available online, but not ready yet to be added to the main Project Gutenberg collection, because of missing data, low-quality files, formats which were not handy, etc. 379 ebooks were available in December 2006, and 2,020 ebooks in February 2009.

December 2006 > eBook #20000 > *Twenty Thousand Leagues Under the Sea* (audiobook), by Jules Verne

In December 2006, Project Gutenberg reached 20,000 ebooks. eBook #20000 was the audiobook of *Twenty Thousand Leagues Under the Sea* (*Vingt mille lieues sous les mers*, 1869), by Jules Verne, in its English version.

If 32 years and 3 months, from July 1971 to October 2003, were necessary to produce the first 10,000 ebooks, 3 years and 2 months, from October 2003 to December 2006, were necessary to produce the next 10,000 ebooks. There were ebooks in 50 languages in December 2006.

December 2006 > *Project Gutenberg News*, the official blog	

In December 2006, Mike Cook launched the blog *Project Gutenberg News* as "the news portal for gutenberg.org", to complement the existing weekly and monthly newsletters. The blog gave for example the production numbers since 2001. The weekly production was 24 ebooks in 2001, 47 ebooks in 2002, 79 ebooks in 2003, 78 ebooks in 2004, 58 ebooks in 2005, and 80 ebooks in 2006. The monthly production was 104 ebooks in 2001, 203 ebooks in 2002, 348 ebooks in 2003, 338 ebooks in 2004, 252 ebooks in 2005, and 345 ebooks in 2006. The yearly production was 1,244 ebooks in 2001, 2,432 ebooks in 2002, 4,176 ebooks in 2003, 4,058 ebooks in 2004, 3,019 ebooks in 2005, and 4,141 ebooks in 2006.

July 2007 > Project Gutenberg Canada

Project Gutenberg Canada (PGC) was launched on 1st July 2007, on Canada Day, by Michael Shepard and David Jones.

December 2007 > Distributed Proofreaders Canada

Distributed Proofreaders Canada (DPC) started production in December 2007. There were 100 ebooks in March 2008, in English, French and Italian.
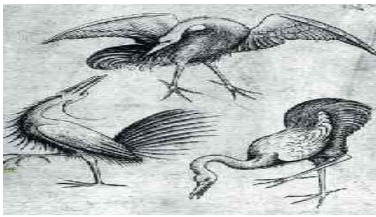
2007 > 15 million ebooks by snail mail

Project Gutenberg sent out 15 million ebooks via CDs and DVDs by snail mail in 2007. The DVD released in July 2006 included 17,000 ebooks. CD and DVD files were also generated as ISO files (since 2005) to be downloaded for burning CDs or DVDs for family, friends, neighbors, schools and libraries.

April 2008 > eBook #25000 > *English Book Collectors*, by William Younger Fletcher

Project Gutenberg reached 25,000 books in April 2008. eBook #25000 was *English Book Collectors* (1902), by William Younger Fletcher.

October 2009 > eBook #30000 > *The Bird Book*, by Chester Albert Reed

Project Gutenberg reached 30,000 books in October 2009. eBook #30000 was *The Bird Book* (1915), by Chester Albert Reed.

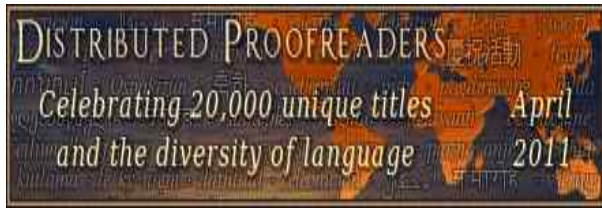**October 2010 > 10th anniversary of Distributed Proofreaders**

Distributed Proofreaders celebrated its 10th anniversary in October 2010, with more than 18,000 books digitized and proofread during ten years by thousands of volunteeers. We sent them a digital bilingual anniversary card from Paris, France.

December 2010 > 33,000 high-quality ebooks

In December 2010, Project Gutenberg offered 33,000 high-quality ebooks , with tens of thousands of downloads every day, websites in the United States, in Australia, in Europe and in Canada, and 40 mirror sites worldwide. eBooks are available in various formats for any electronic device (computer, PDA, mobile phone, smartphone and ebook reader).

April 2011 > 20,000 ebooks processed by Distributed Proofreaders

In April 2011, Distributed Proofreaders celebrated "20,000 unique titles [sent] to the bookshelves of Project Gutenberg, free to enjoy for everybody. Distributed Proofreaders is a truly international community. People from over the world contribute." (excerpt from the website)

April 2011 > 30,000 English ebooks in Project Gutenberg

The 30,000th English-language ebook was posted on 12 April 2011. Its title was *The Religions of Ancient Egypt and Babylonia* (1903), by Archibald Henry Sace. eBooks were also available in 59 other languages. 13 languages were lucky enough to have more than 50 books. 46 languages had up to 50 books. Please join Distributed Proofreaders and contribute.

June 2011 > 36,000 ebooks in many languages

In June 2011, the 14 main languages were English (with 30,569 ebooks on 30 June 2011), French (1,821 ebooks), German (772 ebooks), Finnish (589 ebooks), Dutch (534 ebooks), Portuguese (515 ebooks), Chinese (405 ebooks), Spanish (307 ebooks), Italian (289 ebooks), Greek (144 ebooks), Latin (79 ebooks), Esperanto (74 ebooks), Swedish (68 ebooks), and Tagalog (55 ebooks).

4 July 2011 > 40th anniversary of Project Gutenberg

Text: Marie Lebert
Cover: Jean-Paul Pfirrmann
Copyrighted images: @folio Project, Claude Rayon, Distributed Proofreaders (all websites), Librairie Ulysse, Project Gutenberg (all websites), University of Illinois
Public domain images: Wikipedia