

## Project Gutenberg: A Prototype of Modern Open Movements

**Shamba Dutta**

M.Phil Scholar, DLISc, University of Calcutta, Kolkata, India

Email: [duttshambo@gmail.com](mailto:duttshambo@gmail.com)

### *Abstract*

**Purpose:** *The purpose of this paper is to bring to the LIS professionals, students and researchers some relevant facts about world's first e-book delivery service that was inceptioned years before internet came to exist, and is still continuing as a rich source of many rare and current e-books available in various user friendly formats in today's web environment.*

**Design/methodology/approach:** *The author has explored the Project Gutenberg site and other relevant sources to find its history, philosophy, and the policies, and also various features and statistics of the services it is currently providing.*

**Findings:** *It is found that the Project Gutenberg is an organised and continually updated source of many rare books which are offered for free use to the global readership. It maintains a strong e-book delivery system too. The books are classified under various heads searchable through both online and offline catalogues. The Project has extended its service potential by resorting to the help of its associates like partners, sister projects and affiliates.*

**Research Limitations:** *The present paper is limited to the Project Gutenberg only and does not aim at any comparative study to other open sources of e-books. The article is not analytical type. The full set of data regarding its growth pattern is not available; and therefore, no vigorous statistical method could be applied to find the full picture of the growth from its beginning to the present.*

**Practical Implications:** *The use of the Project Gutenberg as a free e-books source can be promoted.*

**Originality/value:** *Its importance as the prototype of an Open Access Initiative can be highlighted and might be linked historically with the modern OAI and web archiving enthusiasms.*

**Keywords:** *Project Gutenberg, E-Books Platform, Creative Commons, Distributed Proofreaders, Open Access Initiatives, Web archiving.*

**Paper Type:** *Descriptive.*

**1. Introduction:** The modern World of information dissemination is undoubtedly digital and largely based on internet. Digital libraries, repositories and archives are common words today. But the earliest attempt to reach the user electronically, and free of charges was started in 1971; evidently much before internet came into being. The name of the project is Project Gutenberg, named after the great Guru of printing Johannes Gutenberg (1395-1468) whose invention of the printing press marked the beginning of democratisation of information. The aim of the project as sworn by Michael Hart (1947-2001), its founder, was to encourage the creation and distribution of e-Books. In the mission statement of the project, Michael Hart writes, "This mission is, as much as possible, to encourage all those, who are interested in making e-Books and helping to give them away". It was world's first Digital Library. A totally volunteer effort to digitise and archive cultural works, it was in fact the prototype for today's web archiving and Open Access Initiatives (OAI), both in purpose and spirit, notwithstanding the modern technological growth and most importantly, the humble in-house project that started with a single book is still running as a moderate source of free e-books for global access.

**2. Brief History:** Project Gutenberg was started in 1971 by Michael Hart when a student in the University of Illinois, with the digitization of the United States Declaration of

### **How to Cite this Article**

**Dutta, S. (2015). Project Gutenberg: A Prototype of Modern Open Movements. LIS Links Newsletter, 1(7), 2-8.**

Independence, and with the attempt of sending it to the members in the network (ARPANET). It was the first electronic book ever created (Hane, 2004). Actually, Hart was trying to do something fruitful for the reading community that will be worthy of the huge amount of computer time he was allotted (Hart, 1992). The particular mainframe computer he worked with was one of the 15 nodes on ARPANET, the computer network that subsequently became internet (Project Gutenberg, 2015). Hart’s Project Gutenberg activities were based on Illinois Benedictine College up to mid 90s of the last century. Later, Carnegie Mellon University took over the financial burden of the project. In 2004 online catalogue of the available texts was launched, making it more user- friendly. At present it is hosted by *Ibiblio* at the University of North Carolina, at Chapel Hill, the USA. The texts were typed manually. The use of scanned and OCRed texts grew gradually, because typing was too much time taking (Vara, 2005).

**3. Philosophy of the Project:** The philosophy of the Project Gutenberg is based on the belief that the greatest value created by computers would not be computing. It would be the storage, retrieval and searching what was stored in the libraries. Thus Hart made use of the “Replicator Technology” which means that once a book is digitised, then one can get any number of copies from it. This is also true for any other form of documents. Any person at any part of this world or even one in space can get a copy of a book through satellite transmission (Hart, 1992). Another philosophy was that the texts would be made available at a very cheap or virtually no cost. The initial goal of this project was to make available to the public the 10,000 most consulted books by the turning of the 20 century (Project Gutenberg, 2015). To read online or download collection no fee or registration is required. Today some 50000 free e-books are available in the project and more than 100,000 free e-books are available through their Partners, Affiliates and Resources (Project Gutenberg Main Page, 2015).

Early in the last decade there was an attempt to launch Project Gutenberg 2, which was sometimes thought to be the 2.0 version of the project, but was also mentioned as Project Gutenberg II. It was cofounded by Michael Hart, the founder of Project Gutenberg and John S. Guagliardo, founder and the executive director of the World E-Book Library (WEL). But this initiative gave birth to a lot of controversy regarding some of its policies that showed deviation from its original free use policy. The debate over the issue is unfortunately lost now and nothing really retrieves against search by Project Gutenberg 2. Therefore, there is reason to believe that they abandoned the project and Project Gutenberg 2 did not materialise (Hane, 2004).

**4. Growth:** The following data set is a very brief and precise indicator of the project’s growth. Although with a rather slow start, the project shows almost an exponential growth which is prominent from the shape of the curve below. Having the fact at hand that 3042 books were posted from 1971 to 2000, the rate of posting of new books during that period is given in the following matrix.

Year	2001	2002	2003	2004	2005	2006	2007	2008
Books Posted	1244	2432	4176	4058	3019	4141	4049	6392 + 2431 Pre-Prints

**Table 1: Growth of Collection in Project Gutenberg**

The following curve shows the growth in total number of books published by Project Gutenberg over the period from 1994 to 2008.

**How to Cite this Article**

**Dutta, S. (2015). Project Gutenberg: A Prototype of Modern Open Movements. LIS Links Newsletter, 1(7), 2-8.**

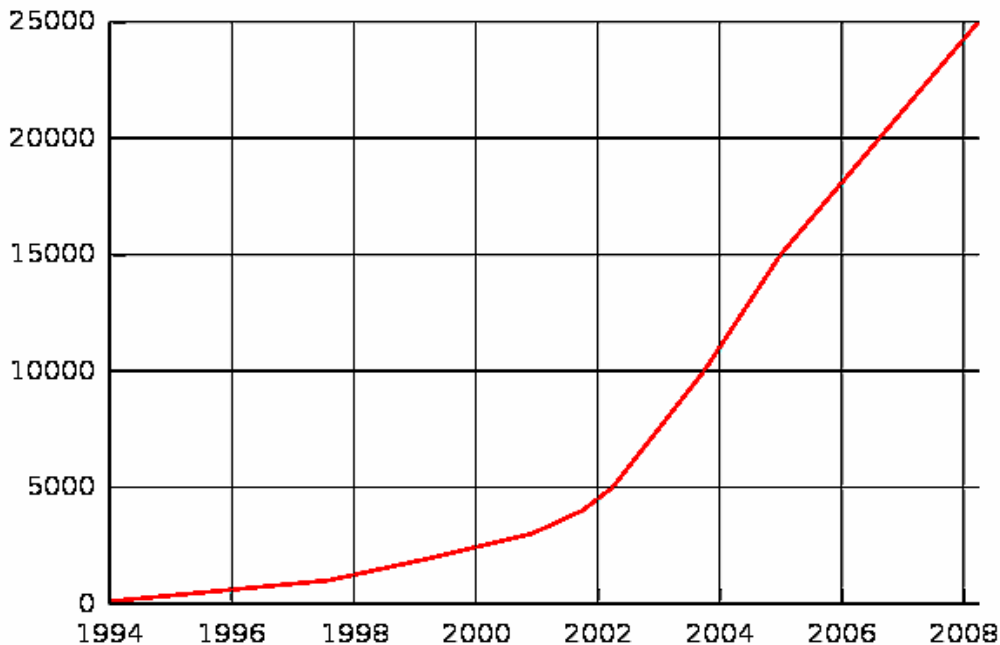


Fig.: Growth of Collection in Project Gutenberg

Courtesy: Wikipedia

In 2009 and 2010 number of books reported to be posted is 3190 and 4408 respectively. The data as of July 20, 2011 show that total 36701 books were in the main Project Gutenberg and 5588 books hosted by the other sister projects. The number of books has now exceeded 49000 (Statistics, 2011). The information about its growth is updated in the Project Gutenberg newsletter regularly. The April 2015 issue of the newsletter (Newsletter April, 2015) reports the following amongst other things:

“Recently, the Press has been imposed to think in encyclop[a]edia entries and other discovery tools for authors, titles and subjects. This adds some great ways to navigate brows and discover books. It also brings additional third party content right to the search page, which can be helpful for students and others seeking additional information.”

The issue also contains a list of e-Books added during February and March, 2015 (Newsletter April, 2015). The latest, August 2015 issue (Newsletter August, 2015) reports the celebration of Distributed Proofreaders for providing their 30,000th title to Project Gutenberg.

**5. The Collection:** Project Gutenberg offers more than 49200 free e-Books (Project Gutenberg, 2015) all of which were previously published by good and prestigious publishers across the world. During digitisation they were carefully proofread by the thousands of volunteers involved in the project. Over 100,000 free e-books are available through partners, affiliates and resources (Project Gutenberg Main Page, 2015).

**6. Search Facilities:** Project Gutenberg offers the following facilities

**a) Book Shelf:** Users may search for books in the book shelf category. Other than English there are books on the shelves in the *Deutsch, Français, Italiano, Português* (Dutch, French, Italian and Portuguese) language. The book shelf is classified under 21 *subcategories* classifying books in broader subdivisions and indexed alphabetically. Books are searchable up to further intentions in the 283 pages contributed to the book shelf. The 283 pages contain alphabetical indexes of books classified by subject headings in greater depth.

#### How to Cite this Article

Dutta, S. (2015). Project Gutenberg: A Prototype of Modern Open Movements. *LIS Links Newsletter*, 1(7), 2-8.

**b) Catalogue:** Project Gutenberg offers both online and offline catalogues. The offline catalogue is for download and use at home. The online catalogue has four access points: *Author, Title, Language, and Recently Posted*. Author, Title, and Language are indexed alphabetically. Language has two subcategories, such as Language with *more than 50 books, and up to 50 books*. Recently Posted category subdivides as: *last 24 hours, last 7 days and last 30 days*. The online catalogue is always updated save the subject search, which may take some time lag for updates. Search services are available through other three search engines: *Anacieto, Yahoo, and Google*, updated periodically. There is a ready reference of top 100 books, and a RSS feed to deliver current awareness to the users.

**c) Types of Search:** Project Gutenberg offers very effective search services. Three main types of searching are possible within the library catalogue under the 'Search' options viz. Popular, Latest, and Random.

*i) Popular:* In this category books are indexed according to their popularity (may be based on the number of hits or downloads).

*ii) Latest:* In this category books are stored as per their dates of release.

*iii) Random:* This is a random index.

**7. Formats:** Books are mainly available in plain ASCII texts, but it is frequently extended to ISO-8859-1 code which is needed to represent special characters and diacritic marks in languages like French, German, etc. But considering the dull appearance of the plain text and the lack of mark up possibility, other formats such as HTML, PDF, EPUB, MOBI, Plucker, etc. are also made available for some books (Project Gutenberg Main Page, 2015) (Project Gutenberg, 2015).

**8. CD and DVD Projects and Document Delivery Services:** In August 2003, Project Gutenberg launched a CD containing 600 best e-books from the collection. The CD was downloadable as an ISO image. It was also made available by e-mail, free of cost. In December 2003, a DVD was created containing almost the full collection of nearly 10000 items. In early 2004 this DVD was also made available by mail. In July 2007, a new edition of the DVD was released that contained more than 17000 books. In April 2010, a dual-layer DVD was brought out that contain nearly 30000 items. Till 2010, the project has mailed approximately 40,000 discs (The CD and DVD Project, 2014) (Project Gutenberg, 2015). CDs and DVDs are downloadable via Bit-Torrent, Edonkey/Emule, Jigdo, FTP and HTTP according to their applicability although Project Gutenberg recommends downloading via Bit Torrent for faster and error-free copy of the image files. These CD and DVD compilations are copyrighted and licensed under the Creative Commons Attribution-Non Commercial-Share Alike 1.0 license. (The CD and DVD Project, 2014) (Attribution-NonCommercial-ShareAlike 1.0 Generic ).

**9. Associates of Project Gutenberg:** Project Gutenberg works on collaborative basis with associates such as Partners, Sister Projects and Affiliates.

**a) Partners:** Project Gutenberg has two partners- Distributed Proofreaders or DP and Ibiblio.

*i) Distributed Proofreaders or DP:* The DP community allow people to share the tasks of proofreading, verifying and formatting the e-books. Readers are invited to join the group of DP.

*ii) Ibiblio:* Project Gutenberg is hosted by Ibiblio. Ibiblio is the public's library and digital archive which is the main e-book distribution site for the project (Hart, 1992) (Ibiblio Home, 2015).

**b) Sister Projects:** These projects function independently with particular national or linguistic focus, sharing the same ideals, and have been permitted to use the Project Gutenberg trademark. They are: Project Guttenberg of Australia; Project Guttenberg of

#### How to Cite this Article

Dutta, S. (2015). Project Gutenberg: A Prototype of Modern Open Movements. LIS Links Newsletter, 1(7), 2-8.

Canada, Project Gutenberg Consortia Centre; Project Gutenberg DE German literature; Project Gutenberg Europe; Project Gutenberg Self-Publishing Portal; Projekt Ruenberg ordic literature; ReadingRoom.ms, which is the home of the Project Gutenberg Pre-Prints, etc (Partners, Affiliates and Resources, 2014).

**c) Affiliates:** Project Gutenberg has the following affiliates

i) *Classical Archives (Classical Archives.com)*: It is an archive of classics and mythological texts.

ii) *Literal Systems (Literal Systems.org)*: It is a source of human audio e-book readings.

iii) *Librivox (Librivox.org)*: It is another source of human audio e-book readings.

iv) *The Internet Archive*: It is an archive of digital books.

v) *Wattpad*: It contains thousands of titles including self published titles, accessible for both computer and mobile. They support essentially every type of mobile phone through their site: <http://m.wattpad.com>.

vi) *Outernet*: It uses free and open satellite transmission to broadcast content (Partners, Affiliates and Resources, 2014).

**10. Rights Management:** Free use of the materials is the policy of the Project Gutenberg project. The word *free* has double connotation. It may mean free of charge or no cost as well as the right to 'do whatever' with the e-books. (No Cost or Freedom?, 2014). Apart from being free of cost materials of Project Gutenberg can be used *freely* in the United States as the collection is not covered by the US copyright law. Nor does Project Gutenberg claim any new copyright over the format, etc. Therefore, it has given the user liberty to do 'whatever' with the books, provided that the Project Gutenberg trademark is not used. Project Gutenberg bears disclaimer about the restrictions in the nature of use of those materials outside the US, which will fall under the purview of the copyright acts of respective countries (No Cost or Freedom?, 2014) (Project Gutenberg, 2015).

Even authors interested in self publishing their books need to send a copyright permission letter giving Project Gutenberg 'perpetual, non-exclusive, worldwide rights to distribute the books in electronic form on a royalty-free basis' (Submitting Your Own Work How-To, 2014). One is also allowed to select suitable licence terms from Creative Commons if that suits the Project Gutenberg policies and terms. However, spread of open and free access to the materials is the spirit that Project Gutenberg kindles. Therefore, Project Gutenberg encourages to 'copy and give away these CDs and DVDs! Keep one for yourself, and give one to a friend ... or a library ... or a school' (The CD and DVD Project, 2014). It is also specified that 'Verbatim copies of the Project Gutenberg 'Best Of' CD may be sold provided that 20% of the gross proceeds are donated to Project Gutenberg, however, none of the DVDs may be sold, as they contain some material which is copyrighted.'

**11. Book Shelf Category: India:** There is a page in the book shelf titled 'India' in Project Gutenberg. Books are available under subdivisions: Biography, Folklore, History, Literature, Nature, Society, Travel and Religion. The *See also* reference directs to the page in Bookshelf 'Hinduism'.

**12. Other Services:** Project Gutenberg has a wiki where relevant pieces of information about the project, its aims and objectives, philosophy etc are available by search. It also has its mobile site <http://m.gutenberg.org/>. Project Gutenberg also facilitating online self-publishing by contemporary authors 'in virtually any genre or subject' (Project Gutenberg Main Page, 2015). The portal is open at [self.gutenberg.org](http://self.gutenberg.org) where an author needs to register himself/herself and log in to the account to upload the work suitable for publication in Project Gutenberg. Project Gutenberg is interested in publishing books already published rather than new and unpublished ones. The RSS feed facility is also available in the self publishing

#### How to Cite this Article

Dutta, S. (2015). Project Gutenberg: A Prototype of Modern Open Movements. *LIS Links Newsletter*, 1(7), 2-8.

portal, at <http://self.gutenberg.org/rss/authors.aspx>. The RSS Feed brings the news of the 10 most recent author contributions.

**13. People's Participation:** Project Gutenberg also gives a call to every reader to help it create more e-books. People can participate by digitizing more books in order to that one has to join the DP pool at <http://www.pgdp.net/c/>; or by volunteering as audio book reader at <https://librivox.org/>. Anybody can help by reporting an error through e-mail to Project Gutenberg or by finding a missing page of an e-book. (Project Gutenberg, 2015). Project Gutenberg also seeks assistance from the reader by way of donation to raise fund to run this nonprofit making service. Anybody can donate an amount through *Paypal* to Project Gutenberg Literary Archive Foundation.

**14. Conclusions:** Project Gutenberg may not have grown a real giant archive, but its free use policy has made it popular. Books here are easily searchable by clicking on the hyperlinks and downloadable in any given format. However, because of copyright reasons, many recent books could not be included. This is a problem for the *Indian* category of books too. Yet it continues to be a repository of many old and rare books, its main focus being the old copyright free books. But as a prototype of all OAI, Project Gutenberg is still alive in its humble existence. It may also be considered the herald of today's open licence or open access paradigm. Thus the idea occurred to a student 44 years ago may be thought to be the beginning of a movement in which line the fundamental concept of right to access to information or documents in the internet or web would change in the later decades.

## References

- Attribution-NonCommercial-ShareAlike 1.0 Generic* . (n.d.). Retrieved August 20, 2015, from Creative Commons: <http://creativecommons.org/licenses/by-nc-sa/1.0/>
- Hane, P. (2004, May). *Project Gutenberg Progress*. Retrieved July 2, 2015, from Information Today: <http://www.infotoday.com/it/may04/hane1.shtml>
- Hart, M. (1992). *The History and Philosophy of Project Gutenberg* by Michael hart. Retrieved 2015, from Project Gutenberg: [https://www.gutenberg.org/wiki/Gutenberg:The\\_History\\_and\\_Philosophy\\_of\\_Project\\_Gutenberg\\_by\\_Michael\\_Hart](https://www.gutenberg.org/wiki/Gutenberg:The_History_and_Philosophy_of_Project_Gutenberg_by_Michael_Hart)
- Ibiblio Home*. (2015, May). Retrieved July 1, 2015, from Ibiblio: <http://www.ibiblio.org>
- Newsletter April*. (2015, April 2). Retrieved July 3, 2015, from Project guttenberg: <http://www.gutenbergnews.org/20150402/pg-monthly-newsletter-2015-04-02/>
- Newsletter August*. (2015, August 6). Retrieved September 25, 2015, from Project Gutenberg: <http://www.gutenbergnews.org/20150806/pg-monthly-newsletter-2015-08-06/>
- No Cost or Freedom?* (2014). Retrieved July 2, 2015, from Project Gutenberg: [https://www.gutenberg.org/wiki/Gutenberg:No\\_Cost\\_or\\_Freedom%3F](https://www.gutenberg.org/wiki/Gutenberg:No_Cost_or_Freedom%3F)
- Partners, Affiliates and Resources*. (2014, August). Retrieved July 2, 2015, from Project Gutenberg: [https://www.gutenberg.org/wiki/Gutenberg:Partners,\\_Affiliates\\_and\\_Resources](https://www.gutenberg.org/wiki/Gutenberg:Partners,_Affiliates_and_Resources)
- Project Gutenberg*. (2015, July 1). Retrieved July 2, 2015, from Wikipedia: [https://en.wikipedia.org/wiki/Prpject\\_Gutenberg#History](https://en.wikipedia.org/wiki/Prpject_Gutenberg#History)
- Project Gutenberg Main Page*. (2015, June). Retrieved July 3, 2015, from Project Gutenberg: [http://www.gutenbergnews.org/wiki/Main\\_Page](http://www.gutenbergnews.org/wiki/Main_Page)
- Statistics*. (2011, July 20). Retrieved July 2, 2015, from Project Gutenberg News: <http://www.gutenbergnews.org>
- Submitting Your Own Work How-To*. (2014, March 12). Retrieved August 20, 2015, from Project Gutenberg Wiki: [https://www.gutenberg.org/wiki/Gutenberg:Submitting\\_Your\\_Own\\_Work\\_How-To](https://www.gutenberg.org/wiki/Gutenberg:Submitting_Your_Own_Work_How-To)

## How to Cite this Article

Dutta, S. (2015). *Project Gutenberg: A Prototype of Modern Open Movements*. *LIS Links Newsletter*, 1(7), 2-8.

*The CD and DVD Project*. (2014, September). Retrieved July 2, 2015, from Project Gutenberg: [http://www.gutenbergnews.org/wiki/Gutenberg:The\\_CD\\_and\\_DVD\\_Project](http://www.gutenbergnews.org/wiki/Gutenberg:The_CD_and_DVD_Project)  
Vara, V. (2005, December 10). *Project Gutenberg Fears No Google*. Retrieved June 28, 2015, from Wall Street Journal Online (Indian Edition): [http://www.wsj.com/public/article/SB113415403113218620-U\\_OqLOmApoaSvNpy5SjNwvhpW5w\\_20061209.html](http://www.wsj.com/public/article/SB113415403113218620-U_OqLOmApoaSvNpy5SjNwvhpW5w_20061209.html)

---

**How to Cite this Article**

**Dutta, S. (2015). Project Gutenberg: A Prototype of Modern Open Movements. *LIS Links Newsletter*, 1(7), 2-8.**