

The gradual digitisation of all human knowledge

Declaration:

I hereby declare that this is my own work and does not use any materials other than the cited sources and tools. All explanations that I copied directly or in essence are marked as such. This work has not been previously submitted.

NATALY ANGULO

04/12/20

Contents

Introduction	2
The History of Print	2
Digitisation of books	3
The Language Barrier of the Web	4
Conclusion.....	5
Appendix	6
Bibliography	6

Introduction

Books have been the modern form we have used for recording both important events and literature; now, books serve as records to preserve art and history. The World Wide Web (created by Sir Tim Burners-Lee) is a resource on the internet intended for extending the access of information to a global scale. What is fascinating about this invention is how the modern world interacts with the old; prior to the digital revolution, knowledge was much localised to single nations or societies. However, since the web emerged, these resources have been made much less limited to geographical borders and now allow for access to information to be universal. While the intention describes the interest of the many, there are social problems in this spread of information; while knowledge is made available to everyone, the result is desaturation of culture and the loss of intellectual property.

The History of Print

From the beginning of written language (in the early years of the 1st millennium), historical events and discoveries have been documented through generations, evolving past cave drawings and more recently through the book. In the 21st century, we estimate One-hundred, and twenty-nine million books exist worldwide. With so many books scattered across the globe, it is difficult to preserve them all. In the ancient world, the library of Alexandria (LoA) was one of the more considerable attempts to localise all knowledge of the time. Scholars from Greece travelled across Alexander the Great's empire to gather scrolls; regarding these documents, the content was copied and thus preserved in the LoA. While this attempted to 'preserve history' in a local knowledge base, the methods of doing so sometimes meant the library confiscated the original scrolls and instead returned mere copies to the owners (without their consent). This method values preservation over the respect of private property in the name of 'prosperity'(The burning of a library of Alexandria: The crime of the age). Unfortunately, the library (the specifics of the event are still debated) burned down in 48BC, destroying at least forty-thousand scrolls according to Livy's History of Rome (written between 63 BCE and 14 CE). The burning of Alexandria is still considered a catastrophe among scholars and historians. While there were more libraries with rivalling importance to that of Alexandria, only a few scrolls were ever recovered from these places that once existed millennia ago. Tragically, pieces of literature did not have to burn to be lost; naturally, the medium disintegrates over time; the life expectancy of 'papyrus' was just fifty years. While this could be avoided by making copies, it could not truly be escaped.

In the modern age, there are thousands of copies of books, as well as recovered ancient literature being preserved by museums and libraries, how can longevity and preservation still be an issue? A typical book will last around 40 years if kept in the right conditions, while a high-quality book may last around 100 years. Still, the process of book restoration extends the lifespan, at best, 200 years. The legacy of human history is validated through its documentation; how can our history stand the test of time? Unfortunately, books and print media cannot promise the security of knowledge; not just due to the natural degrade of paper, but environmental circumstances. The history of books has not been favoured by war or politics, often being burned, taken from the masses, or rewritten to fit an agenda. One

example of this is the 'Nazi book burnings'; mostly left-winged books and those written by Jewish authors were destroyed (as well as any content considered 'un-German'). A modern display of this problem was revealed as a consequence of the devastating hurricane 'Katrina'; "[In] Tulane university, the main library sat in nine ft of water.' (Mary Sue Coleman, president, Michigan University) thus damaging the content of the books there. Worse than this, the school itself was underfunded by the city of New Orleans; this meant that damaged books could not be immediately replaced. Both environmental and social circumstance also play a role in the destruction of books.

Digitisation of books

Since the invention of books, it has been a reliable method of notation for scholars and authors. However, these are of the past; thus far, books have been useful tools for documentation, but we have the means to document more efficiently; digitisation is a technical way of preserving literature and books of all genres. Several libraries use technology to digitise books by scanning their pages to store digitally. As seen with the LoA, scholars would hand-copy books, word for word, in an attempt to preserve their content. Print media has evolved to mass-copy books; it has become much more effective compared to the means of rewriting by hand. Digitising is no different, except it further facilitates the process. Once the pages are photographed or scanned, then technicians could amend any blemishes, and from then, that document, manuscript, or book could live on multiple servers for safekeeping.

'There is no practical obstacle whatever now, to the creation of an efficient index to all human knowledge, ideas and achievements. Through the creation that is of a complete planetary memory for all of mankind.' – H.G Wells, World Brain (1937).

In 1983, H.G Welles, the founding father of the Sci-Fi genre, published the "World Brain" a book predicting a 'Permanent World Encyclopaedia'. The digital revolution gave birth to the internet and subsequently, the invention of the World Wide Web (in 1989). 'The World Wide Web is a collection of webpages that reside on computers connected to the internet'(PM, Heathcote and RSU Heathcote, 2016), its original intention was to aid in the collaboration among scientists by having them share their findings through the web. However, the web has evolved, and there has been a large investment into making it a platform past an information base: expanding to communication, entertainment, and consumerism. The way the web was originally set up allowed for the sharing of information between researchers; it can be said to be a modern attempt of the LoA. Dissimilarly to the LoA, the web has not been exclusive to scholars alone; the web is now available to anyone with a computer and an internet connection. The main difference is that the web allows access on a global scale instead of a specific geographical point. This difference demonstrates the strong divide between the physical limitations of books compared to the virtual potential of the web. The digitisation of books could produce a globalised knowledge base that will democratising access to the entire globe.

'It can be reproduced exactly and full in Peru, China, Iceland, Central Africa or wherever else.'

'It is no dream, or fantasy, it is a plain statement of a contemporary state of affairs.'
– H.G Wells, World Brain.

A case study to prove this is the Harvard library, consisting of 17 million books, they began digitising their extensive private collection in 2010. Now 6 million is available to the public through Harvard's website. Harvard's efforts were impressive but costly, according to Sidney Verba (Former Director, Harvard Library), it would cost 'several hundred dollars just for a single book'. Search engines are tools that help users navigate the web and its services; one of the most used search engines belongs to the tech giant, Google (receiving approximately 5.9 billion searches a day). Google acquires revenue from posting adverts on its free service, that is typically tailored to its users; its mass success as a business has allowed them to expand their services onto email, cloud storage and other services. Google Books was a project that was developed with the intent of scanning and allowing access to all books. As it is today, Google Books resembles more closely to a bookstore compared to its original intent. Google would have become the first in its scale and reach, to distribute knowledge and bring a sense of equality in terms of accessibility in countries where certain resources, such as libraries, cannot be maintained nor updated. Furthermore, the resources needed are few and obtainable for most of the global population; a third of households [in developing countries] having a computer. In contrast, the share of households with a personal computer in developed countries exceed 80% (refer to [figure 1](#)) (Thomas Aloop, Computer penetration rate among household worldwide 2005-2019). With the advancement of these developing countries, it can be guaranteed that there will be a surge in internet users and device owners. Unlike libraries, this service will allow for resources to be accessed simultaneously by multiple users. Here is where Google Books proves useful. The tech giant will make a large profit off the data that will be collected from users of Google Books who in turn are granted free access to a world encyclopaedia.

Google, in a sense, has followed in the footsteps of the LoA in terms of their overall goal. However, another similarity is the way they have treated the authors: during their mass-digitisation (in 2002) Google was able to digitise a total of 25 million books, some at the time were still in copyright, leading to an ensuing legal battle between the authors guild, publishers, and Google. They thought this a negative idea which would force authors to give up ownership of their own works. Like the LoA, Google is disinterested in private ownership, in its vision for 'prosperity'. In 2011, Google lost the case, therefore bringing to a halt their version of the 'World Encyclopaedia'.

The Language Barrier of the Web

The content on the web is primarily in English, targeting the western civilisation. The automated translation tools available are not entirely accurate meaning that most of the content online is left in its original form. An attempt to overcome, volunteers are called to translate the information- Wikipedia is notorious for this. English is one of the more dominant languages worldwide, with roughly 360 million people that speak it as a first language and one billion that speak it as a second. Will the dominance of English on the web have cultural implications for the rest of the world? Language is undoubtedly a part of the culture; however, what could happen if a dominant language such as English were to spread

worldwide? Perhaps an argument can be made that this will inevitably lead to the loss of language where English instead will be used as a global standard (see [figure 2](#)). H.G Wells' vision for the 'World Brain' allows for the accessibility of knowledge for the globe, yet the web's demographic is majorly for English speakers; is the web becoming a monoculture (an 'English (and American) Brain')? The web has recently emerged: its impact can only be witnessed with time. While a universal language may be practical for "human progress", a language is more than a way of communicating. The loss of language is also the loss of culture, the loss of history; the very thing we want to preserve. It has been proven how an invading language in society can influence and change the native language and culture; an example can be seen as a consequence of colonisation. In Latin-America, native languages are fading as younger generations tend towards the adopted-native language, 'Spanish'. Due to mass media consumption, we can see how the Americanisation of an entire population is slowly leading to the desaturation of Latinx culture. From this, we can make an inference regarding how other cultures may be impacted. Perhaps the web is a repeat of this at a much slower pace, once again, culture is overridden by an external imposer.

Conclusion

The advancement of humankind has allowed for the development of preservation to be handled delicately by technology. Even if Google Books was a ham-fisted attempt of its kind and scale, its importance could not be denied. H.G Well's 'World Brain' was a prediction that, to a certain extent, has come true through the Web. Print is becoming outdated in comparison to newer digital mediums. However, the 'World Brain' was not intended to replace books but rather to allow global access to them. Despite the cost of private ownership, perhaps now the American ideal of individualism is being challenged for the benefit of the global community.

Appendix

Figures

Share of households with a computer at home worldwide from 2005 to 2019

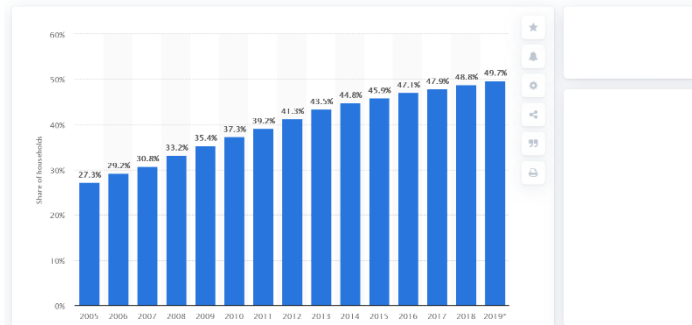


Figure 1, (Mar 2nd, 2020)

In T. Alopp's article on the Computer penetration rate, here you have the increase of worldwide PC ownership.

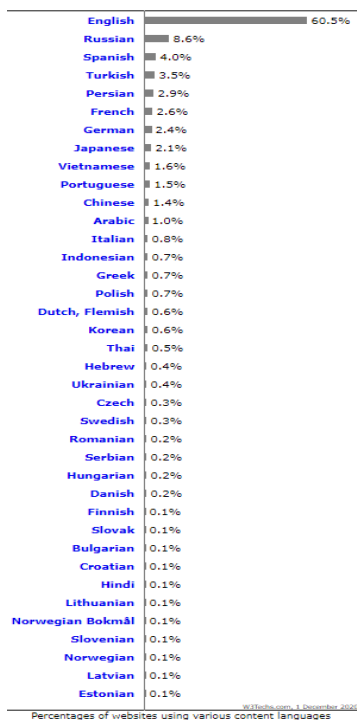


Figure 2, (W3techs.com. n.d. *Usage Statistics And Market Share Of Content Languages For Websites, December 2020*)

Demonstrates the percentage of content in different languages, English being the most used to display content.

Bibliography

PM, Heathcote and RSU Heathcote, 2016,
OCR AS and A-Level Computer Science textbook

Bbc.co.uk. n.d. *What Is The World Wide Web?*. [online] Available at:
<https://www.bbc.co.uk/bitesize/topics/zkccqn39/articles/z2nbgk7>
(Accessed 15th November 2020).

Crystal, D. and Robins, R., 2020. *Language - Language And Culture*. [online] Encyclopedia Britannica. Available at:
<https://www.britannica.com/topic/language/Language-and-culture>
(Accessed 15th November 2020).

Francis, F., Haider, S., Foskett, D. and Estabrook, L., n.d. *Library - The History Of Libraries*. [online] Encyclopedia Britannica. Available at:

<https://www.britannica.com/topic/library/The-history-of-libraries>

(Accessed 15th November 2020).

Google and the world brain. 2013. [film] Directed by B. Lewis. Polar Star Films.

Info Curiosity - Are you curious enough? 2017. *How Many Books Are There On Earth?* - *Info Curiosity*. [online] Available at:

<https://www.infocuriosity.com/how-many-books-are-there-on-earth/#:~:text=A%20book%20is%20also%20a%20literary%20or%20scientific,million%20books%20each%20year%20all%20over%20the%20world.>

(Accessed 15th November 2020).

Leveridge, A., 2008. *Language In Society: The Relationship Between Language And...* / *Ipl.Org*. [online] Ipl.org. Available at:

<https://www.ipl.org/essay/Language-In-Society-The-Relationship-Between-Language-PCJXLSN2AU>

(Accessed 15th November 2020).

McDaniel, S., n.d. *How Far Back Did The Burning Of The Library Of Alexandria Really Set Us?* - *Quora*. [online] Quora.com. Available at:

<https://www.quora.com/How-far-back-did-the-burning-of-the-Library-of-Alexandria-really-set-us>

(Accessed 15th November 2020).

Sitsanis, N., 2020. *Top 10 Languages Used On The Internet For 2020* - *Speakt.Com*.

[online] Speakt.com. Available at: <https://speakt.com/top-10-languages-used-internet/>

(Accessed 15th November 2020).

Techopedia.com. 2017. *What Is The Digital Revolution?* - *Definition From Techopedia*. [online] Available at:

<https://www.techopedia.com/definition/23371/digital-revolution#:~:text=The%20Digital%20Revolution%20refers%20to%20the%20advancement%20of,also%20marks%20the%20beginning%20of%20the%20Information%20Era.>

(Accessed 15th November 2020).

TED-Ed Student Talks, 2019. *How Do Languages And Cultures Interact And Influence Each Other?* | Jennifer Qi | *ACS Athens*. [video] Available at:

<https://www.youtube.com/watch?v=ActWzWajkKc>

(Accessed 10th November 2020).

That Ankh Life. n.d. *The Burning Of Library Of Alexandria: The Crime Of The Age*. [online] Available at:

<https://thatankhlife.com/the-burning-of-library-of-alexandria-the-crime-of-the-age/>

(Accessed 13th November 2020).

Thomas Alopp, *Computer penetration rate among households worldwide 2005-2019, Mar 2nd, 2020*. [online] Available at:

<https://www.statista.com/statistics/748551/worldwide-households-with-computer/>

(Accessed 13th November 2020).

Villa, K., 2014. *How Can We Prevent Cultural Heritage Loss In Latin America? - Latinamericanscience.Org*. [online] LatinAmericanScience.org. Available at: <http://latinamericanscience.org/2014/02/how-can-we-prevent-cultural-heritage-loss-in-latin-america/>

(Accessed 15th November 2020).

W3techs.com. n.d. *Usage Statistics And Market Share Of Content Languages For Websites, December 2020*. [online] Available at: https://w3techs.com/technologies/overview/content_language

(Accessed 15th November 2020).

Wells, H., 1938. *World Brain : Wells, H. G. (Herbert George), 1866-1946 : Free Download, Borrow, And Streaming : Internet Archive*. [online] Internet Archive. Available at: <https://archive.org/details/worldbrain00wells/page/132/mode/2up>

(Accessed 24th November 2020).