

by Włodzimierz Gogołek
Machine (Lack of) Memory

Abstract: The Internet/Web is a storage space for dynamically developing information technology and reflects what is going on in politics, business, education and entertainment. They are no longer passive but take active part in transmission of information between machines and people in a process of effective storage of information. Information technology includes the management of information, its broadcast, reception but also its storage in devices known as computer memory. Its durability is an integral part of the Web, much greater and more long lasting than human memory, but also like human, it is temporary. The results of research presented in this article prove that it is possible to measure this fleetingness. The human memory within one day loses the same amount of information as the Web does in a year. The fleeting of Web information resources, in view of their small size so far (gathered for a dozen years or so) will be a challenge in the future.

Key words: Internet, Web, human memory, Web memory, forgetting, loss of information, digital information resources, threats, Turing's Test

Introduction

Societies have, over the years, coined a functioning model based on the existence of specific hierarchies. This regards organizations, social groups and families. The onset of information technologies, digital information resources and the omnipresent digital communication within society have radically influenced the existing model. The old structure hierarchies have been shaken due to general information abundance, the Web/Internet, computer assisted communication which all the more often overcomes the historical barriers of earlier shaped structures of mutual dependencies. This regards people, small firms, corporations, countries and the entire globe. A growing number of people use the Internet's information potential and easily exchange information using traditional as well as new tools of communication available through the Web. Thanks to growing autonomy of machines communicating with people and their growing capacity for information storage, the computer becomes a partner in the exchange of information and the created virtual world "becomes all the more difficult to separate from the real"¹. As a result, the generally accepted theory that effective communication is reserved for people, homo communicans², has been shaken. Communication is no longer a solely man's distinctive feature.

Modern organizations quickly assimilate this new model of functioning based on communication. On the one hand, aside from traditional information sources, it emphasizes the so called *crowdsourcing*, *smart mobs* or *crowd wisdom*, on the other – it provides fuller knowledge of individuals (personalisation)³. This regards all industries, from education, manufacturing and media to entertainment. The organization's new model of functioning highlights new opportunities based on better cooperation but it also has its drawbacks. The Internet reflects what goes on in

¹ Bronisław Siemieniecki, *Komputery i hipermedia w procesie edukacji dorosłych*, Adam Marszałek Ed., Toruń, 2001, p. 5.

² Bożydar L. J. Karczmarek, *Misterne gry w komunikację*, published by Uniwersytet Marii Curie-Skłodowskiej, Lublin 2005, p. 11.

³ *Wikinomics: How Mass Collaboration Changes Everything (Hardcover)* by Don Tapscott (Author), Anthony D. Williams (Author). Penguin Group, New York, 2007.

politics, business, education and entertainment. It ceases to be a passive transmission medium and becomes more active in the process of information exchange between machines and people as well as in the process of effective information storage.

Determinants

Computers are fundamental links in this technology, determining the functioning of the virtual world, from home PCs, through mobile technology and laptops to supercomputers. And inside them, particularly imperative is their memory capacity.

Each new computer generation is a step closer to simulating what goes on in the real world, making the world of machines more communicative, and this way, effective⁴. There are programmes and devices which combine IT technologies with functions desired by consumers from different backgrounds. As a result, using computers is becoming all the more satisfying, it saves people time and allows them to be more effective in their work. Computers are becoming more user friendly as a result of continuous perfecting and increasing the number of available functions. Computer scientists strive to construct IT tools (equipment and software) in a way that they are intuitive and tolerant of human error.

Machine potential is comprised of computer power but also its compatibility with exterior devices, digital connection and, most of all, proper software. This is the subject of extensive research and experiments, however, not the focus of this article⁵.

Information technology includes the management of information, its broadcast, reception but also its storage in devices known as computer memory (hard discs, floppy discs, flashes, RAM, etc). Putting aside the technologically complicated construction of such memory, what is worth noticing is the continually growing memory capacity. Presently, average PCs contain discs with close to terabyte capacity, there are 50 gigabyte disquettes (blue-ray) and flash memory the size of a tablet with the capacity to store the content of hundreds of books. Thanks to the Web, a large number of computers with their memories is connected creating unheard of before possibilities of gathering and making available vast amounts of information. As a result, created are ever increasing information sources making the virtual world a generally available, very rich information pool.

The usefulness of information stored in computer memory is measured by its speed (the time it takes from the saving/reading the information to its realisation), durability and the possibility (due to web communication) of remote and selective reading.

Putting aside the technological aspects of using digital memory potential, accessibility and speed of information processing, long term storage of data is an imperative and, yet, frequently overlooked feature of "Web memory". Research shows that there are certain similarities between web and human memory, one of which is memory loss.

⁴ Presently (June 2008), the largest supercomputer in the world is the Roadrunner, with one petaflop capacity, the equivalent of 116 640 modern processors. This computer can make as many calculations in one day as all the people in the world equipped with state of the art calculators would make in 46 years, working 24 a day, 7 days a week. From: JOHN MARKOFF, Military Supercomputer Sets Record, The New York Times, June 9, 2008,

⁵ Some of these uses are presented by: Włodzimierz Gogolek, Wprowadzenie do Informatyki dla Humanistów, Defini, Warszawa 2007.

Threats

Memory of artificial, computer devices is a lot more durable than human memory (theoretically computers do not forget!). This will be more visible in the future when the following generations will be able to use digital data compounded over nearly half a century. Presently, data is solely published and stored in digital form which may eventually lead to the so far ignored problem of information archiving. Digital data is saved on various devices (tapes, floppy discs, hard discs), however their durability over a long period of time is questionable. Taking into consideration the fact that less and less information is stored the traditional way, on paper, there is a real threat that information archived digitally will not be available or will cease to exist after some time, ie. several hundred years.

In comparison, thanks to cuneiform writings which were created in IV BC, papyrus - 2500 BC, *kipu* string (Inka writing) and later paper, information endured hundreds and thousands of years. This cannot be guaranteed by the present digital devices, except for discs made of precious metal, as was the case with the disc with information sent out into Cosmos for extraterrestrial creatures⁶.

Another problem regarding digital form of information archiving is the fact that it is continuously undergoing improvement which results in changes of standards and technologies for storing and reading the saved data. This is true for all forms of multimedia, from text to video materials, and there is history to prove it. In the past there were 5 inch discs which nobody remembers anymore, then 3.5 inch ones also quickly going out of style, there are disappearing CDROMs and probably soon the same will happen to DVDs. A proof of the reality of this threat is a BBC initiative from 1986 with a great amount of data on Great Britain, including maps, video materials and texts, stored on special laser discs compatible with, modern at the time, BBC Micro computers. Presently, these discs cannot be used on any computer. A great amount of valuable data was almost lost. After two years of extensive work and a lot of money spent, the data was eventually transferred to modern PCs (WWW.domesday1986.com)⁷. The same situation is with word processors and old search engines. It may not even be a period of ten years and data saved in one form will not be available on more modern versions of the same programme. An example of this are documents saved several years ago on Polish word processors such as Chi Writer or QL Tekst. Outdated, obsolete equipment and special software is necessary in order to open and read documents saved using these devices.

As a result, there seems to be a necessity to create standard saving formats for all forms of multimedia information. Such standards should be universal, precise, not change over time, and be user friendly. A beginning of work in this sphere is the use of virtual computers, the equivalent of older generation computers, simulated by specialist software. An example of such solution is the Universal Virtual Computer (UVC), with software able to decode documents saved in older formats⁸. However, there still remains the unsolved issue of durability and availability of devices for reading (ie. now the majority of modern PCs and laptops do not include an A disc drive).

⁶ Pioneer 11 is heading in the direction of the Eagle constellation. It will pass near the closest star of the constellation in 4 million years. From http://en.wikipedia.org/wiki/Pioneer_plaque

⁷ A New way to stop digital decay, *The Economist*, Sept. 17th, 2005.

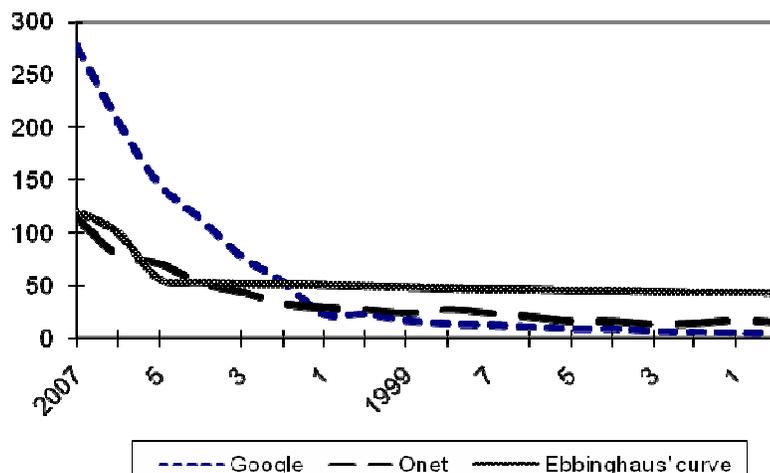
⁸ http://www.kb.nl/hrd/dd/dd_onderzoek/uvc_voor_images-en.html [November 2005].

Human and machine memory

Machine memory fleetingness, being similar to human, results in the fact that some of the information stored on the Web will be lost. Research carried out shows that the forgetting curve, or the loss of information curve, in computer memory is comparable to the classic human memory loss curve (Ebbinghaus' curve)⁹.

The intensiveness (time and amount) of information loss in humans and on the Web can be observed as shown on the chart, 1 day for people is the same as 1 year on the Web. Assuming the above, the Pearson correlation coefficient on Ebbinghaus' curve of time and information storage on the Web, based on the search engine Google example, is 0.93. The same is regarding data from the Onet.pl search engine, correlation value 0.91. These coefficients prove the existence of a line correlation between human and machine memory, comparing one day of human memory and one year of machine (Web) memory.

Chart 1. Human and Web storage of information on Ebbinghaus' curve



Legend: In order for the chart to be more legible, the information amount value (percent of saved syllables) on the Ebbinghaus curve was doubled. The X axis for this curve is a scale in days for humans, for Google and Onet, in years.

Source: by author¹⁰.

Information is stored on the Web for a longer amount of time as it is possible to go back and retrieve earlier saved information. Materials generated on it (e-mail, discussion forum posts) can remain on the Web for many years creating new

⁹ Philip G. Zimbardo, *Psychologia i życie*, PWN, Warszawa 1999, s. 356., Strelau, s. 147.

¹⁰ Research was carried out using data from two search engines, Google and Onet. To measure time durability for storing information on the Web chosen was a function of number of links indicating pages with different years (from 1990 to 2007) as well as pages with dates of publication for the years indicated above, chosen randomly. Research was carried out in April 2008.

challenges for communication understood traditionally¹¹. This way, correspondence or opinions presented cease to be fleeting¹². In the past they could have been forgotten or lost, as happens in case of books, magazines and other traditional information carriers. Modern search engines and other Internet services have created their own Web archives, regardless the will of their creators and have been able to process and use this information since the early 1990s, practically without any time or geographic limitations. An example of this is the Internet Archive which has been storing webpage data for a dozen years (so far it has an archive of 85 billion sites).

This results in an unprecedented earlier responsibility for all information broadcasts which the people have left on the Web. Opinions and commentaries written 15 years ago (often signed by people's real names) are available and can be accessed today, possibly often causing consternation.

On the other hand, the capabilities of storing digital information seem to be inadequate as they involve too short of a period of time. Today, for many people the Web is a basic source of information, however the early materials (text, image, audio, video) stored there, dating back to the early 1990s, may be very difficult to retrieve. Not everything can be made available via regular PCs at the click of a mouse. This is the case with a lot of research materials created in the early 1990s. This problem is currently worked on and the situation is slowly improving. More old data is inserted onto the Web and archived. There is a gradual process of digitally storing older and older materials.

Conclusion

The issues mentioned above indicate that the Web can be viewed as a creation with some characteristics similar to that of human capabilities. This is especially true regarding communication as well as the ability to gather and store information. Technically, machines possess much better memory than humans, however, like human, computer memory is also fleeting. Beside communication and memory, there is one more similarity – intelligence. So far, however, all attempts of computers to pass the Turing Test have failed¹³. Nevertheless, the results of subsequent competitions (Loebner Prize) in this area are promising. Last year's winner, the Ultra Hal Assistant programme, communicating using a natural language, can successfully carry out many office tasks¹⁴.

¹¹Theoretically, machines creating Google archives can erase compromising entries. However, this procedure is complex and it does not offer a complete guarantee that the information will definitely cease to exist.

¹²Materials from 10 years ago until present day are available at <http://www.archive.org/index.php>

¹³ His test determines whether a machine can possess artificial intelligence. It is based on a conversation of a judge with a human and with a machine at the same time. If the judge is not able to differentiate who is who, it means the machine has passed the test.

¹⁴ <http://www.zabaware.com/assistant/> [July 2008].