JLIS.it

Encyclopedias and encyclopedism in the era of the Web*

Gino Roncaglia[(a)]

a) Università Roma Tre, http://orcid.org/0000-0003-4632-3342

—————

—————

ABSTRACT

The paper traces the development of digital encyclopedias through four main stages: 1. The age of the first text-based online services; 2. the age of multimedia CD-ROMs, mainly based on the idea of an offline multimedia encyclopedia; 3. the first generation of web encyclopedias, when the web version initially supplemented and then thoroughly replaced CD-ROMs; 4. the age of data oriented, semantic-aware encyclopedias. This chronological framework is used to discuss how different models of encyclopedias and encyclopedism – including the Wikipedia model – have been intertwined with technological developments on the one side, and with the cultural debate on the new digital ecosystem (and on its role for the publishing industry) on the other.

KEYWORDS

CITATION

* This paper is the English version, substantially revised, enhanced and updated, of a paper I presented at the international conference "Enciclopedismo antico e moderno" held in Bari on April 17th-18th, 2017. The original Italian version of the paper was published in the proceedings of the conference (Roncaglia 2020).

—————

JLIS.it

1. Preliminary remarks

The new digital ecosystem and the rise of the world wide web profoundly changed the nature of contemporary encyclopedias and encyclopedism. Even before the 'Wikipedia era' (which will be extensively outlined later in this paper) all the major publishers operating in the field had abandoned the idea of the utmost centrality of the traditional printed form, replacing or supplementing the latter with online versions. From such a perspective, encyclopedias differ from many other types of editorial products and specifically from narrative and generalist non-fiction books, which remain more closely related – at least for the time being – to the printed format despite the development of electronic books.

The reasons for this development are many: on the one hand, *reference* works can often be considered as database-oriented and intended for occasional consultation rather than for linear reading. So, they are less strictly dependent on the features suggested or required by situations of protracted reading, where paper appears to still retain some advantages compared to digital reading devices. On the contrary, the digital reading environment offers many advantages if we wish to access the relevant content occasionally and in a nonlinear manner. Moreover, online content can be easily modified, and therefore easily updated; digital encyclopedias allow us to add audio and video clips to textual information and images, as well as for cross-references from one entry to another in real time, and for the use of interactive tools (the visualization of mathematical functions in Wolfram-Alpha, a science-oriented online encyclopedia[1], is a good example of this feature). Last but not least, the price factor must be considered: in general, printed encyclopedias have a bulky format and are very expensive, and the marginal cost of a new copy is high; on the contrary, in the case of a digital encyclopedia a new copy has virtually no marginal cost, and the space required for its storage is either very small (if we use a physical support) or inexistant (if the encyclopedia is online).

The process of development of digital encyclopedias, however, took quite a long time and went through several stages, closely connected with the evolution of the digital ecosystem. In my opinion, four main stages of this process can be identified:

- The age of the first text-based online services such as Minitel, Prestel, CompuServe: as we will see, some of them started providing access to online, text-based encyclopedias at the beginning of the 1980s.
- The age of multimedia CD-ROMs, mainly based on the idea of an offline multimedia encyclopedia: it broadly corresponds to the decade elapsing between the end of the 1980s and the end of the 1990s, and reaches its fundamental point in the 'war' between *Encyclopaedia Britannica* and *Encarta*, the encyclopedia produced by Microsoft, which is also the most representative example of the transition from the model of paper encyclopedias to new multimedia encyclopedias.
- The first generation of web encyclopedias, when the web version initially supplemented and then thoroughly replaced CD-ROMs: it is the period of time elapsing more or less between

---

[1] The web site is https://www.wolframalpha.com/. All the web sites and resources mentioned in this paper were last accessed on April 2, 2021. Unless otherwise stated, Wikipedia quotations and references are taken from the English version.

JLIS.it

the turn of the last millennium and the first decade of the current century, and its key product is undoubtedly Wikipedia.

-   The age of data oriented, semantic-aware encyclopedias: it is the age we are currently living in and is linked to the developments of the semantic web; its main feature is the progressive formalization of the idea of highly structured encyclopedic knowledge bases, which rely on ontologies organized in the form of linked data, and on a layer of 'intelligent' software agents used for data retrieval. Some of these agents can respond in an interactive way to questions formulated in a natural language: it is the path that leads – among the others – to the so-called 'voice assistants', such as Google Assistant, Amazon Alexa, Siri, Bixby, Cortana.

We shall see that this evolution is not always linear, and different models of encyclopedias and encyclopedism have been, in several respects, intertwined with technological developments. An example is the Institute of the Italian Encyclopedia 'Treccani', one of the few publishers which continues to sell printed encyclopedias, while enabling access to most content by means of a popular web portal. The brand Treccani retains undeniable authority even in the era of Wikipedia, while *Encyclopaedia Britannica* has not been released in the printed version for a decade and its online version barely survives today.

Despite these differences, I think that distinguishing the four stages as identified above retains its validity, and the in-depth analysis of each of them will be at the core of this paper. In this regard, however, it will often be required to proceed by following specific criteria other than a mere chronological framework.

2. The first text-based online services

The beginning of the age of digital encyclopedias[2] is often identified with the first digital edition of the *Academic American Encyclopedia*, published on laserdisc by Arete Publishing in 1984-85 (purchased by Grolier – a US publisher specialized in encyclopedias[3] – in the same years).

However, the story is more complex and provides us with a good example of nonlinearity in the use of different digital technologies. As a matter of fact, we can travel back in time more than 25 years, to the age of mainframe computers. Possibly the first, little-known example of digital encyclopedia is the "electronic Larousse" which was exhibited at the 1958 Bruxelles World's Fair, running on a Control Data 3600 computer with magnetic-core memory (Mollier et Dubot 2012, 537–8, quoted in Loveland 2019, 358–359): the official guide of the fair describes it as follows: "Larousse mettait l'univers des connaissances au service de l'homme. […] Un dictionnaire électronique «répondait» aux questions qui lui étaient posées sur les sujets les plus varies".[4]

---

[2] A useful reference resource on the encyclopedias produced in this period is Kister 1994.

[3] Grolier was later bought by France's Hachette in 1988 and resold to Scholastic in 2000.

[4] *Guide* 1958, quoted in https://www.worldfairs.info/expopavillondetails.php?expo_id=14&pavillon_id=168. Even if not encyclopedic in nature, the still earlier 'informateur électronique' also deserves to be mentioned: it was built in Paris in 1954 by Albert Ducrocq (a fascinating and almost unknown pioneer of cybernetics; he also invented, as early as 1952, a text and poetry-generating machine called 'Calliope'). The 'informateur électronique' was a transistor-based electronic reference tool which, in the words of his inventor, had the goal "d'instruire les jeunes des carrières s'offrant à eux selon leurs goûts, leurs

JLIS.it

A mainframe-based encyclopedia was however of little use without the ability to access it remotely. And for this step we must wait till 1980, when the Ohio College Library Center[5] announced the launch of a new revolutionary project:

> A computer-based home information service that turns an ordinary television set and telephone into a home bank service, an encyclopedia, a library catalog, and a community information source, will be test-marketed in 200 homes in the Columbus area beginning in October. Called Channel 2000, the home information service is being developed and market-tested by OCLC, Inc., and Bane One Corporation. The system uses an ordinary television set, a telephone, and a special adapter unit designed by OCLC, Inc. (*LITA Newsletter* 1980, 4).

Among the contents delivered essentially by an embryonic system of remote access to databases, somehow similar to the French Minitel (which was introduced between 1980 and 1982 and allowed access to a few thematic encyclopedias[6] by 1986) was the *Academic American Encyclopedia*. As it was usual at that time, access took place via a textual interface, and this explains why the encyclopedia was in a text-only version, without multimedia integration, a feature which will be also maintained in the laserdisc version.

The OCLC service seems to be the first of this kind – and the first to be aimed to a general audience accessing an encyclopedia from home[7]. In the following year, a similar service was offered by *Encyclopaedia Britannica* through LexisNexis, which however was "a closed, subscription-based network that allowed legal professionals, among others, to research documents electronically" (Loveland 2019, 369). In 1982, it is the CompuServe Information Service, one of the very first dial-up online service, to allow on-line access to the *World Book Encyclopedia*; in the same year, access to the Academic American Encyclopedia was offered by Dow Jones News and Retrieval. In 1983, CompuServe access to the *World Book Encyclopedia* was replaced by the access to the *Grolier Encyclopedia* (Mailland 2017). In the same year, the *Academic American Encyclopedia* was made available by the Bibliographic Research Service and by the British videotext service Prestel (Loveland 2019, 396; Rice 1985, 418–430). In a few years, the *Academic American Encyclopedia* wins the very first battle among on-line encyclopedias, being offered by "almost all the commercial dial-up information services, including CompuServe, Delphi and Prodigy" (Loveland 2019, 396, with reference to Flagg 1983, 134–6).

---

aptitudes, leurs résultats scolaires, leur preference pour des voyages ou pour un travail sédentaire, leur désir de travailler seuls ou en équipe" (Ducrocq 1993, 29). An image of this curious machine, from the collection Roger Viollet, is available online: https://www.roger-viollet.fr/image-photo/informateur-electronique-paris-le-1er-mars-1954-roger-viollet-roger-viollet-433767.

[5] OCLC: the meaning of the acronym would subsequently change into Online Computer Library Center, and the organization would become one of the main reference libraries in the world, responsible for managing WorldCat, the largest online catalogue at international level, as well as the Dewey decimal classification system.

[6] Cf. Menning 1986. For a comparison between French Minitel and the US dial-up services available at the time cf. Mailland 2017.

[7] The otherwise useful Loveland 2019 does not mention this experience and wrongly states that "The first company to put an encyclopedia online was the Encyclopaedia Britannica" (*Ivi*, 369).

JLIS.it

The first characteristics of the digital ecosystem to be exploited in this kind of text-only, dial-up online services were full-text search, as well as remote and nonlinear access to contents. However, these early examples of digital encyclopedias were still based on content drawn from printed encyclopedias: the development of a new, born-digital encyclopedism will have to wait until multimedia content is added to the equation.

3. Multimedia encyclopedias on CD-ROM

We already mentioned the first, laserdisc-based digital edition of the *Academic American Encyclopedia* published in 1984-85. Already in 1985, the digital version was less expensive than the paper version: the printed version in 21 volumes was priced at 650 dollars, while the one on laserdisc at 89.95. Even considering the cost of the laserdisc player (approximately 450 dollars) which was clearly suitable for other purposes as well, the price balance turned out to be positive from the very beginning (cf. Smith 2017).

In the following years, Grolier would publish a series of new editions of the encyclopedia under different names: *The Electronic Encyclopedia* (1986), *The Grolier Electronic Encyclopedia* (1987), *The New Grolier Electronic Encyclopedia* (1988-91), *The New Grolier Multimedia Encyclopedia* (1992).[8] The 1987 edition moved from laserdisc to an early CD-ROM version, the so-called CD-32, available on Atari computers; in later versions the support of choice was a standard CD-ROM (Bozzi 1996, 224).

While increasingly sophisticated indexes improved access speed – the text required 60 Mbytes and the indexes 48 Mbytes in the 1987 version of the encyclopedia, thus enabling to retrieve information about over 9 million indexed terms in less than three seconds (*Ibid.*) – the content initially remained merely textual. We have to wait until 1989[9] to truly achieve a multimedia encyclopedia, when *Compton's Multimedia Encyclopedia* was published.

Compton's Encyclopedia originated in 1922 as *Compton's Pictured Encyclopedia*, and traditionally focused on the visual element, also related to the fact that it mainly aimed to families and schools. In 1989, Compton was owned by the Encyclopaedia Britannica, Inc. (which held the majority of shares from 1961 to 1993, and then, again from 2002 onward) so, the 1989 CD-ROM can also be considered the first step taken by *Britannica* in the direction of CD-ROM based encyclopedias.[10]

Throughout the 1990s, encyclopedias on CD-ROM multiplied, thus bringing all the major publishers of the sector into the new field. In 1991 *Compton's Encyclopedia* was distributed for both DOS and Mac, and in 1992 Grolier also produced its own multimedia encyclopedia, *The New Grolier Multimedia Encyclopedia*, which was included for free when purchasing a Mac computer thanks to an

---

[8] *Wikipedia, The Free Encyclopedia*, s.v. "Grolier" (accessed April 2, 2021), https://en.wikipedia.org/wiki/Grolier.

[9] As a matter of fact, Italy could perhaps boast a record in this respect: "De Italia", a sort of encyclopedic review on Italian civilization produced on an interactive videodisc by the Agnelli Foundation, dates to 1987. However, it was not a generalist encyclopedia rather an ancestor of interactive electronic books. For a presentation of the work see "Un videodisco-laser per raccontare un paese" 1987.

[10] For the very interesting events concerning the relation between *Encyclopaedia Britannica* and the digital world, see Greenstein and Devereux 2006, and Greenstein 2017. Much of the following information about the "war of encyclopedias" in the 1990s is drawn from these articles. A detailed overview – with useful information on the European markets – is also offered by Loveland 2019, chapter 10.

JLIS.it

agreement with Apple. In 1995, both *Encyclopedia Americana* (again by Grolier, the text and multimedia content were stored in two different CD-ROMs) and *Hachette multimedia Encyclopedia* ("EHM" – at that time, Grolier was owned by Hachette, so, the two products could take advantage from a common technical infrastructure) were published in a CD-ROM version. During the same years, many other publishers entered the new digital market. In Italy, De Agostini's *Omnia* multimedia encyclopedia was one of the most successful: first published in 1997, it was adapted for other countries (Poland, Germany, and Greece) within a few years, and was produced in a version for children (*Omnia Junior*) and a sectorial version (*Omnia Art*). Moreover, again in 1997, the Institute of the Italian Encyclopedia came into play by distributing a CD-ROM incorporating the second edition of the *Vocabolario Treccani*.

The encyclopedias taken into consideration so far were mainly adaptations of printed versions or produced by publishers already active in the field. However, by the mid-1990s, also new players entered the sector, namely companies who were not operating in the field of encyclopedic publishing on paper, rather in new media. We shall see that this change substantially impacted not only the market, but also the features and format of encyclopedias. Publishers of traditional encyclopedias were naturally interested in preventing the new medium from expanding on the market at the detriment of the old medium (and of the related sales network, mostly still carried out by door-to-door sellers) and therefore were inclined to maintain relatively high prices, or to exclusively sell the CD-ROM version jointly with the paper version. *Compton's Encyclopedia* distributed by *Britannica* is a good example of this approach: it was free for those who bought the paper version of the encyclopedia, but the price tag amounted to 895 dollars when it was purchased separately. And when *Britannica* was finally released on CD-ROM in 1994, the fear of undermining the traditional sales network caused that the latter was sold only jointly with the paper encyclopedia at the price of 1,200 dollars, as it had been initially – and ineffectively – opted for.

The new products based on digital media could instead enter the market following thoroughly different strategies: a key example is Microsoft, which published the first version of its extremely successful CD-ROM encyclopedia *Encarta* in 1993 (partly based on *Funk & Wagnalls Encyclopedia*, of which Microsoft had acquired non-exclusive reuse rights, and *Collier's Encyclopedia* and *New Merit Scholar* afterwards, acquired by Microsoft from Macmillan Publishers). In 1985, Microsoft also tried to acquire non-exclusive rights on the entries in *Encyclopaedia Britannica*, proposing to seal a partnership, which Britannica refused. In 1989, *Britannica* and Microsoft were on good terms, and Microsoft was one of the launch partners of *Compton's Encyclopedia*.

When Microsoft launched its own encyclopedia on the market, however, the fragile balance between paper and digital was definitively broken. The first edition of *Encarta* sold only 10,000 copies: the price was still affected by the strategies brought forward by traditional publishers and was too high ($ 295), while the need to postpone the release date from September to March due to technical issues did not allow to possibly exploit the driving force of Christmas shopping. But the following edition, just a few months later, corrected these errors: it was released in October, at a price tag of only $ 99. Success was immediately achieved: the new *Encarta* sold 120,000 copies during the holiday season, 350,000 copies in the first year, one million copies over two years. Since the physical cost of a single CD-ROM amounted only to a few dollars, it is easy to understand that such a considerable increase in sales promptly impacted the profitability of the publishing operation:

The organization grew into a unique business. Similar to an encyclopedia, the variable costs were low while the fixed costs were high. However, the level of those costs supported a very different price point for the end product because the low price point supported high volumes. The cost of materials for the CD and packaging was quite low, sometimes $ 5 a box. The fixed expenses arose from editorial staff, programmers, and additional staff needed to update the content, as well as to tailor the product to different languages and countries. Much like in traditional encyclopedia businesses, therefore, the business strategy became oriented toward generating sufficient sales volume to support revenue that covered the fixed costs of the operation — again, with one key difference, the unit volumes were in the millions instead of the low hundreds of thousands (Greenstein 2017, 1009).

Another important factor to consider is the constant updating of *Encarta*, especially through substantial improvement in the multimedia content: a possibility opened by a production cycle which became much faster and cheaper once the costs for the basic editorial content had been incurred. Of course, advertising of new digital encyclopedias strongly emphasized this element, which became even more central, levering on the model of hybrid encyclopedias – partly on the web, partly on CD-ROM – which appeared in the mid-1990s and which in a few years, would lead to publish encyclopedias exclusively online.

In order to face the challenge represented by *Encarta*, *Encyclopaedia Britannica* was forced into an impossible chase: the attempt was to lower the prices (to $ 995 in 1995, to $ 200 in 1996) but it was too late; between 1990 and 1995 its sales fell by 50%. The last, bold but desperate attempt was to launch an exclusively online edition: in this case, *Britannica* anticipated *Encarta*, which preferred to undertake the path of the hybrid version; but in 1995, when *Britannica Online* was launched (with a price tag of 150 euro per year, plus 25 € for the initial registration) people connecting to the Internet from home were not many and mainly used dial-up based services; furthermore, those few who were lucky to be able to connect quite frequently, already had *Encarta* on CD-ROM. The operation did not bring the desired results, liabilities increased, and in 1996 the brand and resources of *Encyclopaedia Britannica* were sold for 135 million dollars to a private investor, Jacqui Safra. Compared to just a few years earlier, the company's value had collapsed and could almost exclusively be related to the valuation of the brand.[11]

In a nutshell, contrary to common belief *Encarta* disrupted the market of traditional paper encyclopedias and not Wikipedia, which was established subsequently and, if anything, was at the root of the crisis regarding the model of commercial digital encyclopedias ten years later.

As a matter of fact, *Encarta* continued to sell quite well until 2003-2004, on CD-ROM first, in the hybrid version afterwards, and finally online from 2000 onwards; in those years, editions in eight different languages, including Italian, were published, together with a specific version for children (*Encarta Kids*). But gradually, accesses began to decline in favour of Wikipedia. Microsoft would abandon the project in 2009. In January, 97% of accesses to the online encyclopedias from the United

---

[11] The new owners tried to revive it, but with little success: the last printed edition was published in 2010 and today, *Britannica* is available in the online version only, subscription has a cost of about $ 69 a year. Since 2009, *Britannica* has also tried to adopt a model that is closer to Wikipedia, with the possibility for users to integrate information within the entries (after editorial revision); but also in this case, the initiative reached little success.

JLIS.it

States were related to Wikipedia, while the percentage of accesses to *Encarta* reduced to 1.27% (and to *Britannica Online* to 0.57%) (Hopkins 2009; cf. Cohen 2009).

While *Encarta* undoubtedly represented the most successful example of multimedia encyclopedia produced by a company focusing on digital publishing and with no previous experience in traditional publishing, among the 'digital-native' encyclopedia projects one relevant Italian example deserves mentioning: *Encyclomedia*, the encyclopedia directed by Umberto Eco and Danco Singer. *Encyclomedia* was conceived as a comprehensive, CD-ROM-based multimedia history of the European civilization. The project started in 1993 with the establishment of Horizons Unlimited, a company founded by a group of Eco's students from Bologna; the first two CD-ROMs dedicated to the sixteenth and seventeenth century were published with a fairly high price tag (399,000 Italian lira each, equivalent more or less to 235 dollars) between 1995 and 1996 but with advanced features for the time. The CD-ROM on the seventeenth century included about 9,000 voices, 2,000 images, 165 audio clips, 32 animations and 13 videos, as well as 203 ebooks (Floridi 1996, 193). In the following years, the work was completed and delivered in a variety of formats, on CD-ROM and in print (as collateral at newsagents) up to the online version.

4. Online encyclopedias and the emergence of Wikipedia

As mentioned above, the transition from multimedia encyclopedias on CD-ROM to the online versions is a long process: we saw how the very first digital encyclopedias – emerged when CD-ROMs did not yet exist – were conceived for use over a network, and already during the very first years, after the creation of the World Wide Web, many multimedia encyclopedias embarked on the path of hybridization between CD-ROM (and DVD afterwards) and the Internet.

With the web now available, the market for exclusively online encyclopedias opened: if the attempt made by *Britannica* in 1995 was too bold given the number of families who had an Internet connection at that time, the transformation of *Encarta* into an online encyclopedia in 2000 mirrors a situation where the web was the natural horizon for all information searches, and therefore also for the evolution of encyclopedism. But the web changed not only the tools for accessing information, but it also radically modified the ways and forms information content was produced. And this change involved the world of encyclopedias quite soon. Wikipedia, which was founded in 2001, is the best example of this evolution. But already in October 1993, Rick Gates, one of the pioneers in the use of the Internet, suggested the idea of an online encyclopedia built collaboratively by using the potential of the network. This early project had a name, *Interpedia*.[12]

However, we must wait seven more years, until March 2000, for Jimmy 'Jimbo' Wales and Larry Sanger, who brought to light the first experiment of truly free online encyclopedia, *Nupedia*. Unlike what was going to happen with Wikipedia,[13] Nupedia was not opened to public changes and still relied on expert editors, who however agreed to collaborate to the project for free and to freely distribute the entries created, subjecting them to a public peer review process, always among experts from the sector. The main connection between the initial Nupedia project and Wikipedia lies in the

---

[12] On the first ideas of an online encyclopedia, see Reagle and Lessig 2012, chap. 2 (on *Interpedia*, 32–34).

[13] For an analysis of the first projects of the online encyclopedia and the process that leads to Wikipedia, cf. Mako Hill 2013.

idea of a free and open encyclopedia: an idea that was also endorsed by Richard Stallman, the guru of free software, who soon decided to merge into Nupedia his own project for an open online encyclopedia, started in 1999 and called *GNUpedia*.[14]

However, Nupedia's project did not achieve a positive outcome: relatively few entries were created, and the name of Nupedia today is known, above all, because Wikipedia span off from it. It is in fact on Nupedia's server that, on January 10th, 2001, the Wiki software was installed to become Wikipedia five days later, on January 15th, 2001. Of the two founders, Jimbo Wales was an entrepreneur who at the time of the launch of Nupedia first, and of Wikipedia afterwards, ran a search portal for pop music called *Bomis*. Larry Sanger, on the other hand, had been chosen by Wales as the managing editor of Numedia. In the following years, Sanger left Wikipedia, considering the results of its editorial process[15] to be biased, and founded two new projects: *Citizendium*,[16] an online encyclopedia project attempting to merge the advice provided by expert editors and open peer collaboration, and *Everipedia*,[17] an open encyclopedia using blockchain technology to trace the editorial history of entries. Between 2019 and 2020, he left both projects, and established the Knowledge Standards Foundation and a new website, *Encyclosphere*, which aimed to connect within a decentralized network all the web-based encyclopedia projects.[18] Sangers' peregrination through different encyclopedic projects (and into different ideas of what an open online encyclopedia should and could be) is rather interesting, even though often inspired or affected by his own right-wing political ideas.

But let us return to Wikipedia. What exactly does 'Wikipedia' mean? The initial 'Wiki' refers to the software used to manage the website and its editorial process, while the suffix 'pedia' – derived from the Greek παιδεία – is a familiar word and refers to the encyclopedic nature of the project. The term Wiki derives from the Hawaiian language, where 'wikiwiki' means 'very fast' – the Honolulu airport shuttle bus was called 'Wiki - Wiki'. Ward Cunningham, the creator of the first Wiki system, decided to call his software with such a name to emphasize that it was easy and immediate to use.[19] Basically, the term 'Wiki' indicates a program which resides on a server and allows to create or modify pages and websites quickly and easily by writing directly into the browser.

But the peculiarity of the Wiki software – which is what distinguishes it from other websites management systems – is the ability to keep track of all changes made by each single user of the system, allowing to reconstruct the editorial history of each entry and to go back to a previous version of the page or text, when needed. So, the Wiki system is a perfect tool for writing pages and sites in a collaborative way; it has to be stressed that the term Wiki is not just used to refer to the editorial tool used, but also to the websites that are created by using that software. Wikipedia is therefore an encyclopedia created by using a Wiki software and is an example of a Wiki website itself.

There are different types of Wiki software: as mentioned above, the original program was created by Ward Cunningham in 1995, but today the most used program to create Wiki sites – including Wikipedia – is called *MediaWiki*. It is an open-source software application, free and freely modifiable,

---

[14] This part of the story is well reconstructed in the second chapter of Reagle and Lessig 2012.

[15] Cf. Foggin 2019. Sanger also accused Wikipedia of left-wing political prejudice: Cf. Flood 2020.

[16] The website is https://en.citizendium.org/.

[17] The website is https://everipedia.org/.

[18] Cf. https://encyclosphere.org/.

[19] For a general introduction to the main ideas behind the Wiki software, the reference text is Leuf and Cunningham 2001.

and was specifically developed for Wikipedia in 2002, although it is currently adopted by many other web platforms.[20] Its development is managed by the Wikimedia Foundation, a non-profit foundation created in the United States with branches located in a number of other countries, including Italy; its purpose is to take care of Wikipedia and the several related projects that were developed around or in connection with the original one, also by helping raise the necessary funding.[21]

The Wiki platform used by Wikipedia[22] allows users to collaborate in the drafting process of the encyclopedia entries, and to keep track of the history of all the changes made to each entry. And this is exactly the key characteristic that makes Wikipedia different to traditional encyclopedias. In a traditional encyclopedia, authority i.e. the prestige and reliability of the encyclopedia, is a consequence of authorship i.e. the fact that entries are written by acknowledged experts on the matter. In many cases, entries – or at least the main ones – are signed by the authors, so to make them recognizable. And even when they are not signed, the editorial staff, who is composed of a restricted number of professionals, guarantees authority and validation of the content.

However, in the case of Wikipedia, authority does not derive from authorship since the entries are written in collaboration and not drafted by one identifiable author. On the contrary, authority and content validation are the result of a collaborative writing and review process.

But how can such an encyclopedia be reliable given that it is written by no identifiable authors, where anyone can correct and modify the entries as they wish? Many doubts about the actual reliability of Wikipedia derive from this (more than legitimate) question.

To provide an answer, we shall start from a more basic question: is Wikipedia *really* reliable?

Surprisingly, with specific reference to the broadest version of Wikipedia, the version with the highest number of entries and contributors – of course, the English one – the answer is reasonably positive. Sixteen years ago, in 2005, when Wikipedia was still a newly established undertaking, the prestigious scientific journal *Nature* published a research based on a sample of 42 scientific entries taken from both Wikipedia and *Encyclopaedia Britannica* (clearly considered to be the paradigm of traditional and validated encyclopedia). The selected entries were examined 'blindly' by a group of experts who did not know which version of the entry was taken either from Wikipedia or *Encyclopaedia Britannica*, with the result that the divergence between the two sources in terms of quality and reliability was minimal (although *Britannica* retained a slight advantage).[23] An endless debate was raised about this outcome,[24] and in the following years, a number of similar comparisons and reliability analysis on Wikipedia were carried out.[25]

The resulting debate is complex and interesting, but I will not analyze into detail, not to go beyond the scope of this paper; I will rather try to provide a concise synthesis of its conclusions, with all the

---

[20] For an introduction to MediaWiki see Barrett 2009 and Koren 2017.

[21] The website is https://wikimediafoundation.org/.

[22] Among the many publications on the nature and features of Wikipedia, a pioneering text of special relevance from the point of view of the historical reconstruction is Lih 2009.

[23] See Giles 2005 (at the same web page of the paper, it is possible to find also a link to the objections raised by the editorial staff of *Encyclopaedia Britannica,* and the related replies by *Nature*).

[24] Far less favourable for Wikipedia were the results of the analysis carried out by Holman Rector 2008. Of course, the data on which the paper was based thirteen years ago is now to be considered quite old.

[25] For the interest raised and given that it is a more recent paper with extensive bibliographic reference, an example worth of mention is Greenstein and Zhu 2018.

limitations it may imply. The overview is the following: we cannot be entirely certain of the accuracy of each piece of information drawn from Wikipedia (this also applies to a traditional encyclopedia, but the degree of potential inaccuracy of the Wikipedia content is higher). Overall, however, most information delivered by Wikipedia is reasonably reliable, and the reliability rate grows according to the quantity of editorial interventions on the voice and the number of users involved in its drafting and revision.[26] The fact that Wikipedia has a dynamic character and is constantly updated, makes it possible that errors are entered at any time, but thanks to this dynamic character, errors are usually corrected promptly. Furthermore, Wikipedia coverage is now much broader than any traditional encyclopedia, specifically as far as some sectors are concerned (including those related to the so-called 'popular culture', namely television series, movies, comics…). So, it has a coverage that is 'eccentric' in many ways, but way more extensive than traditional encyclopedic models.

However, the original question remains to be answered: how comes that an encyclopedia where anyone can contribute in any possible way results to be reasonably reliable and correct?

The answer paradoxically lies in the large number of contributors to Wikipedia, associated with the ability of the Wiki platform to keep track of the revision history of each entry. Of course, there are the so-called 'vandals', or 'trolls', who deliberately try to enter misinformation. But most Wikipedia users are mostly well-intentioned; moreover, the system allows to immediately identify the most recent changes to each entry, and to correct them either by reverting to the previous version (*revert* procedure) or by modifying the text, when needed; furthermore, the network address of the computer from which incorrect and misleading content was added can be blocked, while an entry can be temporarily 'locked' if necessary, thus preventing further modifications for a given period of time: a feature that can be used when controversial entries are subject to trolling or to a biased attack.

This also explains why the English version of Wikipedia is on average more complete and reliable than those in other languages. Within a tool whose authority mainly depends on the mechanism of collaborative writing and review, the greater the number of collaborators, the higher the probability to have among them contributors with the skills needed to improve the quality of the entry.

And it is again this extended collaboration that helps tackle another issue, which in principle, could seriously hinder an open collaborative encyclopedia from being implemented: the neutrality of point of views. How can we make sure that a voice, and in particular a voice on a debated topic is not subject to biases due to the author's personal point of view? Think of the diverging opinions about the Arab-Israeli conflict between an Israeli and a Palestinian. Well, in a way, the mechanism of Wikipedia *obliges* the Israeli and the Palestinian to work together through the editorial negotiation of the voice. Of course, complete neutrality is an unattainable goal, and it would be naïve to think that any entry in any encyclopedia, including Wikipedia, could be thoroughly neutral and objective. We may argue that the ideals of complete neutrality and objectivity are not only practically unattainable, but that they do not exist at all. Contributors to Wikipedia are basically asked to 'pretend' that neutrality and

---

[26] Cf. Ma, Tao and Hu 2017. See also the examination of the three-article revision process of Wikipedia contained in Tellis 2010.

objectivity do exist,[27] and are asked to pursue this ideal goal. This is one of the five *pillars*, the basic principles, on which Wikipedia is based.[28]

It might be useful to recall those principles, given that they are the higher-level rules that guide the drafting and collaborative editing of entries:

1. *Wikipedia is an encyclopedia*. This implies that the entries must be recognized as 'encyclopedic' – and therefore relevant and of general interest – by the community of contributors. This also means that a contributor to Wikipedia is not permitted to enter an entry about themselves or about a friend, unless they are well-renowned and so relevant that they can claim to have encyclopedic relevance.

It should be observed that the absence of space limitations due to the fact that Wikipedia is entirely digital and online, as well as the very large number of contributors, allow for a fairly broad interpretation of these principles. Many entries – especially in the language versions of Wikipedia other than English – are dedicated to people or topics of only sectorial or debatable relevance. While this could be seen as breaching the first pillar, it is a way to increase the overall coverage of the encyclopedia. At the same time, using a quite 'weak' interpretation of the first pillar allows Wikipedia to cover a significantly wider range of topics than a traditional encyclopedia could ever afford to consider due to the limited space and the limited number of contributors.

2. *Wikipedia is written from a neutral point of view (NPOV)*. According to what Wikipedia declares in this respect,

> We strive for articles in an impartial tone that document and explain major points of view, giving due weight for their prominence. We avoid advocacy, and we characterize information and issues rather than debate them. In some areas there may be just one well-recognized point of view; in others, we describe multiple points of view, presenting each accurately and in context rather than as "the truth" or "the best view". All articles must strive for verifiable accuracy, citing reliable, authoritative sources, especially when the topic is controversial or is about a living person. Editors' personal experiences, interpretations, or opinions do not belong on Wikipedia.[29]

As discussed above, the 'neutral point of view' basically operates as a regulatory ideal, and the task of resolving the inevitable disagreements on the concrete application of the principle is entrusted to editorial negotiation and is facilitated by an existing discussion page for each entry intended for the exchange of opinions between contributors. The discussion page – which can be reached from the 'Talk' tab in each Wikipedia page – together with the version management, is one of the key tools of Wikipedia: occasional users rarely visit the page and often completely ignore its existence, while contributors use it very often.[30]

---

[27] It would be preferable for Wikipedia to explicitly recognize the nature of a purely regulative ideal of the Neutral Point Of View. The historical evolution of the formulation of the second pillar, which can be reconstructed from the history page at https://en.wikipedia.org/wiki/Wikipedia:Five_pillars, seems to (slowly) move in this direction. The original, stronger formulation might have been influenced by Ayn Rand 'objectivist' philosophy, endorsed by Jumbo Wales: Cf. Lih 2009, chapter 2.

[28] The five 'pillars' are outlined at *Wikipedia, The Free Encyclopedia*, s.v. "Wikipedia:Five pillars" (accessed April 2, 2021). https://en.wikipedia.org/wiki/Wikipedia:Five_pillars.

[29] *Ibid*.

[30] On how the 'Talk' page works, see Laniado, Tasso and Volkovich 2011 and Tavosanis 2020b.

3. Wikipedia is free content that anyone can use, edit, and distribute. This principle refers to the Creative Commons Attribution-ShareAlike 3.0 unported licence and to the GNU Free Documentation License (GFDL) which are used for most Wikipedia content.[31]

4. Wikipedia has a code of conduct. This principle recommends mutual respect between the collaborators of the encyclopedia, engaged in a collaborative and non-conflictual project; it might be considered as an adaptation by the Wikipedia project of the principle of charity or maximization of the agreement:[32]

> Respect your fellow Wikipedians, even when you disagree. Apply Wikipedia etiquette, and do not engage in personal attacks. Seek consensus, avoid edit wars, and never disrupt Wikipedia to illustrate a point. Act in good faith, and assume good faith on the part of others. Be open and welcoming to newcomers. Should conflicts arise, discuss them calmly on the appropriate talk pages, follow dispute resolution procedures.[33]

5. Wikipedia has no firm rules. It is interesting to observe that before September 27th, 2010 this statement was written in the form of a typical logical closure principle: "Wikipedia doesn't have firm rules besides the five general principles presented here".[34] The revision of this rule and the choice of dropping the closure principle – achieved after a lively debate among contributors – clearly weakens the status of the five pillars, which are therefore to be seen as regulatory principles rather than as prescriptive norms.

It is also interesting to observe that the formulation of the five pillars is not the same for all the national versions of Wikipedia. From this point of view, the general Wikipedia project should probably be seen as based on a constellation of slightly different and evolving editorial policies rather than as a single, uniformly adopted, and well-defined model.[35]

Are these rules enough to guarantee the success of an encyclopedic undertaking operating in a form that is often considered to be 'amateurish'? As a matter fact, considering Wikipedia as a substantially 'amateurish' encyclopedia is probably wrong today. The success of Wikipedia (both in the number of users and in the reliability of most entries) is linked, on the one hand – as we saw above – to the quality of the editorial negotiation tools offered by the Wiki platform. Wikipedia is backed by a very robust technical tool (as a further example we can point out the use of *bots*, namely platform applications which automatically correct formal errors and help standardize stylistic choices). On the other hand,

---

[31] Cf. *Wikipedia, The Free Encyclopedia*, s.v. "Copyrights" (accessed April 2, 2021). https://en.wikipedia.org/wiki/Wikipedia:Copyrights.

[32] On the principle of charity cf. Feldman 1998.

[33] *Wikipedia, The Free Encyclopedia*, s.v. "Wikipedia:Five pillars" (accessed April 2, 2021). https://en.wikipedia.org/wiki/Wikipedia:Five_pillars.

[34] Cf. the history tab *Ibid.* A previous version of the page had the logically flawed formulation "Wikipedia doesn't have firm rules besides the four general statements above". When formulated as such, the closure was clearly not effective since the fifth rule itself was not included. The history of the revision of the fifth pillar is an interesting example of the idea that Wikipedia policies and guidelines are not carved in stone, a thesis which is part of the actual content of this very pillar.

[35] This also holds if Wikipedia is considered from the point of view of users from different countries, taking into account their different needs and reading behaviors: cf. Lemmerich *et al.*, 2019. On the idea of an 'ethnography' of Wikipedia, cfr. Jemielniak 2014 and, from a linguistic point of view, Tavosanis 2020b.

it must be taken into account that the widespread idea of an encyclopedia entirely built by 'amateurs' is now inaccurate. Among the contributors of Wikipedia, many are reasonably experienced, and as the encyclopedia has increased its user base and has improved its reputation, as a consequence, the number of experts who use it and can contribute to improving the quality of the entries related to their fields of expertise has increased.

Two relevant examples of most competent contributions to Wikipedia helping achieve a qualitative improvement of its contents are the two spheres of librarianship and education. In both cases, we experienced the development of professionally oriented initiatives and projects (such as 'Wikipedia loves libraries'[36] or 'Wikipedia goes to school'[37]) intended to promote a better understanding of Wikipedia as a tool, and to foster the quality of collaboration when drafting its entries. As for the world of Italian libraries, the reasons behind this interest are effectively summarized in two passionate blog posts by Virginia Gentilini and Eusebia Parrotto (Gentilini 2015 and Parrotto 2016); it may be useful to quote some of the passages:

> If our users do not come to us, but go to Wikipedia, then we must go to Wikipedia too. It is part of our work, of our mission. It is not enough to attend a course on online information resources and know that it exists: Wikipedia must also be "our" library, the place where we actively practice our profession. The same care we used in choosing paper encyclopedias, should now be used it in verifying Wikipedia entries, and in improving them not only with content, but also and above all with sources. The sources are in the library, around us. They are the books we have purchased, the collections of periodicals that we keep, the quality resources that our professional and informed approach to the Internet allows us to know. We find them with closed eyes, if you like. It is our job. […]
>
> The next time when looking at Wikipedia we are disappointed at the poverty or unreliability of an entry, regretting the not-so-distant times when knowledge was locked up in authoritative texts, let's get up, go to the reading rooms or to the warehouses and look for an "authoritative" text on the topic. Then we can go back to Wikipedia, click on "edit" and update the entry (Parrotto 2016).[38]

---

[36] Cf. de la Peña McCook 2014 and Lucarelli 2014.

[37] An Italian project: https://www.wikimedia.it/cosa-facciamo/progetti-le-scuole/.

[38] "Se i nostri utenti non vengono da noi, ma vanno su Wikipedia, allora su Wikipedia dobbiamo andarci anche noi. Fa parte del nostro lavoro, della nostra missione. Non basta frequentare un corso di aggiornamento sulle risorse informative in rete e sapere che esiste: Wikipedia deve essere anche la "nostra" biblioteca, il posto in cui esercitiamo attivamente la nostra professione. La cura con cui sceglievamo le enciclopedie cartacee, ora dovremmo usarla nel verificare le voci di Wikipedia, e nell'arricchirle non solo di contenuti, ma anche e soprattutto di fonti. Le fonti sono in biblioteca, intorno a noi. Sono i libri che abbiamo acquistato, le collezioni di periodici che conserviamo, le risorse di qualità che il nostro approccio consapevole e informato alla rete Internet ci permette di conoscere. Le troviamo a occhi chiusi, se vogliamo. È il nostro lavoro. […] La prossima volta che cercando su Wikipedia ci scandalizzeremo della povertà o inaffidabilità di una voce, rimpiangendo i tempi non lontanissimi in cui il sapere era rinchiuso dentro testi autorevoli, alziamoci, andiamo in sala o nei depositi e cerchiamo un testo "autorevole". Poi torniamo su Wikipedia, clicchiamo su "modifica" e aggiorniamo la voce". The English translation is mine.

Among the many available papers and blog posts on the subject of collaboration between Wikimedia and libraries, cf. Lubbock 2018; Ayers and Zanni 2017; Catalani 2017a, Catalani and Feliciati 2018 (a monographic issue of *JLIS.it* on Wikipedia, libraries and archives; among the papers included, of special relevance from our point of view is Parrotto 2018). The topic was discussed during a workshop organized by the National Library of Florence on November 10th, 2017; for a summary, see Catalani 2018 and Gai 2018.

JLIS.it

A quite similar argument can be made regarding the use of Wikipedia in schools and universities and the role of teachers, once again in both terms of improving users' skills and the quality of entries.[39] To provide one example among the many available,[40] at the medical school of the University of California, final year students were engaged in activities led by lecturers (and evaluated as part of the curriculum) aimed to improve the quality of medical entries in Wikipedia (Azzam *et al.* 2017). This interest is fully justified if we consider that already by 2014, 50% of American doctors used Wikipedia as a reference tool for information on medical topics (Beck 2014; a more recent survey in Smith 2020).

In a nutshell, it is true that Wikipedia entries are not written by individual and recognizable experts, but it is also true that a growing number of experts or at least reasonably competent and motivated users collaborate to improve the accuracy of those entries on a regular basis.

Of course, given the very nature of Wikipedia, this does not guarantee that occasional mistakes, as well as more widespread and systematic issues are avoided. In this regard, there was much discussion about the existence of gender biases in Wikipedia entries, partly due to the disproportion between male contributors (who are still the majority) and female contributors or contributors with a different sexual identity.[41] Further biases were identified in the over-representation of Western culture and the under-representation of non-Western or marginalized cultures (Eom, Young-Ho *et al.* 2015), or in the possible existence of commercial, political or ideological interests behind the activities of some of the contributors,[42] as well as in the possibility that even the very procedures used for collaborative editorial review can lead to biased results.[43]

5. Towards linked data

The solution to some of these issues at least may be represented by the further evolution that online encyclopedism has experienced in recent years, and this takes us away from the traditional concept of encyclopedia (well recognizable even in Wikipedia, which makes explicit reference to it in the first 'pillar', as mentioned above) towards the idea of an encyclopedia built as a strongly structured and semantic-rich database based on rigorous formal ontologies. An encyclopedia of this kind is not intended primarily for the use by human agents, but as an information search and retrieval tool to be

[39] On the use of Wikipedia at school, see Catalani *et al.* 2017 and Catalani 2017b. For an interesting survey of relevant international practices, see Virtue 2017. For a discussion on the use of Wikipedia in the context of flipped classroom methodology, see Zou *et al.* 2020. For a not up-to-date but still interesting discussion on the pros and cons of using Wikipedia in universities, see Wannemacher 2009. For interesting experiences based on the use of Wikipedia in the context of university writing courses, see Tavosanis 2013 and 2019. For an updated discussion and bibliography, see Tavosanis 2020a.

[40] An extensive review of international examples of work on Wikipedia in schools and universities is available on Wikipedia itself: *Wikipedia, The Free Encyclopedia*, s.v. "Wikipedia:School and university projects" (accessed April 2, 2021). https://en.wikipedia.org/wiki/Wikipedia:School_and_university_projects.

[41] On this topic, see *Wikipedia, The Free Encyclopedia*, s.v. "Gender bias on Wikipedia" (accessed April 2, 2021), https://en.wikipedia.org/wiki/Gender_bias_on_Wikipedia, and the reference works therein.

[42] Interesting examples of some of the many possible situations of commercial-driven editing of Wikipedia are discussed in Pinsker 2015. On the possible role of historical revisionism in some Wikipedia entries, see Baldo 2017.

[43] See Martin 2018. An interesting survey (with many examples) of the early debate among wikipedians on the reliability and possible drawbacks of Wikipedia's editorial process is in Lih 2009, chapter 8. For a discussion on how to detect both explicit and implicit bias in Wikipedia articles cf. Hube 2017.

JLIS.it

used by software agents. It is left to such agents to 'mediate' between the formal rigour of structured data and the informal queries submitted by the users, or, if required, to act directly on the basis of the recovered data, or to process it according to instructions, or even to transfer it to other software agents. This way, conversational voice assistants such as Google Assistant, Siri, Cortana, Bixby or Amazon Alexa can 'answer' questions and requests users have submitted via a natural language (and often by voice) by interpreting the question by means of a parsing program, identifying its basic elements, using them to query the database and extract the relevant information, so to finally formulate an answer.[44]

Thus, if somebody in New York asks Google Assistant "Who is the President?", the software agent will decipher, first of all, that we are interested in knowing the name of the President in-office (obviously the simpler and more direct interpretation of "who is she/he?" compared to other possible ones, such as "what role does she/he play?"); will interpret the term 'President' – in the absence of further determinations – as referring to the most important and most mentioned 'President' i.e. the main political appointment in the country; will use the location data to geo-reference the request to the United States. The software agent will then submit a formally accurate version of the query to a database broadly based on *DBpedia*, a 'knowledge base' consisting of a structured and formal version of entries taken from Wikipedia,[45] from which the name we are interested in will be recovered. If we ask the question "how old is the President?", Google Assistant will firstly identify the reference to the President of the United States, carrying out the related search as detailed above. Afterwards, it will replace "President of the United States" with the corresponding name (Joe Biden according to the time when I wrote this paper) and will search the database for the relative entry; from this item – which is structured, and therefore organized into fields with values – it will extract the date of birth, and simply calculate the age at the time of the query.

The construction of highly structured databases – or better, of *knowledge bases* – based on formal ontologies is one of the fundamental requirements of the transition from the web mainly oriented to consultation by human agents, to the 'semantic' web, which can also be used (and perhaps will mainly be used) by software agents, and in which the relationships between elements – and so, also among the entries of an encyclopedia – are explicit and in turn formalized. This is the core of the so-called 'semantic web' project, advanced by Tim Berners Lee in the late 1990s.[46] Today we prefer to use the expression *linked data* to underline the strongly relational element of ontologies the idea is based on.[47]

---

[44] For an in-depth overview of how conversational voice assistants work, cf. Young 2021. A discussion of the relationship between conversational voice assistants and 'classical' AI (from the perspective of Turing's test) is in Roncaglia 2014, an interesting proposal on the use of Wikipedia datasets I AI to generate 'natural' responses is in Dinan *et al.* 2019.

[45] The website is https://wiki.dbpedia.org/; for a discussion on the project see Lehmann *et al.* 2015.

[46] For an accessible description of the nature and purpose of the semantic web, useful is the well-known paper Berners-Lee, Hendler and Lassila 2001; the idea was already incorporated in other late 1990s works by Tim Berners Lee, eg. Berners-Lee and Fischetti 1999.

[47] For an accessible introduction to the idea of *linked data* see Bizer, Heath and Berners-Lee 2011; for a more in-depth presentation, see Heath and Bizer 2011.

JLIS.it

There is little doubt that the future evolution of online encyclopedism will head in this direction, but the analysis of the implications of this development – no doubt far-reaching implications[48] – would go beyond the scope of this paper, and would require a much higher level of expertise than mine.

References

Ayers, Phoebe and Zanni, Andrea. 2017. "Wikipedia e biblioteche: una prospettiva globale". *AIB Studi* 57.1:119–125. http://aibstudi.aib.it/article/view/11560/10889.

Azzam, Amin *et al.* 2017. "Why Medical Schools Should Embrace Wikipedia: Final-Year Medical Student Contributions to Wikipedia Articles for Academic Credit at One School". *Academic Medicine* 92.2:194–200. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5265689/.

Baldo, Tommaso (with the working group 'Nicoletta Bourbaki'). 2017. "Riflessioni sulla narrazione storica nelle voci di Wikipedia". *Diacronie. Studi di Storia Contemporanea* 29.1. http://www.studistorici.com/2017/03/29/baldo_numero_29/.

Barrett, Daniel J. 2009. *MediaWiki*, Sebastopol (CA): O'Reilly.

Beck, Julie. 2014. "Doctors' # 1 Source for Healthcare Information: Wikipedia". *The Atlantic*, March 5, 2014. https://www.theatlantic.com/health/archive/2014/03/doctors-1-source-for-healthcare-information-wikipedia / 284206 /.

Berners-Lee, Tim and Fischetti, Mark. 1999. *Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web by Its Inventor*, San Francisco: Harper.

Berners-Lee, Tim; Hendler, James; Lassila, Ora. 2001. "The Semantic Web". *Scientific American* 284.5, 17 May 2001:34–43. https://www.scientificamerican.com/article/the-semantic-web/ (paywall, but the article can be easily retrieved elsewhere on the Internet).

Bizer, Christian, Heath, Tom Berners-Lee, Tim. 2011. "Linked Data: The Story so Far". In Sheth A. (ed.), *Semantic Services, Interoperability and Web Applications: Emerging Concepts*: 205–227. Hershey (PA): IGI Global. https://www.igi-global.com/chapter/linked-data-story-far/55046 (paywall).

Bozzi, Massimo. 1996. *La grande storia del computer, dall'abaco all'intelligenza artificiale*. Bari: Dedalo.

Catalani, Luigi. 2017a. "Biblioteche e Wikimedia: strategie comuni per l'accesso aperto alla conoscenza e la costruzione collaborativa del sapere libero". *JLIS.it* 8.3:100–114. https://www.jlis.it/article/view/12413/11292.

Catalani, Luigi. 2017b. "I progetti Wikimedia per l'apprendimento delle competenze informative e digitali in biblioteca, a scuola, nelle università". *AIB Studi* 57.2:253–263. http://aibstudi.aib.it/article/view/11654/10993.

---

[48] Among the most advanced and interesting – but also debatable – approaches, is the use of artificial intelligence to assess (ad possibly improve) the quality of Wikipedia information; cf. Wang and Li 2020.

JLIS.it

Catalani, Luigi. 2018. "Diffondere la conoscenza aperta per crearne di nuova. Resoconto della giornata "Sfide e alleanze tra Biblioteche e Wikipedia" (Biblioteca Nazionale Centrale di Firenze, 10 novembre 2017)". *Bibliothecae.it* 7.1:391–404. https://bibliothecae.unibo.it/article/view/8456/8173.

Catalani, Luigi *et al.* 2017. *Fare didattica con i progetti Wikimedia*, *Bricks* 4 (monographic issue). http://www.rivistabricks.it/2017/12/19/n-4-2017-fare-didattica-con-i-progetti-wikimedia/.

Catalani, Luigi and Feliciati, Pierluigi (eds.). 2018. *Wikipedia, Libraries and Archives*. In *JLIS.it* 9.3 (Monographic issue).

Cohen, Noam. 2009. "Microsoft Encarta Dies After Long Battle with Wikipedia". *The New York Times* March 30, 2009, https://bits.blogs.nytimes.com/2009/03/30/microsoft-encarta-dies-after-long-battle-with-wikipedia /.

de la Peña McCook, Kathleen. 2014. "Librarians as Wikipedians: From Library History to 'Librarianship and Human Rights'". *Progressive Librarian* 42:61–81. http://www.progressivelibrariansguild.org/PL/PL42/061.pdf.

Dinan, Emily *et al.* 2019. *Wizard of Wikipedia: Knowledge-Powered Conversational agents*. Version 2. Paper submitted to arXiv. https://arxiv.org/abs/1811.01241.

Ducrocq, Albert. 1993. *Le changement global*. Paris: J.C. Lattès.

Eom, Young-Ho *et al.* 2015. "Interactions of Cultures and Top People of Wikipedia from Ranking of 24 Language Editions". *PLoS ONE* 10.3. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0114825 .

Feldman, Richard. 1998. "Charity, principle of". In Klein, Peter and Foley, Richard (eds), *Routledge encyclopedia of philosophy*. Abingdon: Routledge. https://doi.org/10.4324/9780415249126-P006-1 (paywall).

Flagg, Gordon. 1983. "Online Encyclopedias: Are They Ready for Libraries? And Are Libraries Ready for Them?". *American Libraries* 14:134–6.

Flood, Brian. 2020. "Wikipedia co-founder Larry Sanger says online encyclopedia scrapped neutrality, favors lefty politics". *Fox News*, May 21, 2020, https://www.foxnews.com/media/wikipedia-co-founder-larry-sanger-says-online-dictionary-scrapped-neutrality-favors-lefty-politics.

Floridi, Luciano. 1996. *L'estensione dell'intelligenza. Guida all'informatica per filosofi*. Roma: Armando.

Foggin, Sophie. 2019. ""Wikipedia is a broken system," says co-founder Larry Sanger. *150 sec*, May 23, 2019, https://150sec.com/wikipedia-is-a-broken-system-says-co-founder-larry-sanger/11453/.

Gai, Bianca. 2018. "National Meeting 'Sfide e alleanze tra Biblioteche e Wikipedia. Edition 2017' (National Central Library of Florence, 10th November 2017)". In Catalani and Feliciati 2018:167–170.

Gentilini, Virginia. 2015. "Il sogno realizzato del bibliotecario e che cosa fare ora". *Bibliotecari non bibliofili* (blog), January 1st, 2015. https://nonbibliofili.wordpress.com/2015/01/01/il-sogno-realizzato-del-bibliotecario-e-che-cosa-fare-ora/.

JLIS.it

Giles, Jim. 2005. "Internet Enciclopaedias Go Head to Head". *Nature* 439, 14 December 2005:900–901. https://www.nature.com/articles/438900a.

Greenstein, Shane. 2017. "The Reference Wars: Encyclopaedia 's Decline and Encarta's Emergence". *Strategic Management Journal* 38.5:995–1017, https://onlinelibrary.wiley.com/doi/abs/10.1002/smj.2552

Greenstein, Shane and Devereux, Michelle. 2006. "The Crisis at Enciclopaedia Britannica". *Harvard Business Review – Case Studies*, January 1st, 2006, available behind paywall at https://hbr.org/product/crisis-at-encyclopaedia-britannica/KEL251-PDF-ENG and as a paper by the Kellog School of Management at Northwestern University at https://web.archive.org/web/20130313000034/http:/www.kellogg.northwestern.edu/faculty/greenstein/images/htm/research/cases/encyclopaediabritannica.pdf.

Greenstein, Shane and Zhu, Feng. 2018. "Do Experts or Crowd-based Models Produce More Bias? Evidence from Encyclopaedia Britannica and Wikipedia". *MIS Quarterly* 42.3:945–959. http://fengzhu.info/BritannicaWikipedia.pdf.

Guide Officiel Exposition Universelle de Bruxelles 1958. 1958. Bruxelles: Desclée & Cie.

Heath, Tom and Bizer, Christian. 2011. "Linked Data: Evolving the Web into a Global Data Space". *Synthesis Lectures on the Semantic Web: Theory and Technology*, 1.1:1–136.

Holman Rector, Lucy. 2008. "Comparison of Wikipedia and other encyclopedias for accuracy, breadth, and depth in historical articles". *Reference Services Review* 36.1:7–22. https://www.emeraldinsight.com/doi/pdfplus/10.1108/00907320810851998 (paywall).

Hopkins, Heather. 2009. "Britannica 2.0: Wikipedia gets 97% of Encyclopedia Visits". *Hesperian Hitwise*, January 23, 2009, https://web.archive.org/web/20120326111318/http:/weblogs.hitwise.com/us-heather-hopkins/2009/01/britannica_20_wikipedia_gets_9.html.

Hube, Christoph. 2017. "Bias in Wikipedia". *Proceedings of the 26th International Conference on World Wide Web Companion (WWW '17 Companion)*: 717–721. Geneva: International World Wide Web Conferences Steering Committee. https://doi.org/10.1145/3041021.3053375.

Kister, Kenneth. 1994. *Kister's best encyclopedias: a comparative guide to general and specialized encyclopedias*. Phoenix: Oryx Press.

Koren, Yaron. 2017. *Working with MediaWiki* (2nd ed.). CreateSpace Independent Publishing Platform.

Jemielniak, Dariusz. 2014. *Common Knowledge? An Ethnography of Wikipedia*. Stanford: Stanford University Press.

Laniado, David; Tasso, Riccardo; Volkovich, Yana. 2011. "When the Wikipedians Talk: Network and Tree Structure of Wikipedia Discussion Pages". Paper presented at the *Fifth International AAAI Conference on Weblogs and Social Media*, 17-21 July 2011. https://www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/viewPaper/2764.

Lehmann, Jens *et al*. 2015. "DBpedia – a large-scale, multilingual knowledge base extracted from Wikipedia". *Semantic Web* 6.2:167–195. https://content.iospress.com/articles/semantic-web/sw134 (paywall).

Lemmerich, Florian *et al*. 2019. "Why the World Reads Wikipedia: Beyond English Speakers". *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining (WSDM '19)*: 618–626. New York: Association for Computing Machinery. https://doi.org/10.1145/3289600.3291021.

Leuf, Bo and Cunningham, Ward. 2001. *The Wiki Way: Quick Collaboration on the Web*, Boston: Addison-Wesley.

Lih, Andrew. 2009. *The Wikipedia Revolution*. New York: Hyperion.

LITA Newsletter 2, Summer 1980. https://alair.ala.org/bitstream/handle/11213/2104/LITAnewsletter_1980_02.pdf.

Loveland, Jeff. 2019. *The European Encyclopedia From 1650 to the Twenty-First Century*. Oxford: Oxford University Press.

Lubbock, John. 2018. "Wikipedia and libraries". *Alexandria* 28.1:55–68. https://journals.sagepub.com/doi/abs/10.1177/0955749018794968 (paywall).

Lucarelli, Anna. 2014. "«Wikipedia loves libraries»: in Italia è un amore corrisposto…". *AIB Studi* 54.2-3:241–259. http://aibstudi.aib.it/article/view/10108/10146.

Ma, Zhongming; Tao, Jie; Hu, Jing. 2017. "The dynamics of Wikipedia article revisions: an analysis of revision activities and patterns". *International Journal of Data Mining, Modeling and Management* 9.4:298–314. https://www.inderscienceonline.com/doi/pdf/10.1504/IJDMMM.2017.088415 (paywall).

Mailland, Julien. 2017. *Minitel, the Open Networtk Before the Internet. The Atlantic*, June 16, 2017. https://www.theatlantic.com/technology/archive/2017/06/minitel/530646/.

Mako Hill, Benjamin. 2013. "Almost Wikipedia: What Eight Early Online Collaborative Encyclopedia Projects Reveal about the Mechanisms of Collective Action". In Id., *Essays on Volunteer Mobilization in Peer Production,* Ph.D. dissertation, Cambridge (MA): Massachusetts Institute of Technology; draft available online: https://mako.cc/academic/hill-almost_wikipedia-DRAFT.pdf.

Martin, Brian. 2018. "Persistent Bias on Wikipedia: Methods and Responses". *Social Science Computer Review* 36.3:379–388. https://journals.sagepub.com/doi/abs/10.1177/0894439317715434 (paywall).

Menning, Michel. 1986. *L'indispensable pour Minitel*. Paris: Marabout. Digital edition 2016. Paris: FeniXX.

Mollier, Jean-Yves et Dubot, Bruno. 2012. *Histoire de la librairie Larousse (1852-2010)*. Paris: Fayard.

Parrotto, Eusebia. 2016. "Wikipedia, un posto per bibliotecari". *Librarianscape* (blog), January 9th, 2016. https://librarianscape.com/2016/01/09/wikipedia-un-posto-per-bibliotecari/.

JLIS.it

Parrotto, Eusebia. 2018. "Wikimedia projects in public libraries. Feasible ways". In Catalani and Feliciati 2018: 132–140.

Pinsker, Joe. 2015. "The Covert World of People Trying to Edit Wikipedia - For Pay". *The Atlantic*, August 11, 2015. https://www.theatlantic.com/business/archive/2015/08/wikipedia-editors-for-pay/393926/.

Reagle, Joseph M. and Lessig, Lawrence. 2012. *Good Faith Collaboration: The Culture of Wikipedia*. Cambridge (MA): MIT Press.

Rice, James. 1985. "Videotext, Teletext, and the Impact of Television Information Systems on Library Services". In Kent, Allen (ed.). *Encyclopedia of Library and Information Science*, vol. 39, supplement 4, 418–430. New York: Marcel Dekker Inc.

Roncaglia, Gino. 2014. "Computer che copiano: test di Turing, web corpora e filtraggio collaborativo". In Orlandi, Tito (ed.), *Per il centenario di Alan Turing fondatore dell'informatica. Convegno (Roma, 22 novembre 2012)*, Roma: Accademia Nazionale dei Lincei («Contributi del Centro Linceo Interdisciplinare Beniamino Segre» 129). Preprint https://dspace.unitus.it/handle/2067/2614.

Roncaglia, Gino. 2020. "Enciclopedie ed enciclopedismo ai tempi della rete". In *Enciclopedismo antico e moderno*, edited by D'Alba, Vincenzo; Maggiore, Francesco; Maraglino, Vanna; 193–214. Bari: Fondazione Gianfranco Dioguardi («I Quaderni di Varia Cultura», 12).

Smith, Denise A. 2020. "Situating Wikipedia as a health information resource in various contexts: A scoping review". *PLoS ONE* 15.2. https://doi.org/10.1371/journal.pone.0228786.

Smith, Ernie. 2017. "How the CD-ROM Killed the Physical Encyclopedia". In *TEDIUM*, September 11, 2017. https://motherboard.vice.com/en_us/article/vbb994/how- the-cd-rom-killed-the-physical-encyclopedia.

Tavosanis, Mirko. 2013. "Insegnamento universitario della scrittura 2.0 attraverso Wikipedia". In *Tecnologie e metodi per la didattica del futuro. Atti della 27a DIDAMATICA*: 407–410. Pisa: CNR.

Tavosanis Mirko. 2019. "Scrivere su Wikipedia dall'università alla scuola". In Palermo, Massimo and Salvatore, Eugenio (eds.), *Scrivere nella scuola oggi. Obiettivi, metodi, esperienze*: 173–182. Firenze: Cesati.

Tavosanis, Mirko. 2020a. "L'italiano di Wikipedia e la didattica della scrittura". *Lingue e Culture dei Media* 4.1. https://riviste.unimi.it/index.php/LCdM/article/view/14060.

Tavosanis, Mirko. 2020b. "Le discussioni di Wikipedia". *Lingue e Culture dei Media* 4.2. https://riviste.unimi.it/index.php/LCdM/article/view/14333.

Tellis, Lara. 2010. *The Future of Editing: A Textual Analysis of Three Wikipedia Articles*. Clemson (SC): TigerPrint (Clemson University *All Theses* 973). https://tigerprints.clemson.edu/all_theses/973/.

"Un videodisco-laser per raccontare un paese". Unsigned article in *Lettera da TecnoCity* 3, 1987: 13, http://www.byterfly.eu/islandora/object/librib:628488/datastream/PDF/content/librib_628488.pdf

JLIS.it

Virtue, Andrew David. 2017. "Wikipedia as a Pedagogical Tool". *Wikistudies* 1.1. https://wikistudies.org/index.php?journal=wikistudies&page=article&op=view&path%5B%5D=1.

Wang, Ping and Li, Xiaodan. 2020. "Assessing the quality of information on Wikipedia: a deep-learning approach". JASIST - Journal of the Association for Information Science and Technology 71.1:16.28, https://doi.org/10.1002/asi.24210.

Wannemacher, Klaus. 2009. "Articles as Assignments - Modalities and Experiences of Wikipedia Use in University Courses". In M. Spaniol, Li, Q; Klamma R. and Lau, R.W.H. (eds.), *Advances in Web Based Learning - ICWL 2009. Lecture Notes in Computer Science* 5686:434–443. Berlin-Heidelberg: Springer.

Young, Steve. 2021. *Hey Cyba. The Inner Working of a Virtual Personal Assistant*, Cambridge: Cambridge University Press.

Zou, Di *et al*. 2020. "Flipped learning with Wikipedia in higher education". *Studies in Higher Education* 45.5:1026–1045. https://doi.org/10.1080/03075079.2020.1750195.