sustainability

MDPI

Article

Initiatives to Preserve the Content of Vanishing Web Hosting

Karol Król [1] and Dariusz Zdonek [2,*]

[1] Digital Cultural Heritage Laboratory, Department of Land Management and Landscape Architecture, Faculty of Environmental Engineering and Land Surveying, University of Agriculture in Krakow, Balicka 253c, 30-198 Krakow, Poland; k.krol@onet.com.pl

[2] Department of Economics and Informatics, Faculty of Organization and Management, Silesian University of Technology in Gliwice; Akademicka 2A, 44-100 Gliwice, Poland

* Correspondence: dariusz.zdonek@polsl.pl

Abstract: Free hosting services have contributed to the development of the Internet or even acted as a catalyst thereof. This paper aims to answer the questions of what free hosting services represent for Internet users and why initiatives exist to archive content published on free servers. The empirical part of this study attempts to verify whether websites on free servers were designed in an archaic way, which could justify their discontinuation. Initiatives to preserve the content of vanishing web hosting sites are characterized based on a review of various source materials, including the academic literature and Internet resources. This empirical study involved 168 archaic websites, which were analyzed in three dimensions. Marketing components and design were assessed as well. Each assessment dimension was assigned diagnostic variables. The values of the diagnostic variables were standardized using zero unitarization. It was found that the owners of discontinued servers were not interested in creating and maintaining archives. Hence, numerous grassroots initiatives have emerged to salvage their content, although enthusiasm among the archive community seems to have dwindled. Many grassroots archives are available, but a considerable number are no longer supported. In this context, this paper proposes the term 'vanishing hosting'. It provides a nostalgic and sentimental perspective on the termination of free hosting services. The authors noted that free hosting services have largely lost their past import. It is demonstrated that free servers traditionally hosted archaic websites, which justified their deletion from a business perspective. Moreover, the paper presents an outline of a method to revitalize the tourism sector based on changes brought about by global technology to the digital ecosystem. It is proposed that changes in search engine algorithms vicariously contribute to the revitalization of the tourism sector since they often drive the replacement of old websites with newer versions (latest technology and better quality). This leads to the improved online presentation of tourism service portfolios and the general quality of the content in search engine results.

Keywords: vanishing hosting; digital services; free hosting; digital folklore; archaic websites; digital artifacts; digital heritage

Citation: Król, K.; Zdonek, D. Initiatives to Preserve the Content of Vanishing Web Hosting. *Sustainability* **2022**, *14*, 5236. https://doi.org/10.3390/su14095236

Academic Editors: Laura Di Pietro, Gabriella Arcese and Stefano Poponi

Received: 2 February 2022
Accepted: 6 April 2022
Published: 26 April 2022

1. Introduction

The Internet has changed over the years. There are many methods of contributing to the world wide web today. One can set up a social media account and create a page there, create a video channel, create a website (using one of the many available content management systems, templates, or website creators) and host it using a self-managed domain, or employ a website service such as Weebly, Squarespace, WebWave, or Wix [1]. This has, however, not always been the case. In the times of Web 1.0 in the 1990s, many, or most non-professional websites were hosted on free servers. These sites were usually built spontaneously, and their creators often lacked technical and design skills. Free servers hosted millions of 'the ugliest websites' that ever existed, often brimming with garish graphics, flickering images, tile backgrounds, visitor counters, and guestbooks [2].

Many of them lacked professional design. At the same time, these 'awkward' websites allowed people to share their passions to others, form groups, and jointly create content. The technically modern, responsive websites of today typically combine repetitive and established design formats with photographs of impeccably groomed hairstyles and smiles with soulless texts. The content often seems 'sterile', standardized, and corporate in nature [3].

Hosting services, both free and paid, have been available to individuals for years and it looks like they are here to stay. What has changed is their quality and range, to a certain extent, due to growing competition and user expectations. Despite the increasing availability and falling prices of paid hosting services, the market still has room for free hosting. These were particularly popular in the Web 1.0 era, and played a significant role in the growth of the Internet [4]. However, many of these hosting services do not exist anymore, in a phenomenon which can be described as 'vanishing hosting'.

Free hosting was used by rural tourism facilities in Poland to an unprecedented scale to present and promote lodgings from 2000 to 2019. The most popular free hosting services among owners of rural tourism facilities were those provided by Republika WWW and Miasto Interia. These services were discontinued after search engine algorithms changed [5,6]. This fact may have significantly contributed to the revitalization of the rural tourism sector, since many owners of tourism facilities were forced, to some extent, to update their websites.

Websites on free servers were of poor quality [7], but when paid services replaced free hosting, websites were often substituted with standard-compliant ones. In this regard, 'a piece of the Internet was replaced', as illustrated by websites dedicated to rural tourism facilities in Poland, for example. Therefore, one can assume that changes in search engine algorithms contributed to the revitalization of the tourism sector since they often drove the replacement of old websites with newer versions (of better quality). This led to the improved online presentation of tourism service portfolios and increased the general quality of the content in search engine results. Still, archaic websites were unique and provided historical insight into how tourism services were changing, for example. They also consistently documented changes in farm infrastructure (extensions, modernizations, repairs, or improved service standards). Therefore, online content must be archived, since it is a valuable source of historical information about socioeconomic, spatial, or cultural changes [8,9].

This paper consists of two parts: theory and empirical research. The theoretical part attempts to answer the question of what free hosting services represented for Internet users and why initiatives exist to archive content published on free servers. Moreover, the authors tried to determine whether the use of free hosting was based only on technical, pragmatic reasons, or were feelings and emotions such as nostalgia or sentimental attachment involved. The causes may be more prosaic as well, such as the lack of skills to transfer content to a new platform. To tackle these questions, hosting services are profiled and categorized as free and paid, and their prominent features are itemized. Next, the term 'vanishing hosting' is defined, and examples of free hosting discontinuations are described, together with their consequences. Initiatives to conserve content lost due to shutdowns of free hosting services are also characterized. The causes and practical implications of this problem are described together with deliberations on the future consequences of this phenomenon. The empirical part of this paper attempts to verify whether websites on free servers were designed in an archaic way, which could justify their discontinuation.

The remainder of this paper is structured as follows. Section 2 describes hosting services and their functions, divided into paid and free services. Then, the term 'vanishing hosting' is discussed. Section 3 focuses on selected, formerly popular hosting services that were shut down for various reasons. Section 4 presents a number of initiatives to preserve the content of vanishing hosting. Section 5 provides the methodology applied in this research, focusing on assessing the quality of free hosting websites. Section 6 presents the research results. Additionally, the results of synthetic and algorithmic measurements

are juxtaposed with the perception of archaic websites among users of free hosting services. The final section presents the conclusions, limitations, and further research.

2. What Is Web Hosting?

Hosting involves the provision of separate server storage with a service package (software), which determines how the disk space can be used and how convenient it is. The service consists of infrastructure (hardware) and software. Just as with any other service, there are some characteristics that determine the quality and attractiveness of hosting services [10].

Each hosting service has certain parameters. They can be classified into those linked directly to the server's software (technology and scripting compatibility) and parameters of the infrastructure. For the average user, the key is such limits as server space, data transfer (monthly, annual, other), number and types of databases, or the maximum number of e-mail accounts. Availability of other services, such as a software installer (for example to install a content management system, CMS) or website wizard are important as well. Other vital attributes are data security, backup copies, or technical and customer support.

Basic hosting package services include FTP (File Transfer Protocol) accounts and an FTP client (file manager) or databases. Databases can be used with such applications as CMS. Hosting package software can manage domains, add or park them, create subdomains, and redirections. With hosting packages, users can manage e-mail accounts and e-mail settings as well. Most services can be configured on a basic or advanced level [11].

There are two types of hosting services: free and paid. Paid hosting guarantees server reliability and efficiency and comes in several forms. The most popular of them are shared (standard) and dedicated hosting. Shared hosting is the most common option, offering the best ratio of price-to-service. This service is the most popular among micro and small enterprises.

2.1. Paid vs. Free Hosting

Although providers of free hosting have improved the service quality in recent years, the differences between paid and free hosting remain significant (Table A1). Free hosting does not come with hotlines or helpdesks. Messenger communication is a rarity as well. Few providers offer free hosting packages without advertisements. Also, it may be difficult to choose a provider since service parameters are described in an insufficient, often minimalistic manner or not published at all. It is completely different in the case of paid hosting services, where the provider has to earn trust, especially with the first impression, usually via a website.

Users of paid hosting enjoy much better-performing services. Moreover, paid offers are described in more detail. Websites of paid hosting providers have telephone numbers, address data, and registration numbers so that they can be fully identified. The customer can contact the company by telephone, e-mail, or a messenger. All this contributes to the feeling that actual specific people are behind the service. In the case of free services, all of the user can do is to report errors via e-mail or look for help on forums.

The user of free hosting has no guarantee of uninterrupted server availability. Terms and conditions of free services allow the provider to cease service, delete user accounts at convenience, or interrupt services due to upgrades. What is more, providers of free hosting reject liability for server failures and resulting data loss. Users are responsible for backup copies. Furthermore, providers usually consider free hosting a 'customer acquisition cost'. Users have access to a free service, but it is strictly limited, and they are regularly encouraged to upgrade to a paid package.

The lack of quality (reliability) guarantee and hotline or helpdesk makes free hosting unattractive for commercial use. Free hosting services are usually employed in amateur projects such as private, family, or hobby websites.

2.2. What Is Vanishing Hosting?

The website hosting market is changing rapidly. Service providers rise and fall. The added values of hosting packages grow. Competition pushes prices down and service quality up. Markets are dominated by large players, although the number of smaller, local providers is considerable.

Still, no hosting lasts forever. Obviously, no hosting server is guaranteed to offer 100% availability. Each hosting provider, paid and free both, may be shut down. The difference lies in the obligations under free and paid hosting contracts.

Websites come and go every day. The amount of irrevocably lost resources, files, contents, documents, or multimedia, grows by the day. This state of affairs may stem from a multitude of reasons, from technical, economic, to legal or human causes. The process seems natural or even inevitable from the global perspective. It is the matter of transience, which seems to be unavoidable and apply to digital resources as well [12].

Online services provided through websites grow old, lose users, and are discontinued. The Internet is a dynamic environment. The way we use the network keeps evolving. Terminal devices and recipients change. It may be more profitable to develop new services than upgrade existing ones. It is a global problem that will persevere as long as Internet technologies are developed [13,14]. The same applies to paid and free hosting services.

Vanishing hosting is a more descriptive term for free hosting services. It was free hosting services that contributed to the development of the Internet as a catalyst at the turn of the century. Vanishing hosting produces no symptoms as it progresses. The infrastructure grows old and worn out. Users grow scarce, and server performance declines. Vanishing hosting in unprofitable and its fate has been sealed. It is uneconomic. No-one knows when vanishing hosting actually started to vanish. Was it due to a lack of development funding or adverse changes in service configuration?

Hosting vanishing starts small and progresses slowly but steadily. Having reached its last days, vanishing hosting tries to make users notice its demise. Its voice is too low, beyond hearing for many. This is why so few users of vanishing hosting are ready for the last act. For most of them, it comes as a sad surprise. Others will not even notice.

3. You Cannot Use These Hosting Services Any More

Internet services were born in garages or computer laboratories [15]. Some of them were created by passionate amateurs or professionals as grassroots, institutional, social, or business initiatives. They were test grounds for hobbyists or final products of sophisticated strategies and project goals [16]. Still, as products, they are subject to the natural life cycle, starting with an idea and budding project, to finished product and implementation, development and use, user base growth, decline of interest, and end of life (replaced with another technology or another project solution) [17–19].

Many hosting providers disappear unnoticed. This is mostly due to the scale and time of their business: the establishment of the brand on the market—in customer, user, and recipient awareness. If the termination has no legal faults, users usually select a different, alternative service available on the market. Still, the end of such service providers that had been part of the market for years, provided global services, and had many users all over the world can be noticed by the media and leaves a trace. Hence, the end of service provider's business can be more or less spectacular. The same can be said of hosting services, although the termination almost always causes a reaction among users. This is due to the specific nature of hosting services. It is a 'space rental' service for content that is usually of value to the creators.

3.1. Deleted City: A Story of GeoCities

Internet services gained particular momentum from 1999 to 2001. The sudden popularity of hosting services was contemporary with an increased competition in the new technology sector. Before the stock exchange downturn in 2000 and 2001, the general public was excited about the 'new economy' based on the Internet ecosystem [20]. The growth

perspective for the e-economy was so enticing that many investors invested enthusiastically in various dot-coms, especially when there was a 'com' in their names. The dot-com bubble (Internet bubble period) was the sudden sky-rocketing of stock of new technology companies, which then slumped in 2000 and 2001, causing numerous bankruptcies. Many investments were spectacular disasters at the time. The dot-com bubble burst had significant economic and social consequences, and is still analyzed today [21].

The shutdown of GeoCities is an example of how the termination of free hosting services reverberated virtually all over the world. GeoCities began in 1994 in Beverly Hills, California (USA). The service was created as Beverly Hills Internet by David Bohnett and John Rezner. It offered free hosting and tools for building websites [22]. In 1996, the users had two options to create a website. They could use a simple creator with templates or a more advanced editor for more sophisticated websites. By autumn 1998, users could create their websites in five new ways: from the form-based and affiliated 'Intel.com Web Page Wizard' to the GeoBuilder [22]. The founder of GeoCities was convinced at the time that 'we all have something to share with each other, which enriches both their lives and ours as well' [23].

GeoCities quickly gained in popularity as it enabled users to create personal websites about themselves, their hobbies, and lives. GeoCities gathered over a million people who published content in theme-specific 'virtual districts'. EnchantedForest held websites for children, HollywoodHills offered sites about celebrities, MotorCity discussed motor vehicles, and so on. Until 1998, GeoCities was one of the most popular websites online [4].

Free hosting was enough for the growing expectations of users of the developing World Wide Web. They sought their place in the network to be its part. The name Beverly Hills Internet was replaced by GeoCities in 1995. The expansion of the server space was compared to a never-ceasing geographical or territorial expansion. The digital space was compared to physical space where one can live free of charge. Such places on the Internet became 'private living space' for many people, although it was just a piece of server storage. It was well-received by users and the mass media. Five weeks after GeoCities opened, it got over 600,000 hits. By summer 1995, it was hosting 1400 websites. Then, the numbers skyrocketed [22]. By mid-1998, the site was one of the top ten places on the web and was growing by 18,000 new users a day. GeoCities was listed in August 1998. A year later, it was acquired by Yahoo! for USD 4.6 billion [24]. After 1999, when Yahoo! changed the service specification, the number of users of GeoCities started to drop. GeoCities was shut down in 2009.

As a consequence, all files of GeoCities users were gone with most websites and multimedia content irreversibly lost. The relatively early warning to the users did not help. Many of the e-mail notifications were not received.

At the beginning of the 1990s, when the Internet became part of the every-day life, GeoCities brought a sense of community. GeoCities users built communities and helped each other grow their websites to constitute a whole. It was there that users made their first, often stumbling steps in the world of online publishing. At that time, GeoCities was one of the largest digital content collections ever [25].

GeoCities played a crucial role in the shaping of the network we know today. GeoCities was one of the first hosting platforms for 'user-generated content'. It was a time when creators of websites had no generally acceptable design standards to follow. It was in this 'vacuum' that users experimented with graphic design and materials they shared. Diaries, 'under construction' messages ridiculed today, blinking or sliding text (marquee), visitor counters, and background music were characteristic of the period [4].

3.2. Almost like GeoCities: Miasto Interia

'Miasto Interia' (Interia City), was a Polish counterpart of GeoCities. It was a free hosting service, also known as 'Strefa Interia' (Interia Zone). Website addresses took the form of subdomains: user-name.w.interia.pl or user-name.w.interiowo.pl. The servers were shut down on the 21 May 2018. User accounts and websites in the interiowo.pl domain

were closed on the 19 April 2018, with accounts and websites in the strefa.pl domain shut down on the 21 May 2018. In a way, these dates set the limit of the economic feasibility of large-scale free hosting in Poland, as such model started to generate more costs than benefits. This mostly occurred due to changes in Google algorithms, resulting in lower search scoring of websites that used free hosting [7].

The decision to discontinue Interia-operated free hosting was objectively justified. The user base dropped, and Interia's free servers became the testing ground for amateurs and home-grown hackers, with the majority of websites characterized by poor quality [7]. Moreover, a large proportion of websites created from 2000 to 2007 were not updated at all. On the other hand, the content of Miasto Interia was an amalgamate of 'digital oddities', similarly to GeoCities.

3.3. Republika WWW—The Last to Leave

The discontinuation of Miasto Interia reverberated loudly in the Polish media. Although the Polish hosting market had other providers such as prv.pl or friko.pl, only Republika WWW matched Miasto Interia in terms of popularity. Websites on free servers of Republika were available at user-name.republika.pl.

Republika WWW was established in 1999. In 2003, it had over 300 thousand users. Apart from the most popular hobby sites, reflecting private interests of users, Republika WWW's servers hosted websites of small and medium businesses, grassroots and local-government organisations, schools, public institutions, and Internet communities.

Users of Republika WWW's hosting could create their own website free of charge, with the help of guides and knowledge articles about website design provided by the company. Furthermore, in 2000, Republika began to offer an innovative website building tool, 'WebMajster'. WebMajster was an online WYSIWYG (what you see is what you get) editor, with only GeoCities providing a similar tool at the time. Republika WWW expanded its service and tool offer from 2001 to 2005, resulting in a surge of new users. At that time, Republika WWW was one of the most popular free hosting service providers in Poland. Republika's team strongly encouraged user activity. They held contests for the most interesting websites and promoted their creators. Republika WWW, its blog, and online chat gathered a significant group of loyal users.

In February 2018, users of Republika WWW were notified that the portal would be shut down. In response to popular demand, the date was postponed until 30 March. Republika WWW was very popular among the users of the Polish Internet, but it had been losing customer base for years, making its maintenance uneconomical [7].

Republika WWW was yet another brand that went offline after many years. The discontinuation of free hosting services was a part of a wider change scheme, which involved the shutdown of a popular blogging platform, Blog.pl. Another blogging platform, Blox.pl, was closed on the 29 April 2019. All of the content contained on those hosting platforms vanished. The decision was justified by the diminishing popularity of blogs resulting from the expansion of social media.

Bloog.pl was established in 2005 and the expectations of users changed over the 12 years of its activity. Blogging platforms failed to follow the trends and were replaced by social media. Other providers were aware of this change as well. After twelve years, Agora closed down its social media service ludzie.gazeta.pl in early 2019. The reasons included decreasing popularity and the General Data Protection Regulation coming into force. Due to the above, the providers believed it was infeasible to continue maintaining such blogging portals.

3.4. Goodbye, Digart and FortuneCity

Digart was founded on the 2 February 2002. What was special about it was that it was a Polish portal with a community familiar with the local market principles and ready to assist its members. Onet, the owner of the portal, closed it down at the end of January 2018. All user files were lost.

Digart.pl was one of the first portals in Poland where authors of images, prose, and poetry could publish their work, amateurs and professionals alike. The website drew on its American counterpart, DeviantArt. Digart was quickly embraced by those who sought to promote their work and exchange opinions with people of similar interests. Nevertheless, Digart was losing user base in its last years, in the end turning into nothing more than a sentimental memory. When Internet users in Poland used Grono.net instead of Facebook, and Gadu-Gadu instead of Messenger, Digart.pl did what Instagram does today. The 'natural causes' of its death included the deluge of low-quality content (the service was free) and developmental stagnation. With time, its functions were taken over by better-known international portals. The founder of the portal, Piotr Kaczor, commented on its closure in such words: 'It's a shame to look at the inevitable death of a portal which was there for a half of my life and had such a great impact on the Polish Internet'.

In turn, FortuneCity was another popular free hosting service provider from New York. It was established in 1997 by two British entrepreneurs, Richard Jones and Dan Metcalfe, and closed on the 30 April 2012. The official reason behind this was the growing cost of server maintenance. The users were given an opportunity to create backup copies of their files. Nonetheless, just as in other cases, the e-mail about the shutdown failed to reach many of them, with a large part of the content from FortuneCity servers irrevocably lost.

3.5. America on Line (AOL) Hometown

AOL Hometown is a hosting service operated by AOL, one of the largest Internet service providers in the United States. It offered the users tools such as 1-2-3 Publish and Easy Designer, allowing to create website without the knowledge of HTML. The AOL Hometown hosting service was the most popular at the beginning of the 21st century. It was shut down on the 31 October 2008.

AOL Hometown hosted pages created by AOL members and non-members, providing this service both as part of its ISP package and for free. As with other free hosting services, users of AOL Hometown had to provide their e-mail addresses and agree to have advertisements displayed on their websites. AOL Hometown was structured as a portal and divided by topic. Users added websites by selecting their appropriate category and community. Most websites hosted on free AOL Hometown servers were amateur projects. The website usually consisted of one home page and several simple pages written in notepad or created using a template. AOL attempted to improve the appeal of its service with a Member Hall of Fame, a list of the best home pages, or Perfect Picture, a service allowing users to publish their profiles online [26].

4. Initiatives to Preserve the Content of Vanishing Web Hosting

The history of web archiving goes back more than 20 years, with the first initiatives launched in 1996 by the Internet Archive, the National Library of Australia, and Sweden [27]. Also, the National Library of France (BnF) undertook its first web archiving experiments in 1999 [28]. According to the International Internet Preservation Consortium, 'web archiving is the process of collecting portions of the World Wide Web, preserving the collections in an archival format, and then serving the archives for access and use'. A website can be captured and stored, but preservation of its content ensures it will still be accessible over time. Given this long-term perspective, web archiving requires a strategic approach, as it requires much in terms of technologies, systems, policies, and procedures to make it more than merely harvesting and storing online content [28].

Owners of discontinued hosting services were largely not interested in creating archives. Usually, they merely notified users about the possibility and procedure for archiving content of their discontinued accounts. However, the information often failed to reach the users, and many among those who received it were unable to create backup copies. Therefore, after many free hosting services were shut down, grassroots initiatives came up to preserve content scheduled to be deleted.

4.1. Archive Team

The exploration of the digital remains of GeoCities is a particular challenge for historians digging through Internet archives [22]. Part of the digital heritage of GeoCities users of was saved thanks to the effort of the Archive Team and the Internet Archive. These efforts made it possible to delve into a section of the early web even today.

The preservation of about a terabyte of 'GeoCities universe' data was considered important enough for some to establish a task force of 'technical wizards' motivated by a profound belief that the whole existing 'digital culture' is worth saving.

From the time the closing notification was sent and GeoCities was officially closed, a group referring to itself as the Archive Team managed to preserve almost a terabyte worth of GeoCities data. On 26 October 2010, on the first anniversary of what D. Rourke [29] bluntly called a 'digital holocaust', the Archive Team commenced the operation of a digital archive, 'geocities.archiveteam.torrent' [29]. This collection is compressed in a UNIX file system with both 7zip archives and tape archives (gtar). 'If you're a bit of a data tourist and just want to waft in the scent of a web era gone by, please go to one of the Geocities mirrors that were put up in the wake of the end of Geocities' [30].

4.2. GeoCities Special Collection 2009

Some content of GeoCities is available from the Internet Archive as well. The Internet Archive has been a not-for-profit digital library, and a home to a giant archive of the public web since 1996. The web archive is viewable for free via the Wayback Machine. The Wayback Machine of the Internet Archive, with its original slogan 'surf the Web as it was', was conceived and presented in part as a solution to the 404 problem, the response code signifying that the file or web page was not found [31].

The Internet Archive launched several special deep collection crawls, including specific sites nominated by the public, over the last few months when GeoCities was operational, to help make the archive of GeoCities sites as deep and thorough as possible. GeoCities Special Collection 2009 [32] is a special collection of GeoCities websites, indexed from July to October 2009. As the indexing was based on publicly available directories and links to GeoCities websites, not all websites hosted by GeoCities were archived. This omittance mostly concerned the websites that were rarely visited and had few backlinks [12].

4.3. OoCities and Geocities Archive

OoCities is a digital archive of the early web. The purpose of OoCities is to preserve websites hosted by GeoCities. The initiative was founded on the belief that Web 1.0 era websites are a unique source of information. These websites were of cultural importance, attracting significant attention and representing an uncommon style of online publications from the 1990s.

OoCities gives access to many GeoCities websites preserved from 20 to 27 October 2009. Nevertheless, due to both time and financial restrictions, OoCities archivists managed to save 'merely' two million websites. Today, they are hard or even impossible to access since OoCities lack the funds to share them.

Just as with OoCities, GeoCities Archive is a portal in the memory of GeoCities. Its creators attempted to archive the artefacts of the early Internet culture. GeoCities Archive is an initiative to retrieve and provide access to the lost content. In turn, Geocities.ws is a disputable initiative to preserve the content of GeoCities and FortuneCities. For the most part, it promotes paid hosting services, and archive projects seem to be just an addition.

4.4. Internet Archeology: The Enter 99 Archive and One Terabyte of Kilobyte Age

Enter 99 comprises 99 entry pages hosted by GeoCities. Being a certain collection of 'works of art' Enter 99 attempts to preserve the phenomenon of the entry page, once a ubiquitous part of every website.

The Enter 99 archive was created by Jacob B. Engblom and Ryder Ripps. It is part of the Webgrabs section of InternetArchaeology.org. Webgrabs presents the layout, content,

and images of unique websites from the late 1990s, but its primary objective is to preserve websites hosted on GeoCities.

Internet Archeology was established in 2009. Its primary goal is to preserve digital artefacts and emphasize their role in the history of the Internet. The portal only collects images believed to best reflect the technological and cultural changes. Files in the archive (mostly JPGs or GIFs) are categorised by topics. Regrettably, InternetArchaeology.org has not been updated for several years.

The digital cultural heritage preservation movement includes artistic initiatives as well. One of them was the One Terabyte of Kilobyte Age exhibition by Olia Lialina and Dragan Espenschied. It was held in London in 2013 and featured thousands of GeoCities websites, with selected exhibits available online [33].

5. Materials and Methods

In its early days (Web 1.0), the Internet was a one-way communication channel, a system of interrelated hypertext documents. Web 1.0 websites were static and prevented interaction, with users mostly able to browse published information. Characteristic technologies of the era include FTP, IRC, Usenet, E-mail, SQL, SGML, and File Servers. Web 1.0 pages usually followed HTML 2.0, HTML 3.0, or HTML 4.0, and ranged from 700 to 900 px in width to match CRT monitors of the time [34]. The Web 2.0 era saw a dynamic development of CMS, social networks, and bi-directional communication. Web applications facilitated global content generation and information exchange [13]. In turn, Web 3.0 is a transformation of the network into a database where information can be published using artificial intelligence, data analytics, geoinformation, 3D visualizations, and augmented reality. Web 4.0, on the other hand, connects various devices in real time, with development equated to interconnecting as many of them as possible [35]. Each of these concepts entails a change in development techniques, technology and hardware, and new ways of access (both online devices themselves and the content accessible through them). This led to design techniques that provide responsiveness and in-depth personalization, followed by new development concepts, such as long-page design or one-page design. It was a shift from machine experience to user experience. Website editors paid increasing attention to the quality of websites, development techniques (particularly concerning search-engine optimisation, SEO) and content quality (Content is King) [36].

It is not only the dwindling user base that encourages closures of free hosting services. New google algorithms favour carefully edited, diversified, and unique content [5,6], while adaptation of websites to mobile devices and valuable links grew more important. Unfortunately, most websites on free servers were amateur projects. Many of them have not been updated for significant period of time, giving the impression of being 'abandoned' [7].

This study analyses a sample of 168 websites hosted on free servers of Miasto Interia and Republika WWW. Firstly, addresses of the websites were collected (Figure 1). Next, attributes characteristic for archaic Web 1.0 websites (diagnostic variables) were selected [13]. Then the data was analyzed using statistical methods.

The research was based on digital artefacts- websites that do not exist anymore but have been archived in the Internet Archive. Their addresses were obtained from online directories, katalog.onet.pl (accessed on 12 December 2017) and katalog.wp.pl (accessed on 12 December 2017).

As we are all digital archaeologists [37], our field first and foremost helps identify and describe technology changes regarding online content publication and changes in website design. It is mostly based on case studies. Internet archaeology can 'excavate' digital traces left in online resources. This type of research can provide important insight into the dynamics of digital content creation [38].

Users leave a digital trace when accessing the web, a digital artefact of variable nature. The inherently numerous digital artefacts scattered around the web can provide precious information as long as they are collected, ordered, and analysed. Such information can assist in the development of 'internet archaeologist' knowledge on digital cultural heritage [39].

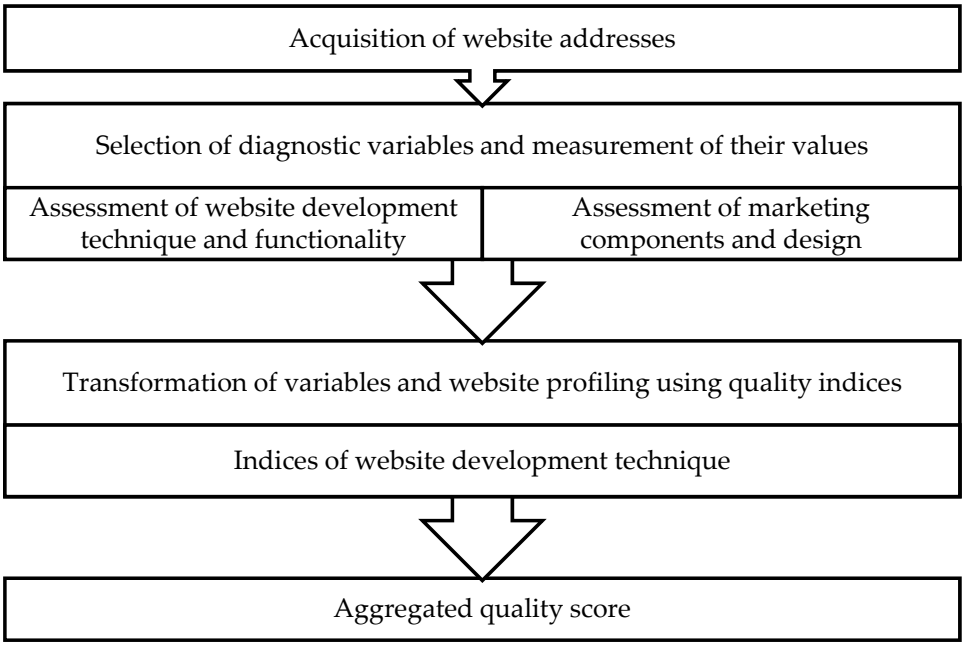

Figure 1. Diagram of the research process.

The websites were analysed in three parameters: (1) website development technique, (2) functionality and usability, and (3) marketing components and design. Each assessment dimension was assigned a diagnostic variable [7].

5.1. Assessment of Website Development Technique

The research employed an exploration method referred to as the cognitive walk-through. It is, generally, a method of expert usability assessment. It involves noting down findings made during the performance of typical tasks, most often relevant to the intended function of a website [40].

The exploration involved copies of websites archived by the Internet Archive. The analysis was focused on both graphics and the structure of the HTML code. The HTML code of websites archived by the Internet Archive additionally includes components of the interface of the archive itself, making it harder to explore the code. To tackle this challenge, the URL parametrisation method was employed. It involves preceding an URL address with the command: 'View-source:' and inserting 'id_' after the timestamp (Table 1).

Table 1. Example of URL address parametrisation.

view-source:https://web.archive.org/web/20090820081158id_/http://deicka.w.interia.pl/ [1]

[1] accessed on 12 January 2022.

Regarding the website development technique, the authors assessed selected design attributes such as compatibility with mobile devices, HTML specification, content management methods, and components enhancing functionality and interactivity (Table 2).

One of the international organisations providing technical standards for creators of websites is the World Wide Web Consortium (W3C). W3C standards allow for creation of websites that are displayed the same way regardless of the software used to access them. Generally followed W3C standards include HTML specifications and recommendations for creating a website accessible for people with disabilities (Web Content Accessibility Guidelines, WCAG). While the guidelines of W3C are not obligatory, they are respected by creators of websites and web applications all over the world.

Table 2. Website development technique assessment attributes.

Group of Variables	Diagnostic Variables	Score Interval and Unit	Type of Variable
Development technique, 5 diagnostic variables	X_1—compatibility with mobile devices (responsive design)	0–1 point	S
	X_2—HTML5 or XHTML Strict (DTD) specification	0–1 point	S
	X_3—content management system	0–1 point	S
	X_4—components enhancing functionality and interactivity	0–1 point	S
	X_5—whole website in Adobe Flash or components of Adobe Flash	0–1 point	I
	X_6—website based on tables or frames	0–2 points	I

S, stimulus; I, inhibitor.

Until recently, website content was published using the FTP (file transfer protocol). The editor (administrator) of a website modified the content of hypertext files stored on a local hard drive and uploaded new files to the server, implying basic language tag knowledge. This changed as content management systems became widely used. The content management system (CMS) is a web application, program, or a suite of programs which can be used by people without any IT background to create a website, update it, publish content, and develop the site [17].

5.2. Assessment of Functionality and Usability, Marketing Components, and Design

Websites can have a range of specific functions facilitated by their, more or less useful, features. Primary website functionalities of include a contact form, booking form, shopping form, or payment form as well as an interactive map.

The set of functionality and usability attributes included the availability of a travel map (location of the facility), contact or booking form, the possibility of online payment, a search bar, and other tools that improve the website's usability (Table 3).

Table 3. Attributes of the assessment of website functionality and usability.

Group of Variables	Diagnostic Variables	Score Interval and Unit	Type of Variable
Functionality, usability, 8 diagnostic variables	X_7—contact form	0–1 point	S
	X_8—booking form	0–1 point	S
	X_9—guestbook	0–1 point	S
	X_{10}—online payment	0–1 point	S
	X_{11}—search bar	0–1 point	S
	X_{12}—tools improving usability for visually impaired users according to WCAG 2.0 (font size change, contrast change)	0–2 points	S
	X_{13}—website in a foreign language	0–1 point	S
	X_{14}—static or dynamic map	0–1 point	S

S, stimulus; I, inhibitor.

It was assumed that such facilities as a search bar, tools improving usability for visually impaired users, or a map, improve the usability of a website, which classifies them as stimuli.

Regarding the assessment of marketing components and design, the authors identified multimedia or social media components on the website. Moreover, selected design attributes were assessed as well (Table 4).

Table 4. Attributes of marketing component and design assessment.

Group of Variables	Diagnostic Variables	Score Interval and Unit	Type of Variable
Sources of traffic and marketing communications, 5 diagnostic variables	X_{15}—presence of such multimedia as spherical panoramas or virtual tours	0–2 points	S
	X_{16}—video	0–1 point	S
	X_{17}—social media components	0–1 point	S
	X_{18}—blog or forum	0–1 point	S
Design, 4 diagnostic variables	X_{19}—animated GIF images	0–1 point	I
	X_{20}—*marquee* objects	0–1 point	I
	X_{21}—visitor counter	0–1 point	I

S, stimulus; I, inhibitor.

The availability of multimedia, social media components, and a blog or a forum improves the website usability. The use of gaudy, garish, animated GIF files, marquee objects, or a visitor counter is characteristic of archaic websites. Hence, these attributes are inhibitors.

5.3. Statistical Methods

Values of the diagnostic variables were standardised using zero unitarization. Zero unitarization can standardise numeric intervals of variables and convert their moduli into non absolute values. As the analysed variables varied in nature, the unitarization employed Equation (1) for stimuli and Equation (2) for inhibitors [7].

$$x'_{ij} = \frac{x_{ij} - \min_i\{x_{ij}\}}{\max_i\{x_{ij}\} - \min_i\{x_{ij}\}} \quad (1)$$

$$x'_{ij} = \frac{\max_i\{x_{ij}\} - x_{ij}}{\max_i\{x_{ij}\} - \min_i\{x_{ij}\}} \quad (2)$$

where:

i=1, 2, . . . , n; j = 1, 2, . . . , m,

x'_{ij}—normalised diagnostic variable,

x_{ij}—unnormalized diagnostic variable,

$\min_i\{x_{ij}\}$—the minimum value of the unnormalized diagnostic variable,

$\max_i\{x_{ij}\}$—the maximum value of the unnormalized diagnostic variable.

Standardisation of variables allowed the authors to describe each website with an aggregate quality index, a sum of standardized variables (AQI). The AQI value ranges from 0 to 21 points.

6. Results

From the technical perspective, websites hosted on free servers of Miasto Interia and Republika WWW were outdated. They were built according to obsolete specifications such as HTML 4.0, characteristic of websites created at the beginning of the 21st century. Only four websites conformed to the currently valid HTML 5.0 specification (Table 5). As many as 73 websites were developed using an HTML editor. In 36 cases, it was Microsoft Front Page, which is seldom used today. Some websites were created using Pajączek and Webmajster software, popular in Poland from 2000 to 2007.

Most of the websites used iso-8859-2 encoding, which is proper for content in Polish. Thirty-seven websites were found to use the windows-1250 encoding, which is not used today. Only 15 websites used the recommended utf-8 encoding.

Table 5. Number of websites using each HTML specification.

HTML Specification	HTML	HTML 2.0	HTML 3.0	HTML 4.0	XHTML	HTML 5.0	Total
Number of websites	74	1	3	61	25	4	168
Percentage (%)	44.0	0.6	1.8	36.3	14.9	2.4	100.0

None of the websites were compatible with mobile devices. As many as 106 websites were based on tables (tag <table>), while frames (tags <frameset> or <iframe>) were used on 20 websites. Only one website had a CMS. Websites using tables or frames were characteristic of the Web 1.0 times.

Only 33 websites had dynamic elements, 25 of which were Flash objects (not recommended today). Animated GIFs and marquee objects were relatively rare, found 16 and 17 times, respectively. Archaic visitor counters were found on 67 websites. Interactive tools such as booking, contact, or payment forms were rare. Only 4 contact forms were found.

Guestbooks were identified on 16 websites. Only 12 websites had English versions. Travel map was found on 54 websites. None of the websites offered the option to change contrast or font size. Only 3 cases of videos and 5 occurrences of social media components were identified. None of the websites had a blog or forum.

The websites were scored 1 to 12 AQI points from the maximum of 21. Only one site had a score of 12 points. The most common score was 4 to 7, amounting to 19% to about 33% of the maximum score (Table 6).

Table 6. Number of websites by AQI value.

AQI Value	1–3	4–5	6–7	8–9	>10	Total
Number of websites	11	80	65	8	4	168
Percentage (%)	6.5	47.6	38.7	4.8	2.4	100

Low AQI values are indicative of the archaic development technique, lack of interactivity, and poor functionality. The investigated websites were mostly static and served to provide information only.

7. Discussion

The observations made during the study demonstrated that some owners of agrotourism facilities in Poland prepared new versions of their websites as a response to announced discontinuation of hosting services (service providers informed users about their plans and provided advice on what to do). As a consequence, some websites had notices about new websites displayed until the files were removed from the server. In turn, not (global) changes in search engine algorithms but rather the prospect of losing their online presence motivated owners of agrotourism facilities to replace their websites.

The discontinuation of free hosting in Poland was a starting point a quality change in some areas of the Polish Internet. Changes in the algorithms drove the improvement of Internet quality in many dimensions and on many levels. In this case, it affected national service providers (changed scope and type of hosting and directory services) and the industry, as the initial agrotourism-related websites were replaced with more modern versions hosted on paid ccTLD (country code top-level domain) servers.

The end of free hosting services results in a gap in the worldwide web. Król and Zdonek [12] referred to such incidents as 'digital bumps' or a 'digital tsunami', which reverberates throughout the Internet, damaging its structure, resulting in dead links and link rot. Websites on free servers were also referred to as 'the forgotten Internet' [7]. Research shows that many of them are no longer updated, and seem to have been abandoned and forgotten.

Setting financial and marketing aspects aside, the discontinuation of free hosting platforms is indicative of the relatively low impact of the people who used them. This

phenomenon entailed the removal of many websites belonging to individuals, but also entrepreneurs, schools, libraries, or NGOs that used this infrastructure. All websites, be it private, amateur, or of public institutions, were treated the same and shut down. This experience led to public institutions refraining from the use of free hosting services, with other reasons comprising the decrease in hosting prices, and new regulations such as those regarding collection and storage of sensitive data. Businesses also do not use free hosting anymore, as the hosting cost savings do not equalize the lack of quality guarantee, which often hampers brand building.

7.1. Why Is It Important to Preserve the Content of Vanishing Hosting and for Whom?

Discontinuation of digital services, including hosting, blogs, accounts in social media, or websites can inspire objection, irritation, resentment, anger, or even depressive disorders [41], especially if the service was used for a long time or became an addiction [42,43]. Many users look back on their first interaction with the Internet with emotion and sentiment.

Sentiment analysis and opinion mining is the field of study that focuses on people's opinions, sentiments, evaluations, attitudes, and emotions in written language. It is one of the most active research areas in natural language processing and is also widely studied in text mining, data mining, and Web mining. This research has spread outside of the computer science to the management sciences and social sciences due to its importance to business and society as a whole. The growing importance of sentiment analysis coincides with the growth of social media platforms such as forums, reviews, blogs, micro-blogs, and social networks. This represents a huge volume of opinionated data recorded in a digital form for analysis [44].

7.2. 800 × 600 Pixel Nostalgia

Longing is born from the unavailability of the object [45], with nostalgia describing a sentimental attachment to the past. Usually, nostalgia is a positive, self-sufficient, and social emotion with key psychological functions. Nostalgic narratives reflect more positive than negative effects and are socially grounded [46]. This feeling is a response to negative mood and discrete emotions of loneliness, void, longing, and loss, improving social bonds and self-esteem [47].

Sentimental attachment is a liking and attachment to an object or entity, related to emotionality. When feelings of the user, consumer, or general recipient are deciphered from large datasets such as online comments, they can provide valuable insight into products and services [48].

Sentimental, expressive, and emotional phrases depict the power of sentiment, the strength of longing for youth and carefree days. The nostalgic sentiment promotes a younger perspective on self, which may be beneficial for the psychophysical state [49].

The memories of the first website published on free servers arouse sentiment among users. For many of them, the 1990s was the time of fascination with new technologies and the first contact with the web, as reflected by the sentiment analysis of opinions of forum users (paraphrased) [50]:

- In 1997–1999 (I was 14–16), I was a teenage 'community leader' at GeoCities. I will always love GeoCities—it was not only my introduction to building for the web but to online communities in general.
- GeoCities was my first, very precious experience with the online community.
- I was writing my HTML on paper during classes at school, during lunch, and after school. It was how my first websites were created.
- The memories and anecdotes about GeoCities make me feel sad about it not being around anymore.
- I spent a lot of my youth with Geocities, Angelfire, EzBoards, AIM, and ICQ. It's all almost gone now. We should have preserved much more of the content.

- I hope someone will someday get their hands on unformatted hard drives of some old hosting companies and revitalise the content. It's weird to feel so nostalgic for immaterial stuff.
- It's also weird to reminisce about youth. Despite all of the new tech, the wealth and career I've built, and the people I've met, I sometimes wish I could relive the early 2000s. It feels like all of the adventure and newness has been sucked out of the world.
- I miss the times when each website seemed like an independent Solar system, light years of empty space away from all others. Who you were and what you did on one site had nothing to do with your life on another one? The modern web seems to be a never-ending series of stages where everyone holds their own show for numb spectators instead of a real community.
- I remember looking for a good 'location' for my website for a while and then the many hours learning HTML. I was 12 and had plenty of spare time.
- I had GeoCities when it was called GeoPages. Yup, I'm old!
- GeoCities was really amazing to me. A no-tech guy who just wanted to put stuff online.
- It's a shame so many Web 1.0 sites are gone. I support projects to salvage archaic websites, although I'm a bit sad not all of them can be preserved.
- Access to the archives is unbelievable and brings back so many memories!

Discontinuation of hosting services often can be personal, with exploration of archaic content serving as a 'look back' [51]:

> But after seeing, and searching it, I was floored to be able to find my old page from when I was 16. I'm now 26. The amount of laughter and sheer embarrassment I felt was monumental. This was such an utter joy to experience again! I sheepishly made a page for my girlfriend at the time, Danielle Pryor. And boy, reading over it brought back a flood of fun and very awkward memories! But, oh so good!

For many users, the memory of hosting service shutdown arouses regret and sadness due to the lost content [50]:

- I had a GeoCities website in high school. With 'wooden panels', animated GIFs (a rotating skull and so on), a visitor counter, blinking text, and a 'midi' audio file. I'm so sorry it's lost.
- It's insane how the whole early Internet is actually destroyed and there is no turning back. I was a kid at the beginning of 2000 and many things I had access to were not preserved. It makes me feel sad.

The preservation of digital content is justified as digital archives keep content important not only for Internet archaeologists or institutions but also individual users [51], with comments such as: My dad had an AOL hometown page. He passed away this year, only 52 years old, and very suddenly. I liked checking in on his old site now and again. It was a comfort. Obviously, I was devastated to find it was gone forever, but thanks to this link, not so! The memory of the beginning of the Internet arouses nostalgia regarding the role of the user and content presentation. Today, web is unarguably more dynamic, with a clean aesthetic that barely shifts behind the waves of content that wash over its surface. However, the user has been relegated to the role of material shuffler [29].

Free hosting accounts have not always been treated as a form of personal, even private space. Did hosing service providers such as Miasto Interia or Republika WWW spark such emotions among their users? Desk research suggests otherwise. It may be due to the fact that the most popular free hosting services in Poland were discontinued relatively recently, almost ten years after GeoCities, when the paid hosting market and social media were already sufficiently developed. The number of tools for publishing content online, such as free CMSs, Wiki, or website generators and wizards, grew significantly. While the problem of 'vanishing websites' hosted for free on Miasto Interia and Republika WWW servers has been noticed [7], there were no 'major grassroots initiatives' to preserve the content. Only some private archives have databases of website addresses formerly hosted by these providers [52].

8. Conclusions

One of the aims of this study was to verify whether free hosting shutdown, which in Poland largely took place in 2019, was justified. Due to the relatively long time of operation, many archaic websites, often developed using Web 1.0 techniques, survived in the Polish 'living Internet' relatively long. Service providers did not archive content from discontinued hosting accounts, but a significant portion is available in the Internet Archive.

The results of the synthetic and algorithmic measurements demonstrated that websites hosted on free servers were of poor technical quality and mostly archaic. Consequently, they offered no value to hosting providers due to the low traffic generated. They could, however, have value for the users, arising sentimental feelings in many of them.

Free hosting services gave millions of people the opportunity to create a free website available to the whole world. It was a complete novelty in the 1990s, with some of the websites presenting exceptional value. Should humanity be unable to access these archaic websites, it would mean a loss of a piece of history written by regular people, not designers, programmers, or developers. These arguments support the efforts to preserve such content.

Nonetheless, some files from the discontinued hosting services were salvaged, and some websites were moved to other servers. It was possible, among others, thanks to tools allowing to export data to any medium. But the Internet is not only a collection of digital resources. A certain 'constituent of the web' cannot be exported and saved as a file, the relations forged in virtual communities that bloomed in the Web 1.0 era. Even if files are preserved and websites moved to other servers, the elaborately woven cloth of human relationships can be hard or even impossible to restored.

Not all users were wholeheartedly committed to the creation of content and its publication online. Still, for a specific group of users, the discontinuation of free hosting services was painful as it entailed the breakdown of online relationships and 'digital communities' For such involved users, the shutdown of the hosting meant eradication of a piece of digital cultural heritage, deletion of content worth countless hours of creative work, and silencing of hundreds of thousands of voices of 'regular people' expressing their beliefs on websites. The content represented its period, and thoughts and dreams of people who are now much older or gone. Free servers of the 1990s often kept records of honest, true stories.

Most users, probably, did not give the discontinuation of free servers a second look, considering their content and form obsolete. Despite this, many archive, artistic, institutional, and social initiatives were launched to preserve it. In the 1990s, free hosting services were a window on the world, a door to a different dimension. Teenage users were often personally and emotionally committed to the creation of the content. Many of them associate retro websites with their youth. Their memories are nostalgic and sentimental. Their words touch upon the essence of transience. To lose the archaic content would be particularly painful. After all, exploration of websites salvaged from vanishing hosting is an opportunity to look into the past, representing a piece of a certain continuum.

Limitations and Further Research

This study contains an algorithmic indicator-based analysis of selected development attributes of archaic websites (digital artefacts). However, the research did not cover digital folklore and the content of archived websites, such as photographs. In the context of tourism sector revitalisation, such research aims to provide evidence of socioeconomic changes recorded in pricelists, descriptions of services, and photographs of infrastructure, residential buildings, yards, nature, etc. Therefore, there is a certain research gap that can be addressed by exploring such resources in digital archives.

Superficially, while archaic websites are worthless, digital artefacts have some characteristics that help to date them (put them on the timeline of eras through association of typical features) and determine their purposes and socioeconomic environment, the reality of their existence and functioning. As for websites, such attributes include metadata, development techniques typical of certain times (years), and selected design elements. Content of archaic websites can record changes in service standards, tourism facility infras-

tructure, or cultural landscapes. Therefore, archaic websites and digital archives can be a source of historical information. On the other hand, digital space provides opportunities to experience a destination, object, or phenomenon beyond the real-life reach of many users (due to financial, geopolitical, or other reasons). Information on socioeconomic and cultural changes available in digital archives can inspire further research. The opportunity to represent, present, and experience places, phenomena, or objects in the digital ecosystem can have a similar effect.

Author Contributions: Conceptualization, K.K. and D.Z.; methodology, K.K.; software, K.K.; validation, K.K. and D.Z.; formal analysis, K.K.; investigation, K.K.; resources, K.K. and D.Z.; data curation, K.K.; writing—original draft preparation, K.K.; writing—review and editing, K.K.; visualization, K.K.; supervision, K.K. and D.Z.; project administration, K.K.; funding acquisition, D.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Silesian University of Technology, grants no. 13/010/-RGJ22/0063 and 13/010/BK_22/0065.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All trademarks and registered trademarks mentioned herein are the property of their respective owners. The company and product names used in this document are for identification purposes only.

Acknowledgments: The author wishes to express his gratitude to the reviewers for their constructive criticism, which contributed to the final content of the paper. The paper was written at the Digital Cultural Heritage Laboratory (https://culturalheritage.urk.edu.pl accessed on 2 February 2022) part of the Department of Land Management and Landscape Architecture at the Faculty of Environmental Engineering and Land Surveying of the University of Agriculture in Krakow, Poland. The research was carried out as part of the scientific project entitled: Digital cultural heritage of rural areas (2021/05/X/HS3/00859) financed by the National Science Center Poland.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Appendix A

Table A1. Comparison of selected attributes of paid and free hosting.

Paid Hosting	Free Hosting
	No costs
Subscription fee	For amateur or test websites
Suitable for professional and commercial applications	Longer URL (a subdomain)
	Potential advertisements
Helpdesk and regular server backup copies of files and databases	No helpdesk (only e-mail contact usually)
	No reliability guarantee, use 'as is'
Domain name depends on the user and availability	Questionable service stability and availability
	High risk of failure
No advertisements	Risk of service terms modification
Service Level Agreement	Negative impact on image and user trust, hampers branding
Stable service terms	
Possibility to resell the URL (website)	Impossible to resell the website/blog/URL
Usually high technical quality of servers	Questionable technical parameters of servers
	Smaller range of services and worse service parameters such as minimum storage and bandwidth

Source: own work based on [53].

References

1. Fuchs, C.; Hofkirchner, W.; Schafranek, M.; Raffl, C.; Sandoval, M.; Bichler, R. Theoretical Foundations of the Web: Cognition, Communication, and Co-Operation. Towards an Understanding of Web 1.0, 2.0, 3.0. *Future Internet* **2010**, *2*, 41–59. [CrossRef]
2. Król, K.; Hernik, J. Digital Folklore of Rural Tourism in Poland. *Sustainability* **2022**, *14*, 1165. [CrossRef]
3. Ageeva, E.; Foroudi, P.; Melewar, T.C.; Nguyen, B.; Dennis, C. A Holistic Framework of Corporate Website Favourability. *Corp. Reput. Rev.* **2020**, *23*, 201–214. [CrossRef]
4. Baker, J. *GeoCities and Diaries on the Early Web*; Ben-Amos, B., Ben-Amos, D., Eds.; Indiana University Press: Bloomington, IL, USA, 2020; ISBN 9780253046994.
5. Kingshott, D.; Keck, A. What Panda 4.0 means for online businesses. *J. Digit. Soc. Media Mark.* **2014**, *2*, 153–158.
6. Patil, A.; Pamnani, J.; Pawade, D. Comparative Study Of Google Search Engine Optimization Algorithms: Panda, Penguin and Hummingbird. In Proceedings of the 2021 6th International Conference for Convergence in Technology (I2CT), Maharashtra, India, 2–4 April 2021; pp. 1–5. [CrossRef]
7. Król, K. Forgotten Agritourism: Abandoned Websites in the Promotion of Rural Tourism in Poland. *J. Hosp. Tour. Technol.* **2019**, *10*, 431–442. [CrossRef]
8. Tetiana, B.; Maria, K.; Andriy, P. Development of Method of Search and Identification of Historical Information in the Social Environment of the Internet. In Proceedings of the 2017 12th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT), Lviv, Ukraine, 5–8 September 2017; Volume 1, pp. 196–199. [CrossRef]
9. McGrew, S. Internet or Archive? Expertise in Searching for Digital Sources on a Contentious Historical Question. *Cogn. Instr.* **2021**. [CrossRef]
10. Tajalizadehkhoob, S.; Korczyński, M.; Noroozian, A.; Gañán, C.; van Eeten, M. Apples, Oranges and Hosting Providers: Heterogeneity and Security in the Hosting Market. In Proceedings of the NOMS 2016 IEEE/IFIP Network Operations and Management Symposium, Istanbul, Turkey, 25–29 April 2016; pp. 289–297. [CrossRef]
11. Li, Z.; Yang, D.; Li, Z.; Han, C.; Xie, G. Mobile Content Hosting Infrastructure in China: A View from a Cellular ISP. In Proceedings of the Passive and Active Measurement, Berlin, Germany, 26–27 March 2018; Beverly, R., Smaragdakis, G., Feldmann, A., Eds.; Springer International Publishing: Cham, Switzerland, 2018; pp. 100–113. [CrossRef]
12. Król, K.; Zdonek, D. Peculiarity of the Bit Rot and Link Rot Phenomena. *Glob. Knowl. Mem. Commun.* **2019**, *69*, 20–37. [CrossRef]
13. Cormode, G.; Krishnamurthy, B. Key Differences between Web 1.0 and Web 2.0. *First Monday* **2008**, *13*. [CrossRef]
14. Benito-Osorio, D.; Peris-Ortiz, M.; Armengot, C.R.; Colino, A. Web 5.0: The Future of Emotional Competences in Higher Education. *Glob. Bus. Perspect.* **2013**, *1*, 274–287. [CrossRef]
15. Berners-Lee, T.; Cailliau, R.; Luotonen, A.; Nielsen, H.F.; Secret, A. The World-Wide Web. *Commun. ACM* **1994**, *37*, 76–82. [CrossRef]
16. Król, K. Digital Cultural Heritage of Rural Tourism Facilities in Poland. *J. Cult. Herit. Manag. Sustain. Dev.* **2020**, *11*, 488–498. [CrossRef]
17. McKeever, S. Understanding Web Content Management Systems: Evolution, Lifecycle and Market. *Ind. Manag. Data Syst.* **2003**, *103*, 686–692. [CrossRef]
18. Hynes, G.; Reynolds, V.; Hauswirth, M. A Context Lifecycle for Web-Based Context Management Services. In Proceedings of the Smart Sensing and Context; Barnaghi, P., Moessner, K., Presser, M., Meissner, S., Eds.; Springer: Berlin/Heidelberg, Germany, 2009; pp. 51–65. [CrossRef]
19. Mo, J.P.T.; Lorchirachoonkul, W. Lifecycle Design and Support of Intelligent Web-Based Service Systems. *Int. J. Agil. Syst. Manag.* **2016**, *9*, 135–153. [CrossRef]
20. Buenstorf, G.; Fornahl, D. B2C—Bubble to Cluster: The Dot-Com Boom, Spin-off Entrepreneurship, and Regional Agglomeration. *J. Evol. Econ.* **2009**, *19*, 349–378. [CrossRef]
21. Kraay, A.; Ventura, J. The Dot-Com Bubble, the Bush Deficits, and the US Current Account. In *G7 Current Account Imbalances: Sustainability and Adjustment*; University of Chicago Press: Chicago, IL, USA, 2007; pp. 457–496.
22. Milligan, I. *Welcome to the Web: The Online Community of GeoCities during the Early Years of the World Wide Web*; Brügger, N., Schroeder, R., Eds.; The Web as History; UCL Press: London, UK, 2017. [CrossRef]
23. Ocamb, K. David Bohnett: Social change through community commitment. *Frontiers* **2012**, *16*, 18.
24. Motavalli, J. *Bamboozled at the Revolution: How Big Media Lost Billions in the Battle for the Internet*; Penguin Putnam: New York, NY, USA, 2004.
25. Milligan, I. *Finding Community in the Ruins of GeoCities: Distantly Reading a Web Archive*; Institute of Electrical and Electronics Engineers: Piscataway Township, NJ, USA, 2015.
26. Papacharissi, Z. The Presentation of Self in Virtual Life: Characteristics of Personal Home Pages. *J. Mass Commun. Q.* **2002**, *79*, 643–660. [CrossRef]
27. Schroeder, R.; Brügger, N. *Introduction: The Web as History*; Brügger, N., Schroeder, R., Eds.; The web as history. Using web archives to understand the past and present; UCL Press: London, UK, 2017; pp. 1–19.
28. Vlassenroot, E.; Chambers, S.; Di Pretoro, E.; Geeraert, F.; Haesendonck, G.; Michel, A.; Mechant, P. Web Archives as a Data Resource for Digital Scholars. *Int. J. Digit. Humanit.* **2019**, *1*, 85–111. [CrossRef]
29. Rourke, D. The Impulse of the Geocities Archive: One Terabyte Of Kilobyte Age. Furtherfield. Available online: http://bit.ly/Furtherfield (accessed on 31 March 2022).

30. Torrents (2019), "Geocities—The Torrent", Archiwe Team. Available online: http://bit.ly/ATTorrent (accessed on 31 March 2022).
31. Rogers, R. Doing Web History with the Internet Archive: Screencast Documentaries. *Internet Hist.* **2017**, *1*, 160–172. [CrossRef]
32. GeoCities Special Collection. Saving a Historical Record of GeoCities. *Internet Archiwe.* Available online: https://archive.org/web/geocities.php (accessed on 31 March 2022).
33. One Terabyte of Kilobyte Age Photo, Op. Tumblr. Available online: https://oneterabyteofkilobyteage.tumblr.com/ (accessed on 31 March 2022).
34. Choudhury, N. World wide web and its journey from web 1.0 to web 4.0. *Int. J. Comput. Sci. Inf. Technol.* **2014**, *5*, 8096–8100.
35. Aghaei, S. Evolution of the World Wide Web : From Web 1.0 to Web 4.0. *IJWesT* **2012**, *3*, 1–10. [CrossRef]
36. Müller, J.; Christandl, F. Content Is King—But Who Is the King of Kings? The Effect of Content Marketing, Sponsored Content & User-Generated Content on Brand Responses. *Comput. Hum. Behav.* **2019**, *96*, 46–55. [CrossRef]
37. Morgan, C.; Eve, S. DIY and Digital Archaeology: What Are You Doing to Participate? *World Archaeol.* **2012**, *44*, 521–537. [CrossRef]
38. Harrison, R. Exorcising the 'Plague of Fantasies': Mass Media and Archaeology's Role in the Present; or, Why We Need an Archaeology of 'Now'. *World Archaeol.* **2010**, *42*, 328–340. [CrossRef]
39. Nicholson, S. A Framework for Internet Archeology: Discovering Use Patterns in Digital Library and Web—Based Information Resources. *First Monday* **2005**, *10*. [CrossRef]
40. Nielsen, J.; Mack, R.L. *Usability Inspection Methods*; John Wiley & Sons: New York, NY, USA, 1994.
41. Banjanin, N.; Banjanin, N.; Dimitrijevic, I.; Pantic, I. Relationship between Internet Use and Depression: Focus on Physiological Mood Oscillations, Social Networking and Online Addictive Behavior. *Comput. Hum. Behav.* **2015**, *43*, 308–312. [CrossRef]
42. Błachnio, A.; Przepiórka, A.; Pantic, I. Internet Use, Facebook Intrusion, and Depression: Results of a Cross-Sectional Study. *Eur. Psychiatry* **2015**, *30*, 681–684. [CrossRef]
43. Younes, F.; Halawi, G.; Jabbour, H.; Osta, N.E.; Karam, L.; Hajj, A.; Khabbaz, L.R. Internet Addiction and Relationships with Insomnia, Anxiety, Depression, Stress and Self-Esteem in University Students: A Cross-Sectional Designed Study. *PLoS ONE* **2016**, *11*, e0161126. [CrossRef]
44. Liu, B. Sentiment Analysis and Opinion Mining. *Synth. Lect. Hum. Lang. Technol.* **2012**, *5*, 1–167. [CrossRef]
45. Boym, S. *The Future of Nostalgia*; Basic Books: New York, NY, USA, 2002.
46. Sedikides, C.; Wildschut, T.; Arndt, J.; Routledge, C. Nostalgia: Past, Present, and Future. *Curr. Dir. Psychol. Sci.* **2008**, *17*, 304–307. [CrossRef]
47. Wildschut, T.; Sedikides, C.; Arndt, J.; Routledge, C. Nostalgia: Content, triggers, functions. *J. Personal. Soc. Psychol.* **2006**, *91*, 975. [CrossRef]
48. Villarroel Ordenes, F.; Ludwig, S.; De Ruyter, K.; Grewal, D.; Wetzels, M. Unveiling what is written in the stars: Analyzing explicit, implicit, and discourse patterns of sentiment in social media. *J. Consum. Res.* **2017**, *43*, 875–894. [CrossRef]
49. Abeyta, A.A.; Routledge, C. Fountain of Youth: The Impact of Nostalgia on Youthfulness and Implications for Health. *Self Identity* **2016**, *15*, 356–369. [CrossRef]
50. Geocities Archive. Hacker News. Available online: http://bit.ly/Hacker-News (accessed on 31 March 2022).
51. Schlosberg, J. AOL Hometown Shutting Sown, and Taking a Bit of Bronze with It. Joel's Humanistic Blog. Available online: http://bit.ly/2PJUjI9 (accessed on 31 March 2022).
52. Data Base of Archaic Web Pages (Interia City and Republika WWW). Available online: http://homeproject.pl/data-base.xlsx (accessed on 31 March 2022).
53. Król, K. Internet Domains Promoting Agrotourist Farms. *Turyzm/Tourism* **2017**, *27*, 23–29. [CrossRef]