**Chapter 3: 404 File not found: where are film promotion sites archived online?**

Keywords: Digital Archives; Archive; Internet Archive; The Webby Awards; Blog; Transmedia Archaeology

The internet is now a central location for film promotion and the film website has become the place for trailer releases, curated media coverage as well as ‘making of…’ featurettes (Walden, 2017:80). But film websites not only function as hubs for promotion, but they also contribute to film storytelling and have developed their own media-specific aesthetics, narrative techniques, and forms of audience engagement (Ibid). For film franchises, websites now routinely bridge the narrative interstices between film releases, detailing story worlds, profiling characters and providing plot primers to induct audiences and whet their appetites for forthcoming features (Gray, 2010:48). To the extent that commentators have observed the blurring of the boundaries between a film’s content and its marketing, film sites are clearly contributing to the contemporary film experience(Grainge and Johnson, 2015:4). Yet in the academic field of film studies, scholarly distinctions between commercial and creative content remain stubbornly fixed, where marketing and promotional content ( sometimes termed branded content) is regarded as a commercial activity but ‘creative’ content is lauded as art (Grainge and Johnson, 2015:5).

One probable reason for this oversight is that paratextual promotional materials have a short shelf life. Once a film has journeyed through its release windows in the cinema, onto DVD and Blu-ray formats, and streaming platforms, their job is done, they are effectively redundant and the fate of the film site becomes precarious. Studios will often lock discussion boards, so that no further contributions can be made. And finally sites are taken down and disappear from the public domain, as studio marketing divisions seek to focus attention on new film releases. The result is that film sites become ephemeral and this ephemerality jeopardises the possibility of critical appraisal or evaluation of their contribution to the film experience. Since the mid-1990s when the first film sites were created, the first generations of film sites have effectively disappeared. The result is their role remains largely unacknowledged, their development uncharted, and their history untold. It is this absence which is the concern of this chapter. In setting out to develop a better understanding of these unsung artefacts, we must first consider where film sites might be collected, recorded, and archived? And, once located, what these digital archives can reveal about the development of online film promotion?

To understand the context to this investigation, the chapter starts by looking at the development of web archiving in a general sense and what has been written about archiving and preservation. The chapter goes on to distil this literature into a set of conceptual and technical *principle*s through which to consider the nature of the digital archives investigated here. Four archives in which film sites were discovered are selected for closer examination from each of the phases of development of online repositories: the *Internet Archive*; *Digitalcraft.org*; The *Webby* Awards, and the *Movie Marketing Madness* blog (now *Cinematic Slant*). The archives’ contents are examined for what they can tell us about the development of film sites, and what is revealed about digital archives. So, to begin, it would be valuable to consider the development of digital archiving practices, to better understand both the nature of digital archives and how they are different from conventional analogue archives.

**3.1 Digital Archives**

A report compiled by the UK’s Joint Information Systems Committee (JISC) defines digital archives as collections of born-digital artefacts of interest (Dougherty et al. 2010). Archives may take the form of collections representing material of interest to the nation; an entire domain collection; a collection based on a specialist subject; or a record of a management system for commercial or legal purposes (Ibid.). Collections may be created manually, or with automation tools, and be stored either on or offline (Ibid.). But an archive should be accessible and have an interface that allows users to view objects in the archive, at first hand (Ibid.). Although the history of the development of web archives is a brief one, barely amounting to three decades, two distinct approaches have emerged to the plotting of this history. The first is proposed in an article titled ‘Unpacking Archival Silences’ (2013) by Anat Ben-David where she identifies four main phases of development and recognises the connection between web archiving practices and evolving technologies as the key driver of change. While the second approach stems from the work of Richard Rogers described in *The Sage Book of Web History* (2019) where he argues that the development of web archiving practices generates distinctive historiographical perspectives such as a biographical approach or an event-based approach. He goes on to suggest that these perspectives are hard-wired into different web archiving practices and therefore, he argues different web archiving traditions afford particular approaches to understanding the history of web archiving. What is interesting about these two approaches is, firstly their accord about the phases of this development, and secondly how their approaches yield such different insights into these emerging practices.

According to Rogers, the earliest web archives were link lists or directories, which took the form of lengthy lists of websites and their uniform resource locator (URL) addresses (Rogers, 2013:61). In the text-based operating systems of the early web, indexes were compiled by internet portals like *Yahoo.com* and *AOL*’s (America Online) *Open Directory Project*. Requests for inclusion in these indexes were decided by editors who curated sites into categories for the link lists, which were hyperlinked to the sites (Rogers, 2013:62). The *Open Directory Project* was maintained by a community of volunteer editors who curated sites and maintained their directory, ensuring hyperlinks were functional. In this way portals like *Yahoo* and *Open* *Directory Project* pioneered ways of archiving the web, developing their systems of classification as well as navigation (Ibid.). However these link lists differed in one fundamentally advantageous way from conventional library-based classification systems because , unlike a book in a library, at *Yahoo.com*, a website could be duplicated and located in more than one category, as well as be viewed by multiple people simultaneously.

From 1996 to 2000 the first generation of web archives were in some regards extensions of libraries, both in terms of method and in practice, with their emphasis on collection (rather than access), preservation and cataloguing of sites. The key exponent of this approach was the *Internet Archive*, whose founder Brewster Kahl refers to the archive as precisely that - a ‘digital library’ with its aptly named search engine technology, Alexa, inviting parallels with the mythical library of Alexandria in ancient Egypt. To address the ‘404 File Not Found problem’ (a term used by computer systems to indicate that the page requested cannot be located), one of the strategies for collection employed by the archive was a prototype form of crowd sourcing (Rogers, 2018:44). Alexa Internet made a browser tool available which took the form of a small button named ‘Wayback’ that would pulsate if the website visited had not been found and, where possible, would supply the missing page from the *Internet Archive* (Ibid.)

In exchange for this service, the user would be invited to agree to allow their internet search history to be surveyed so crawlers could visit and archive logged sites. Originally the Internet Archive’s search facility, the Wayback Machine was designed to enable searching by the site’s URL alone. Although since 2016 it has become possible to search using key words that appear in the URL or metatext too. Rogers observes how the *Internet Archive*, in effect, organises the history of the web into standalone single site histories. Search returns are delivered in a calendar and timeline format which enables the searcher to see all the dates the site was crawled and archived, as well as the frequency of the crawls. This form of organisation lends itself to a particular kind of web historiography that Rogers terms a biographical tradition (2017:161). Indeed one of the affordances of the *Internet Archive*’s approach to record is that the lifespan of a single site can be tracked in this format which has value for tracking the duration of film promotion campaigns.

Once web content retrieval tools became available from around 2000 onwards, collections could be assembled for specific research projects (Ben-David, 2013:11). Researchers began using the web as an archival resource for studying social and political phenomena from presidential elections to papal transitions, as well as natural disasters (Ibid.). Collections were made of sites focussed on specific events for studies too, such as Christine Hine’s ‘virtual ethnography’ of the Louise Woodward alleged child abuse case (2000). But the most well-known archive research based on a collection was undertaken by Kirsten Foot and Steven Schneider who tracked online articulations of bereavement following the 9/11 attacks in New York in 2001 (2005). The researchers collected 25,000 sites over a period of four months relating to the 9/11 attacks (Rogers, 2019:45). The rationale of the collection in operation at this time was the development of ‘web spheres’ - a collection of web sites which could be defined topologically, as well as temporally around an event (Rogers, 2019:44). What is innovative about this approach is that it is medium-specific and capitalises on the nature of the web with its dynamic collection approach which discovers new websites through hyperlink analysis – a principal Rogers advocates for the development of internet-sited enquiry (Rogers, 2019: 45).

Initiatives to create national level web archives began with Kulturarw3 web archive in Sweden and the Pandora and Tasmanian archives in Australia as early as 1996, shortly after the advent of the commercial web (Costa, Silva, Gomes, 2017: Ben-David, 2013:12). Over the next decade what Ben-David terms the ‘national turn’ in the development of web archiving commences and other national level organisations followed suit developing infrastructure for collection, preservation and, to a lesser extent, access (Ibid.). Although archives such as the *Internet Archive* and the *Internet Memory Foundation* seek to preserve web content from all over the world too (Costa, Silva, Gomes, 2017). In 2003 UNESCO published a ‘charter on the preservation of digital heritage’ that recognised the inherent ephemerality of born-digital cultural forms and that the risk of digital heritage becoming lost is an issue of world-wide concern. It goes on to state that the disappearance of digital information constitutes an impoverishment of the heritage of all nations and calls upon member states to secure its preservation for the benefit of present and future generations ( UNESCO, 2003). In 2010 UNESCO endorsed a universal declaration on archives stating that digital archives play ‘an essential role in the development of societies by safeguarding and contributing to individual and community memory (UNESCO, 2010).

The imperative to include digital within the remit of archiving policy is endorsed by the United Nations (UN), but in practice it is the member states that must implement the charter so this historiographical tradition takes a national perspective, in that the web is conceptualised in terms of geographical terrain, sovereignty and national institutions and it extends its existing analogue remit to keep public records and national heritage to the internet. Newly conceived national digital archive policy sought to demarcate and ‘save’ the national web beginning with what was determined to be the part of the official public record and sites pertinent to current conceptions of what constitutes the national heritage (Rogers, 2017:165).

The US Library of Congress began archiving what it termed ‘born-digital web content’ in 2000 (Library of Congress, 2018). This included sites deemed to be of national significance, social media pages and national events (Rogers, 2019: 45). In this regard the aim of these collections is to create an official national history and by this definition, as Rogers observes, the web becomes a ‘national story’ (Rogers, 2019:46). To support these initiatives, the International Internet Preservation Consortium (IIPC) was established which leads on the development of best practices for archiving and open-source tools (Costa, Silva, and Gomes, 2017). National digital cultural heritage policies were established, accompanied by legal frameworks to put them into operation.

In 2004 the British Library developed a policy to ‘collect, make accessible and preserve web resources of scholarly and cultural importance from the UK domain’ and began the collection the following year (British Library, 2016). This policy was underpinned by legislation that enables the library to acquire a copy of any digital resource published in the UK, including not only websites, but social media, and to make that copy available to the public at designated UK deposit libraries (Ibid.).[[1]](#footnote-1)

Digital legal deposit legislation was designed to enable the creation of a comprehensive collection, an authoritative bibliographic record and ease of access with the aim of archiving the entire UK web domain (Ibid.). Between 1996-2012 the focus of national web archive initiatives was on building the infrastructure for the preservation of web archives and, to date, more than 19,000 websites have permission from their owners so they can be viewed from anywhere in the United Kingdom via the UK Web Archive website (Ibid.). However, the vast majority of sites collected under the legislation, whose numbers run into the millions, can only be accessed via building-based reading rooms in the deposit libraries (Rogers, 2019:45). The reason for this apparent anachronism is the constraints of copyright law (Ibid.).

From 1997 onwards, one of the most popular forms of research across the disciplines is the use of web collections for longitudinal studies, drawing either on data extracted from existing institutional archives like the *Internet Archive*, or bespoke archives. Ben-David terms this form of web archive ‘temporal web collections’ because they are generated by periodic sampling over time to facilitate a diachronic perspective (Ben-David, 2013:22). But the main reason for custom built archives either by researchers or specific interest groups is that such archives often focus on interests that fall outside the traditional selection criteria for national institutional web archives.

The most recent form of web archive to emerge has been described as the paradigm shift resulting from the advent of web 2.0 technologies and the emergence of social media platforms and applications (Ben-David: 2013:6). Conventional conceptions of the archive were transformed by second-generation web applications by facilitating individuals to create their own archives, which Rogers characterises as an ‘autobiographical’ period of archiving (2019:46). The traditional model of the archivist as an ‘agent’ selecting artefacts for preservation, collection, and the cultural record, stands in marked contrast to the archival practices of today by individuals, afforded by storage capacity and digital communication technologies. In short, this fourth phase of development brings with it the capacity to archive oneself, particularly since social media are outside the purview of previous web archiving strategies.

From 2003, self-archiving practices often took the form of personal diaries like blogs and, with the advent of *Facebook*, *LinkedIn* and smart phone applications, a fresh archival challenge becomes evident. Unlike websites that are available on the open web and accessible without passwords, social media platforms and smart phone applications are kept behind personal login passwords. The result is these forms of communication that have become so pervasive in our culture are much more difficult to archive and impossible to make accessible (Rogers, 2017:165). Indeed, to do so would require the archivisation of the hardware as well as the communication record which brings with it a whole host of further problems as we shall see when we look at the Digital Craft archive initiative.

Social media platforms have not developed a consistent approach to the question of their preservation (Rogers, 2019: 48). *Facebook* and *LinkedIn* regard themselves as social media networks rather than content sites, so they do not permit the crawlers deployed by the *Internet Archive*(Ibid).[[2]](#endnote-1) While user generated sites like *YouTube* and *Pinterest* do allow archiving of their material and *Twitter* actively encourages its archivisation by donating historical tweets to the Library of Congress for storage (Ibid.). However on an individual level, it has been argued that the practice of archiving has been fundamentally redefined from ‘the practice of a social agent’ to a ‘social practice’ (Uricchio,2009: 144). Anyone with access to the web can become an archivist in the age of social media and it is these developments in archival technologies that will shape what we can know in the future. So, what are implications of the development of web archiving for a historical understanding of online film promotion? To address this question, this chapter now moves on to considers how archives and archivisation have been conceptualised by some of the key theorists in the field to see what insight these writings may provide.

**3.2 Conceptions of Archives**

Returning to Michel Foucault’s proposition that our understanding of the past is shaped by the ‘present state of knowledge’ outlined in the previous chapter, it is interesting to note that he uses the term ‘archive,’ not to describe a physical repository of historical artefacts, but to explain what is known about a subject (Foucault, 2002:5). Foucault refers to this as a ‘system of discursivity’ with all its accompanying enunciative possibilities and prohibitions (Foucault, 2002: 145). For Foucault, an archive is the first ‘law of what can be said’ on a specific subject and to illustrate, he points to the way knowledge is organised by subject and how their discipline’s accepted methodologies determine what lies within and what lies outside the concerns of the subject (Ibid. : Foucault, 2002:14). Promotional resources like film sites are regarded as commercial materials of little intrinsic value and, for the most part, located outside film studies’ definition of what is worthy of scholarly attention (Grainge and Johnson, 2015:4). As a result, since the advent of the internet film sites have, by and large, been disregarded and there has been little critical evaluation of what they contribute to the film experience. Clearly the current state of knowledge in film scholarship plays a significant role in determining what the future can know about the subject and raises wider questions about how historical knowledge is constituted.

The relationship between the past, present, and future is of interest to Jacques Derrida too. While Derrida’s work was published in 1995, right at the beginning of the period under consideration in this book, at the time when the internet entered the public domain, he is concerned with physical archives. Derrida recognises that archives play a role in determining what can be known about an artefact. In *Archive Fever* (1995) he suggests that, while archives may be concerned with the preservation of the past in the immediate sense, in effect, they constitute an ‘open letter’ to the future (40). In this respect he is in accord with Foucault that whatever is archived in the present conditions what is remembered in the future. Derrida writes: ‘the question of the archive is not ... a question of the past. ………It is a question of the future…. the question of a response, of a promise and of a responsibility for tomorrow’ (1995:36). So, by this definition, archival preservation of film sites is a critical precursor to any form of evaluation of the artefact itself.

Derrida traces the term ‘archive’ back to its Greek antecedent, *arkhe* and he identifies the three main principles that define archival practices as *commandment, consignment,* and *commencement* (1995:1). To explain, *commandment* is the point at which the archive maker (*archon*) is recognised for this undertaking, and receives authorisation to undertake the task (Derrida, 1995:2). This means the archival process is socially sanctioned, culturally validated and, by implication, whatever resides in this archive is confirmed as a legitimate representation of the past (Ibid.). The process of *consignment* takes place when decisions about what will be selected for preservation in the archive are made, along with decisions about the processes of identification, collection, ordering and classification of artefacts (Derrida, 1995:3). While the third process of *commencement* recognises that one of the basic motivations of the archive is to establish the authority of provenance (Derrida, 1995:2). Derrida is interested in how these processes impact on the artefact being archived (1995:18). For example, at the most basic level, the archival process transfers an artefact into a new (physical) space and assigns it a different identity and value from its original function. Archives have the effect of establishing confirming the worth of the artefact, as well as its value as a ‘historical truth’(Derrida, 1995:59). So, Derrida observes ‘archivisation produces as much as it records’ (1995:17).

Digitisation has brought with it a renewed interest for scholars in the material conditions under which media operate. Friedrich Kittler pointedly describes media as ‘machines’ rather than ‘texts’ to underscore this (1990:233). He asserts that one of the consequences of digitisation is that media characteristics have, in effect, become ‘surface effects’ providing an interface for human interaction, but the real meaning of digital media is its facility for storage(Kittler, 1999:1). In this regard, Kittler shares Foucault’s interest in excavating the conditions in which knowledge arises (1990:5), and concludes that media ‘produce what it allegedly only reproduces’ which suggests that any examination of archives must scrutinise not only its artefacts, but also its technology, design, and architecture (Kittler, 1999:145).

Like Kittler, Lev Manovich maintains that storage capacity is the basic attribute of an archive (2001:220). In *The Language of New Media* he explains how, in essence, all digital media are primarily a collection of data files and are therefore best understood as databases (Manovich, 2001:218). He notes that it was in the mid-1990s when the storage facility of the ‘universal media machine’ became available, that the incentive to archive began (2001:224). He observes ‘everything is being collected: cultural asteroids, DNA patterns, credit records, telephone conversations, it does not matter’ (Ibid.). Indeed, Manovich hypothesises that data indexing will become a cultural pastime, or even profession in the future (2001:225).

Two decades later, Manovich’s observation has been borne out, as can be seen in the proliferation of platforms with archival architectures on the internet such as *Flickr* (2004), *Facebook* (2004)*, Tumblr* (2007)and *Instagram* (2010), and the inference of these developments is the de-institutionalisation of the archive. Now an archive can be created independently, and *archonic* power assumed rather than formally ‘commenced,’ in the Derridean sense of the word, even though its value may only be to the individual, rather than, culturally or institutionally recognised. In turn, this challenges assumptions about which cultural forms should be regarded as culturally significant which has bearing on ephemeral forms like online film promotion.

In essence, what these writings indicate is that archives are basically characterised by three fundamental processes: storage, system, and selection. Besides the physical storage of the data/artefacts collected and preserved for the record, all archives, whether digital or analogue are driven by intention or rationale, and therefore archival collection processes are inherently systematic too. Following on from this, archival collections are always selective because they are based on a set of beliefs about what will be (and, implicitly, what will not be) important in the future. Furthermore, while all collections are subject to boundary restrictions, selection infers curation and value conferred upon the object through this process as well as the creation of a place in which to situate the collection.

As archives undertake these processes of storage, systematic collection and selection, the principles and purposes that define the nature of the archive and determine its usage become apparent. Moreover, these principles and purposes do not just inform the practical level of organisation, as implicitly archiving is also a historiographical process too. Just as a library is not just a collection of books but a system by which books can be found, archives are defined by their systems and their processes and procedures, as much as their contents, and these processes are part of the meaning of an archive (Russell. 2012).

A further observation worth making is that archival principles and purposes are always based on a view of the world, and a conception of their future uses and values, and in this way, they are discursive in the Foucauldian sense. In the creation of archives, there is always going to be an evaluation of the balance between effort and worth. In the past, few archives have been interested in film promotional ephemera because ephemera were not regarded as having any value. Furthermore, the ephemerality of contemporary film promotion is complicated by transmedia productions which have blurred distinctions between promotion and content. While some components of a production remain available, others, like promotional sites may disappear because transmedia paratexts have not been regarded as worth preserving. Thus, rendering the archives of transmedia production only partially available for future scholarship.

If digital files are by definition a form of archive, it may be useful to make a further distinction about intention and how archives are created from deliberate to inadvertent collection. A deliberate collection sets out to collect a specific artefact, and selection criteria determine what is collected. Whereas an inadvertent collection may have a broad remit to collect material, and coincidentally create an archive-type collection, or even have no remit and inadvertently collect an artefact. And finally, deliberate collections may be subdivided into two further categories: the universal and the particular - universal in the sense that all published artefacts are collected, and particular in the sense that selection is determined by the criteria for collection. Having established these principles and parameters of archivisation, these features can form a framework which can be used to appraise and evaluate the archives selected for closer consideration.

To get a clearer sense of the various kinds of archival settings where film promotional sites are archived online, Ben-David’s phase-based mapping of web archives described in the Digital Archives section provided a guiding framework for the next stage in the investigation. One archive from each of the four historical periods was identified so that a cross section of digital archives could be examined. The first case study is an ‘index’ archive and is probably the best known - the *Internet Archive* established in 1996. The second case study is a digital cultural heritage policy-led initiative by the Museum of Applied Art in Frankfurt, Germany to create a collection of digital-born artefacts known as *DigitalCraft.org*. The third case study is the *Webby Awards* that have been celebrating digital creativity and, specifically for my purposes, film sites since 1997 (Webbys, 2020). And the fourth case study is a blog called *Movie Marketing Madness* (now *Cinematic Slant*) started by an individual enthusiast, Chris Thilk in 2003. Having identified a cross-section of distinct kinds of archive, the next task is to look more closely at their archival practices, and how they can be measured against the principles identified in the archive writings.

**3.3 Web Index Archives – *Internet Archive***

Established as a non-profit organisation by the computer scientist and internet entrepreneur, Brewster Kahle, the *Internet Archive* is widely regarded as the most complete archive of the internet to this day. Originally it was set up to preserve historical collections in digital formats, its web archive has become its biggest collection and currently numbers 826 billion web pages (Internet Archive, 2023).[[3]](#footnote-2) The archive is compiled by automated domain archiving, using software tools that periodically download all publicly accessible websites, and present ‘snapshots’ of these sites as an archival record (Internet Archive, 2017). So, in this sense the *Internet Archive* is not selective, but it is systematic. Although the archive maintains a 6-month embargo on collecting new sites, so time-sensitive content like marketing and promotion may be particularly vulnerable as it may be taken down before it has been recorded (Barone, Zeitlyn & Mayer -Schoberger, 2015).

In Kittlerian terms, a collection of materials only becomes a true archive when it is searchable by its users and data can be retrieved. The archive’s search engine, known as the *Wayback Machine,* was not developed until 2001 and unlike Google’s text-based search facility, in the first two decades of its existence the *Wayback Machine* search engine searched by web address as indicated by the ‘http’ prompt in its search box (Rogers, 2013:65). Limited index information indicates that the archive is set up primarily for collection, and access was clearly an afterthought because without knowing a site’s URL address, searching can be challenging. Since 2016 this search facility has become more user-friendly and it is now possible to search using key words from the URL or metatext. But from the perspective of this chapter’s concerns, the *Internet Archive* is an inadvertent archive, rather than a deliberate archive of film promotion sites.

Since 2010 *Wayback Machine*’s search returns have been presented in the form of a calendar (see Figure 3). Along the top of the screen, there is a bar chart spanning from 1996 to the present day, and the ‘snapshots’ taken by the *Internet Archive* during that year year are identified in the bar chart. Moving the cursor over a specific year, brings up a year’s calendar with the days illuminated by small blue circles indicating when the site was recorded. Moving the cursor over a blue circle brings up a record of how many times the site was captured on the day. An exploratory search for the film *Event Horizon*’s (Paul Anderson, 1997) promotional site via the *Wayback Machine* yielded an interesting example of an early site (see Figure 3).

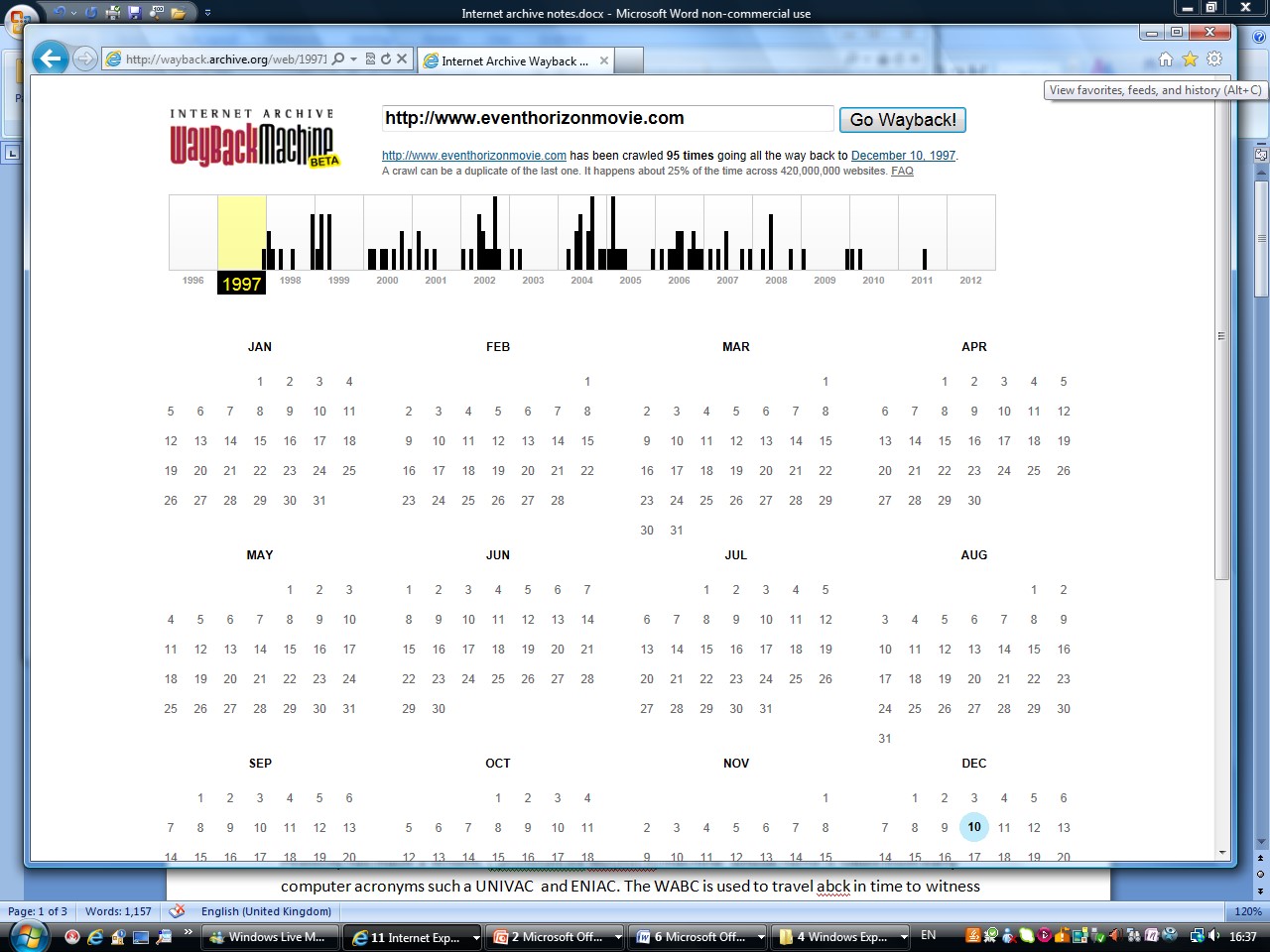
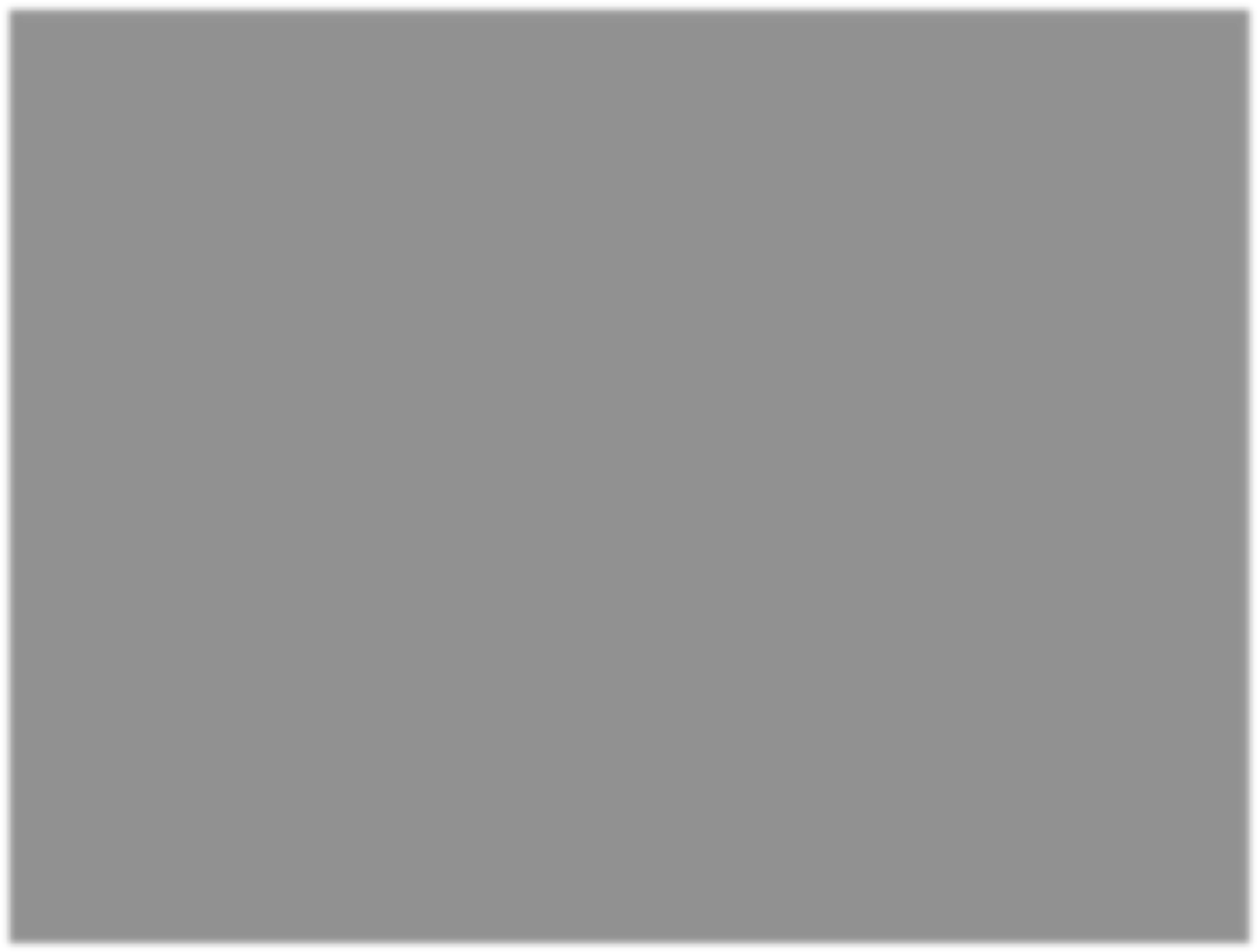


Figure 3. Wayback Machine search returns delivered in calendar format © Internet Archive

The site’s landing page advises visitors about its free-to-use plug-in software for different computer operating systems. But the site’s intended audience is not clear. From one perspective the site seems to be addressing film journalists with electronic press kit (EPK)-style information including pages of production notes and a full set of credits. While from another perspective the site addresses its prospective audience inviting them to enter the film’s in-movie storyworld for a ‘mission briefing’ and visitors are instructed to get their helmets on and ’be ready for anything!’ as well as plough through an extensive essay presented in a tiny font on the scientific phenomenon of black holes in space. These different registers address both the press and public audiences at the same time and the preservation of sites like this offers the possibility of encountering early design formats for in-movie worlds, first-hand as well as emerging web social protocols. Clicking on the site’s ‘cockpit’ link however brings the whole experience to an abrupt halt because it does not work. The hyperlink is broken as the page content (assets) has been moved or deleted. When the link destination changes or is deleted, the connection between one page of a site and another ceases to exist or defaults to the current site, creating a link between the past and the present that internet historian, Richard Rogers describes as creating ‘jump cuts’ through time (2013:66). In other words, the archive privileges spatial navigation of the web over temporal coherence and the effect on archived sites is that they become connected to versions of themselves across time which Rogers describes as a ‘multi-temporal archival experience’ but the upshot is that the site becomes a jumble of webpage fragments (2018:44).

Following this initial exploration, I repeated the search for several other film sites from the same year including *Men in Black* (1997), *Titanic* (1997) and *Alien: Resurrection* (1997), and the problem is evident in them all. The *Wayback Machine* had archived the core site but was unable to archive all its assets. So, while they have not disappeared entirely, these sites hardly constitute a complete or even stable object for the researcher. Moreover, because of the nature of algorithmic code, what appears on the screen cannot be understood as the ‘original’ artefact either (as discussed in chapter 2) because its code and data are different with each iteration, bringing into question the very notion that digital materials can be archived (Rogers, 2017:163). So, this exercise brings into question the conventional understanding of an archived document in the online context.

While Kahle’s ambition is for the *Internet Archive* to become a universal collection, in practice it has become an enormous collection of fragments (Kahle, 2007). The archive has vast storage capacity, but the nature of the object it seeks to preserve eludes complete collection and preservation. This is because the conception of websites is fundamentally flawed as they are not standalone entities but rather are made up of material from elsewhere. Furthermore the *Internet Archive*’s crawlers can only, *de facto*, take snapshots of surface content in the searchable web which only represents a fraction of the web (Wright, 2009). Moreover, the *Internet Archive* respects robots.txt files, which site owners can apply if they do not want their content archived. All these factors illustrate Kittler’s point that the materiality of technology needs to be understood to comprehend its limitations as an archival setting (Kittler,1999:145). However, Rogers advises that despite its limitations if the researcher follows the *Internet Archive*’s modus operandi, it can be an insightful tool and retain ‘archonic’ authority (2013:68).

The main outcome produced by this archive is a mode of historiographical information charting the lifecycle of the website over time. The *Internet Archive* and the *Wayback Machine* have organised the history of the web into a collection of website biographies (Ibid.). By so doing the *Internet Archive* illustrates Derrida’s third archival process of *commencement* that establishes the provenance of the artefact by charting its lifespan (Derrida, 1995:2). Archaeologists make a distinction between *use-lives* and *life spans* and suggest that once the former is over and it is no longer used, an artefact becomes ‘archaeological’ (Tringham and Ashley, 2015). What the *Internet Archive* is able to demonstrate is how far the latter exceeds the former in the case of film promotional websites. Media archaeologists can use the *Internet Archive* to understand version histories, user activity as well as website design conventions. Recently, the *Internet Archive* developed a suite of data analysis tools which identify the various kinds of image files, text and applications that make up a site. So, for example, the 1998 site for *Alien: Resurrection* consisted of 75% text to 25% graphics, whereas the site for the next *Alien* film, *Prometheus* in 2012, contained just over 30% text with far greater percentage of graphic content illustrating how film site design has evolved during the lifetime of the franchise.

**3.4 Digital Cultural Heritage Policy Archives – *Digitalcraft.org***

From the mid-1990s, national libraries and museums began developing digital cultural heritage policies (Ben-David, 2013:14). The *Museum für Angewandte Kunst* (Museum of Applied Art) in Frankfurt, Germany whose remit is to conserve and preserve applied arts and craft was concerned that despite the impact that digital technology was having on all sectors of society, there had been little recognition of this form of media production as a craft (Institute of Network Cultures, 2014). As a result, *Digital Craft.org* was established in 1999 to develop a collection of online communities, games, emulators, and websites for the museum (Ibid.).

This archival initiative set out to tackle the numerous technical challenges such as platform obsolescence and software updates and the project’s director, Franziska Nori, decided that the only way to create a permanent collection of web design and ensure full functionality was to make offline versions of the digital artefacts (Ibid.). The project had to take into consideration a number of archival challenges, not least of which was questions of legality. Domain owner permissions and copyright holder clearance was sought as copyright law was not designed for digital artefacts (Ibid.). An exhibition of web-based artefacts could be licensed for display in the museum, but not on the web, because copyright concerns not only the artefact, but the commissioning partner, as well as the designer. Moreover, some websites incorporated add-ons which were licensed separately (Ibid.). Even housing the collection online turned out to be prohibited, as web publication was too broad to be covered by a licence agreement and so, ironically, like counterparts at the British Library and Cambridge University Library, this collection of web-based artefacts was confined to a building-based exhibition in the Frankfurt museum (Digital Craft, 2003b). A situation that is inherently contradictory, when in the post-geographical context of the internet, archived web material is only accessible ‘on site’ in a fixed location.

Unlike the *Internet Archive*’s automated collection, for *DigitalCraft.org* the question of what to include in the collection was deliberate (Institute of Network Cultures, 2014). Although film sites were not the sole focus of the collection, within the database there were eight notable film sites including sites for Darren Aronofsky’s *Requiem for a Dream* (2000) and Christopher Nolan’s *Memento* (2000) (Ibid.). Based on the host institution’s conception of design craft, the selection criteria included the originality and uniqueness of concept; the quality of visual representation; usability; content and context; technical innovation and inventiveness in navigation. Together these criteria illustrate Derrida’s concept of ‘consignment’ where decisions about what is selected for preservation are taken according to specific initiatives and the nominated institution’s agenda (1995:3).

The wider challenge was to ensure an archiving system that accorded with international museum description standards advised by the International Council of Museums (CIMI), Dublin Core and the Art Museum Image Consortium (AMICO) (Digital Craft,2003a). To do this the Digital Craft team developed inventory descriptions recognising the features of digital artefacts including programming languages and plug-in software (Digital Craft,2003c). Lastly, catalogue information outlined what was distinctive about each artefact as web design illustrating how the museum’s curatorial processes infer appraisal and confer value upon the object (Institute of Network Cultures, 2014).

The complete archive consisted of two collections: a building-based collection of fifty websites housed on the project’s servers and an online database of one hundred websites from different fields (Ibid.). Catalogued entries included provenantial information about the site’s designer, copyright holder, country of origin, date of entry into the collection, programming language, as well as notes of plug-ins to view the site’s full operation and a link to the live web version (Digital Craft, 2003c). The catalogue also provided a description of the ‘experience’ of the site and outlined how the site operated in relation to the film (Ibid.). In this way, the collection fulfilled the first of Derrida’s archival processes - *consignment* when selection and systems for storage were created, together with processes of identification, ordering and classification (1995:3).

However, while the *Digitcraft.org* archive set out to address the challenge of preserving digital artefacts, the initiative had its limitations. To be stored in the collection, sites had to be stripped of components deemed unsuitable for server storage such as databases, competitions, and chat functions (Institute of Network Cultures, 2014). An approach indicative of the project’s pre-web 2.0 mindset which regarded websites as stand-alone objects. The artefacts were further compromised by the fact that sites were archived without other components of the promotional campaign like trailers, posters, and TV spots, or indeed the film they were designed to promote. By archiving sites like this, the artefact’s meaning and value is pared down to that conceived by the institutional agenda which in this instance regards film sites as digital craft, disregarding the artefacts role as film promotion or expanded forms of storytelling. But this does illustrate Kittler’s assertion that storage media are not just physical repositories but discursive artefacts too, and points to the way all archives play a role in determining what we know (1999:145).

Just three years after this pioneering project began, a new incoming Museum Director reversed the policy decision to maintain this collection, closed the department and dismantled the building-based collection (Nori, 2012). Nowall that remains is the online catalogue hosted by the City of Frankfurt’s server to maintain public access and whilst the project’s Director still maintains a *Digitalcraft.org* email address, this case study illustrates that archives can be ephemeral too (Ibid.). To all intents and purposes, when funding was withdrawn and the museum severed its ties with the collection, this archive lost its ‘archonic’ authority and could be regarded as a failure. However, for the purposes of this book’s examination of the archivisation of digital artefacts, there is much to glean from this case study. What was clear was that the process of archival selection was based on a set of ‘archonic’ conjectures about what will be important in the future and the truth is that it is always in contention. So, consequently the archival process was shaped by institutional agendas and politics that determine what people in the future may know about the past.

**3.5 Temporal Web Collection- The *Webbys***

The next archival case study is what Ben-David terms a ‘temporal web collection,’ that is to say either a bespoke collection or archive generated from existing resources, by periodic sampling over time (2013:22). Web design awards have proliferated over the last twenty years and these awards often hold a ‘hall of fame’ commemorating award winners and while this is not their primary function, these archives are an inadvertent consequence of a deliberate process. Moreover, some of these awards have systemically created a record of their winners which may provide the basis of a longitudinal perspective on the development of film sites (Ibid.).

A survey of web design awards revealed that the *Webby Awards*, established in 1996, are the longest running awards of their kind. Hosted by the International Academy of Digital Arts and Sciences (IADAS) the *Webby Awards* are modelled on the Oscars for film and seeks to honour ‘the best of the internet’.[[4]](#footnote-3) A by-product of this annual contest is its Archive and Gallery, which is, in fact, a searchable database of award winners spanning more than twenty-five years to date, earning its self-declared recognition as ‘a capsule of internet history’ (Baio, 2012). While the *Webby Awards* is an unintentional archive, its processes of appraisal and judgements by expert panels and popular votes are both selective and systematically judged. Moreover, by adopting the term ‘archive’ the *Webby Awards* has become a self-appointed custodian of web design history and the collection assumes the Derridean authority of a record (Walden, 2017:84).

The *Webby Awards’* category for‘movie and film’ covers ‘sites dedicated to promotion, celebration and presentation of films, movies and film culture.’ From 1996 to 2004, film site winners and nominees were known only by the title of the film they promoted, rather than as cultural artefacts in their own right. However, by 2005 there was a shift in the mode of address, and websites begin to be attributed to design agencies (albeit in small print with no hyperlinks) and explanatory notes on the design process became an additional feature of the listing. In 2012 credits were given to designers in the winners and nominees lists, and by 2017 many agencies were archiving their own work and the *Webbys* link to agency sites where their web design is displayed with extended commentary.

A developmental arc of this new form of promotion is discernible in the *Webbys winner’s* archive. Over the period film websites transform from avant garde net art sites, such as Richard Kelly’s *Donnie Darko* (2001) and Darren Aronofsky’s *Requiem for a Dream* (2002),to fully integrated transmedia events such as the sites for franchises like *The Hunger Games* (2012-15) which can be operational over the lifespan of the franchise rather than specific films. Web site conventions begin to emerge such as the franchise ‘universe’ in *Jurassic World.com* and *StarWars.com* (Webby winner in 2017 and 2018). Another convention to emerge is web-based counter narratives (in which film narratives are told from different perspectives) for character-driven films such as *The Simpsons* (2007), and *District 9* (2009); While website ‘experiences’ provide a foretaste of the film such as *Sully: The 208 Seconds Experience* (2016) or *Independence Day: My Street* (2016) where visitors type in a street address and watch how the site uses Google Street view imagery to generate fake news footage of an alien bombing attack of the location. What is evident, even from this snapshot of sites, is that websites have become an increasingly integral component of the film narrative and the contemporary film experience.

The *Webbys* Archive and Gallery claims, ‘if you want to see the past, look in here,’ but they are unable to provide a comprehensive record, as the awards’ process of selection has the effect of filtering the present and claiming value for selected nominated works (Webbys 2016). There are notable gaps and absences (Walden, 2017: 84). According to *Variety*, in 1999 *The* *Blair Witch Project* site made such an impact on film audiences that it prompted a shift in industry attitudes towards the internet as a site for film marketing and promotion, yet the site is conspicuous by its absence from the *Webbys* archive (Hayes and Graser 2000). Similarly, *The Beast* was significant as it was widely regarded as the first ARG (Alternative Reality Game) designed to promote Steven Spielberg’s *AI* *Artificial Intelligence* (2001), yet it does not figure in the *Webbys* ‘Hall of Fame’ either (Walden, 2017:84). Likewise, much has been written about 42 Entertainment’s alternative reality experience for Christopher Nolan’s *Batman Begins* (2005) but it does not make the archive either. Such absences remind us that, contrary to its claim, this archive cannot enable its visitor to ‘see the past’ - only an edited version of it (Webbys, 2016).

In 2005 when ownership of the *Webbys* changed hands the number of itsaward categories was increased significantly and entry fees were introduced. Subsequently the *Awards* have been widely criticised for excluding ‘worthy candidates’ and leaving the field dominated by well-resourced film studio marketing divisions (Baio, 2012). Whilst this has clearly become the case, close reading of these awards can still yield insight into how awards shape the development of film sites . Awards proclaim their winners as models of excellence for adoption or adaptation. *Webby Award* winners, nominees and honourees consecrate certain ways of doing things and as a result models of design are copied and become conventions. Foucault asserts that ‘monuments’ (like awards) are discursive entities that tell us as much about the factors that gave rise to their appearance, as the past (2002:155). So, in the absence of other critical discourse such as journalism or academic commentary, these awards function as discursive practices. As a result, awards may not only validate existing film website designs, but shape the future forms they will take.

**3.6 Social/Participatory Web Archives –*Movie Marketing Madness* Blog**

The fourth archive represents a paradigm change in online archiving practices resulting from the technological development of the social or participatory web (Ben-David 2013:6). The *Movie Marketing Madness* blog was established in 2003 at a time when blogging was the new state-of-the-art form of web publishing (Walden, 2017:85). A ‘blog,’ a term created through a conflation of web and log, enables individuals to compose and share a log of content and, while structures vary, a blog’s principal component is the ‘post’ – that is a piece of written commentary, image or video linked to other sources of information. Blog-style archives are inherently archaeological as they store collections of writings in reverse chronological order, with the most recent posts published at the top of the homepage and by scrolling down the reader can delve into past posts in the blog’s archives (Rettberg-Walker 2008:8). Structurally speaking, the blog’s archive is an index to past posts, organised by date, category and lists of links, similar to the first web archives (Ibid.).

*The Movie Marketing Madness* blog illustrates the importance of human actors in digital archiving initiatives. Freelance writer and content strategist, Chris Thilk began posting on the Blogger platform, a free weblog publishing tool owned by Google and hosted at Blogspot.com, in 2003 as a past time. This kind of activity has been described as ‘techno-volunteering’ - that is to say, a self-appointed individual who chooses to devote time and effort, as well as their own finances to the maintenance of an online resource (De Kosnik, 2016:41). Initially, Thilk had no set approach to writing about film promotional campaigns but by the end of the year he had developed a clear editorial policy, that he describes as a ‘manifesto of intent,’ illustrating Manovich’s prediction that data collection will become a popular recreational pursuit (Thilk, 2005a; Manovich, 2001:225). Thilk’s *modus operandi* was to focus his attentions predominantly on marketing campaigns for Hollywood movies and on a weekly basis he selected one campaign and gave an account of its components, including trailers, posters, online and social advertising, cross-promotions, media, and publicity (Thilk, 2005).

In time, *Movie Marketing Madness* garnered attention from the film studios and the blogger became an ‘influencer’, attaining what Derrida refers to as *commandment,* whereby the archivist is recognised for their undertaking and receives a kind of ‘authorisation’ from the industry *post hoc.*[[5]](#footnote-4) As Thilk set out to write a blog not create an archive, his blog is an advertent rather than deliberate archive; but I would argue its longevity means that it has become a unique record of online film promotion.

The blog provides links to film sites and these links often suffer the same fate found in other online archives. But the value of this archive lies in its insight into the ways in which film promotion sites have developed over the period from EPK (electronic press kits) formats to more integrated film marketing experiences today. Blog posts chart evolving forms of engagement between producers and audiences, as producer’s ‘official’ film sites link to fan sites, and other social networking sites from *Friendster* to *Facebook*. The blog also illustrates how film websites formats have had to be mindful of different Internet access speeds and software requirements, as well as the incorporation of changing plug-ins and applications such as *Flash*, *QuickTime*, *Google* maps, and later to social media and mobile apps (Thilk, 2017). Clearly the use of these technologies in film promotion illustrates Kittler’s assertion that ‘media produce what they allegedly only reproduce,’ with the shift from what are known as stand-alone ‘destination’ websites to socially integrated multi-media entertainment sites actively seeking to generate audience communities through technical affordances (Kittler, 1999:145).

The most interesting blog posts are those that illustrate emerging film site conventions. Horror films seem to contain elements designed to startle and surprise. For example, the viral campaign for *Resident Evil Apocalypse* (2004) in which mobile phone users received hoax messages saying they had been infected with a ‘T-virus’ and were directed to an online destination which turned out to be the film’s official site (Movie Marketing Madness, 2004). Or the site for Quentin Tarantino and Robert Rodriquez’s double bill, *Grindhouse* (2007):on arrival at the site, when the image loads, the ‘film’ breaks and burns revealing the entrance to the site where the visitor can grab a machine gun and ‘shoot up’ the lobby (Movie Marketing Madness, 2007). While promotion for science fiction and fantasy films appears to be often accompanied by in-movie website worlds that take the form of fictional institutional or commercial sites, as can be seen in the ‘Paranormal Studies lab’ website for *Ghostbusters* (2016), ‘The Institute for Human Continuity’ for Roland Emmerich’s *2012* or Weylandindustries.com for Ridley Scott’s *Prometheus* (2012). Instances like these illustrate how online film promotion has expanded its role to include story world exposition.

Thilk’s blog provides a unique contemporaneous commentary on film promotion campaigns and an opportunity to track how websites have changed over time (Walden, 2017:85). However, inevitably, an archive created by one individual has limitations too. Coverage is confined to English speaking, mostly American films, and sites from other parts of the world or in languages other than English barely figure in the blog at all. This means that in practice the campaigns under consideration tend to spring from a limited number of Hollywood studios and marketing agencies.

But the most significant aspect of this archive is that the blog redefines conventional understanding of who undertakes archival practices. Archiving is now widely regarded as a profession and the work of archivists is undertaken by trained practitioners who hold post-graduate qualifications awarded by accredited professional bodies and universities. But in this instance, Thilk writes this blog as a labour of love and assumes the authority of Derrida’s *archon* in a manner that is akin to the tradition of nineteenth century independent archivists. When individuals like Lt-General Augustus Pitt Rivers, founder of the Pitt Rivers Museum in Oxford, simply assumed archonic authority to set up an archive of artefacts related to his fields of interest (Ibid.). It makes sense that if archiving becomes an everyday practice because of the affordances of the web, then what gets archived will change too. From this perspective, individual interests will change what can be known in the future about the past. What this example clearly shows is that cultural memory is longer solely under the control of institutions but is subject to what has been described as ‘distributed logic’ - a new generation of web-based archives (Uricchio, 2009:142).

So, having established where film websites have been collected, recorded and archived, and who has been undertaking these archival practices, as well as getting a sense of how each of these archives shapes the preservation of film websites, and what we may know about film websites, in the next section a transmedia archaeological survey is undertaken to examine some of these archival sites and to explore their capacity for collection, recording and preserving of film websites, in practice.

### 3.7 A Transmedia Archaeology: searching for District 9’s promotional site, D-9.com

The *D-9.com* promotional campaign for Neill Blomkamp’s *District 9* (2009) was selected as the focus of this archaeological investigation as it is a highly regarded site that has won numerous accolades at *The Hollywood Reporter’s Key Arts Awards*, the *Webby Awards* and the W3awards, among others. It has attracted attention further afield too when Jason Yim, founder and Executive Creative Director at *Trigger Global* who headed up the campaign, gave a presentation at the American Film Institute’s *Digifest* he told the audience the campaign started at Comic Con 18 months before the film had its cinema release and indeed before the film had been shot (AFI, 2009). The site has also garnered academic attention when the transmedia scholar Henry Jenkins wrote about the campaign on his blog (AFI in Walden, 2017:83).

Neill Blomkamp’s *District 9* is a science fiction set in South Africa where a spaceship full of alien creatures has landed and are being contained in refugee-style camps in a slum district on the fringes of Johannesburg. *District 9*’s transmedia promotional campaign consisted of four websites under the title *D-9.com*, as well as several videos on *YouTube*, a telephone answering service, various games, billboard advertising and even a ‘guerrilla advertising’ campaign of signs affixed to park benches in America’s biggest cities proclaiming, ‘Bus bench for humans only’. The campaign was the first to have an AR ( augmented reality) experience and all the main characters had social media profiles on *Facebook* and *Twitter*.

The central plank of the D.9 campaign was to invoke the story world as an ‘as-if-real’ extension of the film, deploying a similar approach to the *Blair Witch Project* site. Providing an entry point to the film’s story world was a corporate website for the company Multi-National United (MNU) whose job is to contain the aliens. The site was designed to be interactive, engaging audiences in the story world with job vacancies to apply for, online training programmes to complete, and featuring a community watch site mapping the district to monitor alien activity. This illustrates what Jenkins identifies as one of the key principles of transmedia storytelling, allowing audiences to learn more about the film’s story world and interact with it as if it is real (2009c). But what made this site particularly engaging though, was that there were two versions of the site: one for humans and one for aliens.

The story world was manifest on *YouTube* too with a fictional corporate video and public information-style videos on how to deal with encounters with aliens. As well as fulfilling an expository function mapping this story world, a further ingredient was added to the transmedia campaign. A separate site called *MNU Spreads Lies* reported events from the alien’s perspective and thereby constructed a ‘counter-reading’ to the narrative of the film, illustrating Jenkin’s transmedia storytelling principle of subjectivity, whereby audiences can see the story world through the perspective of an alien character, Christopher (Jenkins, 2009c). This ‘as if real’ site contradicts the other on-line material, as well as the film itself and, by so doing, opened up space for ambiguity about how reliable the information on the ‘official’ site (and indeed the film) might be (Jenkins in Walden, 2017:83). All told, this transmedia campaign distinguished itself from conventional film marketing by elaborating on the film’s narrative through storyworld building, as well as articulating character subjectivities and, as a result, extended the pleasures of the fiction across media platforms.

So the credentials of the site as an interesting example of transmedia marketing and promotion were clearly established, however, examination of the site was halted when in May 2013, *D.9.com* was taken down from the Sony Picture’s studio site leaving only the film’s cinema trailer, and was replaced with an online ‘shop window’ displaying DVD and Blu-ray formats for sale (Ibid.). This brought the research project to an abrupt halt and illustrated the perils of researching ephemeral media artefacts like film promotion sites. When I began looking further afield on the web, it became clear that the studio’s digital housekeeping had not completely erased *D-9.com* from the web (Ibid.). On *YouTube* the in-world videos were still available to view, and the fan blog *Movieviral.com* had created compilations of videos on its *YouTube* site (Ibid.). The film’s *Facebook* site was still live too, containing clips of deleted scenes and ‘Making of’ features from the time of the film’s release. The *Facebook* pages for the main characters - Wikus Van der Merwe and the alien, Christopher - were still in place too, although they contained little more than a few pictures, and all references to the website were gone (Ibid.). Over time the Facebook pages became populated with advertisements for forthcoming films of a similar generic disposition, and were eventually re-purposed to promote Blomkamp’s next film, *Elysium* (2013) (Ibid.).

The search for *D-9.com* moved on to the *Webby Awards* as the film site was the recipient of an award in 2009 (*Webby Awards*, 2009). However, in the event, I discovered that whilst the awards hold a thumbnail image of the site and provide a link to it, the archive does not contain the winning film site itself (Ibid.). While the *Webby Awards* promise visitors to its gallery that they will be able to ‘see 19 years of the web’s best’, in practice, it cannot deliver on this promise because these artefacts do not reside in this archive (*Webby Awards* 2016). In practice, the archival record can only provide links and these may be moved, reorganised, or deleted, leading to ‘link rot’ or worse still, being taken down and disappearing altogether (Walden, 2017:84). As the *Webbys* do not hold the copyright of these film sites, they cannot host the sites as these assets remain the property of the film producers (Ibid.). So, after initially looking quite promising, closer examination revealed that, for the most part, the Webbys is little more than a list or record of the winners which in semiological terms reduces this web-based archive’s status from indexical connection to the past to just an iconic one. So, *de facto*, it operates more like a catalogue than an archival repository in the conventional sense of the word.

The next archive that promised the possibility of providing access to the *D-9* site was the blog, *Movie Marketing Madness*. However, this archive proved to be a disappointment too, as it turned out that in August 2009 the blogger, Chris Thilk did not choose to cover the marketing campaign for *District 9* (Walden, 2017:85). Although the blogger does not explain this particular omission, given his stated editorial policy, there are several possible reasons why *District 9*’s campaign was not covered (Ibid.) Firstly, its release coincided with the release of Quentin Tarantino’s *Inglourious Basterds* (2009) and this was the blogger’s preference as it was simply a more high-profile film (Ibid.). More broadly , with the advent of *Twitter* at this time, the digital landscape was changing. Thilk observed how visitor traffic to the site had declined, as RSS traffic increased enabling subscribers with updates via email, without having to visit the *Movie Marketing Madness* site for themselves (Thilk, 2009). In other words, ways of accessing websites were changing from audiences visiting ‘destination’ sites, to notifications sent to readers email addresses, and this resulted in a slowdown in the blog’s traffic and Thilk’s posting activity for a time (Ibid.).

This omission (amongst others) reveals the limitations of this form of archiving and indeed raises questions about the role of subjectivity in archiving (Walden, 2017:86). The blog archive is dependent on a single individual, and whilst the blog was produced in a systematic way with a clear editorial policy, what we saw over the years is that, with the pressures of work and private life – thilk gets married, becomes a father and changed jobs – his blogging activity ebbs and flows according to the time he has available (Ibid.). At one point, he simply runs out of enthusiasm and this, together with the fact that the blog’s host is hit by a security breach infecting his site with malware, results in him closing down *Movie Marketing Madness* for a period of three years until 2015, when he started posting again (Ibid.). Moreover, whilst his ‘archive’ is still publicly available, it has since been renamed *Cinematic Slant* and is preserved on his personal blog (for the time being, at least), but it suffers from the same link rot found in the award archives that render this archive a set of commentaries without a collection.

The last port of call in this search for the *D-9* websitewas the *Internet Archive* (Ibid.) but as I had found in initial searches, the ability to travel back in web history promised by the *Internet Archive*’s *Wayback Machine* did not prove as easy as its name might suggest (Ibid.). For its first decade, the *Wayback Machine* searched by web address (URLs) rather than search term (Rogers in Walden, 2017:87). So, while some sites shared their film’s title and were straightforward to locate, the site for *District 9* was titled *D-9.com* which made it more challenging to find (Walden, 2017:87).

According to one of the founders of Creative Agency, *Trigger Global*, Perry Wang, the *D-9.com* site was launched in July 2008 but it was not recorded by the *Internet Archive* until 27 August 2008 and the first archive snapshot of the site showed an empty black page featuring just the MNU logo and a still live link to a free downloadable Adobe Flash player plug-in (Wang in Walden, 2017:87). However, with the snapshot taken a year later on July 25 2009, I finally found what I was looking for (Ibid.). Following the trailer, hidden from view, was the *D.9* website (Ibid.). The archived site had been stripped down to its core (Ibid.). Several features of the D.9 movie experience were gone including ‘*The D-9 MNU Alert Game*’, ‘*MNU Training Simulation’*, as well as the ‘*MNU Spreads Lies’* site and the ‘*Maths from Outer Space*’ site (Ibid.). But the main site was archived, and, in part, operational (Ibid.). While the *Internet Archive* may have its limitations, it had archived the site and the *Wayback Machine* search engine provided access to it (Ibid). Moreover, through the calendar presentation of archived snapshots, the lifespan of this artefact becomes visible and can be measured. By tracking the site on the *Wayback Machine*, I was able to trace its lifespan through to 17th March 2013 and the site was taken down around the beginning of May 2013 which meant that it had a life spanning more than 5 years, vastly exceeding the conventional 6-week long campaigns of its analogue predecessors (Walden, 2017:87). From May 2013 the site was replaced with a link to the Sony Pictures site where all that remains now is a still from the film, a plot summary, and details of purchase options for home viewing formats.

.

Once its promotional role was fulfilled, this hybrid form of content and marketing which had contributed so much to the film experience, and had been widely recognised for doing so, was unceremoniously taken down from the web. While the archives examined in this chapter clearly valued this new kind of artefact, in practice they were unable to preserve it, and the only archive able to provide any kind of access to the website was the *Internet Archive*, and then just in a fragmented form.

### 3.8 Reflections and Conclusions

It has been observed that software-based culture is littered with archival metaphors offering up digital data repositories such as cloud storage, blog hosts, as well as social media platforms such as *Facebook* *Tumblr*, *Pintrest* and others (Parikka in Walden, 2017:88). The assumption with these digital repositories is that they have the facility to safeguard data and provide the panacea to the material challenges encountered in the analogue past, and that the premise underpinning them is their ‘always-thereness’ (Chun in eds. Huhtamo and Parikka, 2011:188). During the course of this transmedia archaeological search, a new generation of digital archives was explored and yet in spite of the archival promise of digital media, clearly for the most part, *D-9.com* has gone dark. It is no longer publicly available to view, or, at least, not in its entirety (Walden, 2017:89). So, this search bears out Chun’s observation that the promise of permanence is a misconception of the capacity of digital archives, as well as Ernst’s assertion that digital archives cannot store artefacts in the conventional understanding of the term.

What this transmedia archaeological search does demonstrate, however is some of the ways in which digital archives differ from their real-world counterparts, as well as some of the consequences for the digital artefacts in them (Ibid.). Materially speaking, archive content exists only as binary data with algorithmic instructions for how to build the site, in response to machine-event or user action (Peacock in Walden, 2017:88). The website’s assets are not embedded in the page. They exist separately and are called for algorithmically when the page is constructed in the browser, and this has implications for the way digital archives can be understood (Ibid.). If software-generated media objects are only manifest in the algorithmic process, they cannot be archived, only documented, and consulted and it is this that makes digital archives dynamic entities (Ernst in Walden, 2017:88). Over the course of the life of a film promotional site like *D-9.com* browsers are replaced and versions of operating systems and software supersede one another. Another possible variable is that cookies allow web objects to be presented to different people differently, and with the participatory web, the ‘social life’ of the artefact is defined by the uses to which it is put by different users at various times, as well as co-created artefacts such as the blog on the *MNU Tells Lies* site (Meyer, 2012:19). For digital archives, these issues raise questions about where digital artefacts begin and end, as well as how the meaning of an object changes over time (Ibid.).

From this encounter it is clear that digital archives are better understood as a network of information in which links are activated and reactivated as required, rather than as repositories of material objects in the conventional understanding of the archive (Ibid.). The consequence is that digital artefacts may be recorded but cannot be archived *per se* (Ernst in Walden, 2017:89). Whilst conventional physical archives continue to be concerned with evidence, information and systems of ordering and classification, in digital archives - like the ones encountered in this search- it is the metadata describing the object that is archived, rather than the material object itself (Walden, 2017:89). So, *de facto*, the first generation of film websites is lost and the concern is that the ephemerality of this new media cultural artefact means they will disappear before they have been appraised for what they bring to the contemporary film experience (Ibid.).

Whilst my interest is in film sites and online film marketing and promotion, the concern about digital cultural memory is not confined to this field, and this investigation also raises broader questions about how culture is remembered and indeed what culture remembers (Ibid.). Clearly, not everything can be archived and indeed, wherever there are archives, collections are always pre-conditioned by selection processes (Ibid.). It is also true that all archives have their agendas, which inevitably lead to blind spots and absences in the record of the past (Ibid.). But in the case of emerging cultural artefacts like film sites, we have an instance of concern. Without putting too fine a point on it, this experience is concomitant with the position articulated in an article in *Wired* magazine titled ‘We need to act to prevent a digital dark age’ which raised concerns about the preservation of data for future generations (Koehl in Walden, 2017:89).[[6]](#footnote-5) What is evident from the archives consulted in this search is that the wherewithal to archive digital culture remains out of reach. Moreover, the implications may not be confined to the world of film-related archives (Walden, 2017:89).

Casting further afield the point is illustrated by the disappearance of *GeoCities* (Walden, 2017:89). Originally established in 1995, *GeoCities* was a free web hosting service on which the first generation of digital pioneers set up homesteads in the form of do-it-yourself homepages made of hand-coded profile pages, gifs, and guest book pages (Gyford, 2009). *GeoCities* is regarded as a significant manifestation of early web culture but a few years after Yahoo! acquired it, the service was shut down in 2009 and almost lost to history (Lialina and Espenschied in Walden, 2017:90). In the event, a collective of volunteers known as *The Archive Team* dedicated to the preservation of digital heritage rescued a proportion of *GeoCities* homes pages which are now preserved in a vast downloadable Torrent file that formed the focus of an exhibition at The Photographer’s Gallery in London in 2013 titled *One Terabyte of Kilobyte Age* by Olia Lialina and Dragan Espenschied. What this cautionary tale illustrates is that the challenges encountered in the archive are not confined to questions of how to preserve the past but raise questions about the need to archive the present too. As I have already stated, the web is now three decades old, so it can no longer be regarded as a ‘new’ media (Walden, 2017:90). In this transmedia archaeological excavation of the recent digital past to search for *D-9.com*, what I encountered was not the artefacts themselves which, materially speaking, have disappeared from the public web, but an imprint of where the artefact had once been. As Ernst reminds us, in media archaeology we need to be mindful of the silences, gaps and the disappearances, as much as the presences (Ernst, 2013:194).

What this chapter has done is establish some of the locations on the web where film promotional sites have been archived. However, tracing the development of the earliest iterations of digital-born promotion since the early 1990s certainly invokes parallels with the challenges faced by media historians of the development of early film and live broadcast TV. As in periods of emergence, innovation and experimentation, there is less focus on record, preservation and prototypes get lost. Anecdotally, a senior creative marketing executive for one of the major American film studios told me in a recent interview that between 1995-2010, he had seen millions of dollars of work simply deleted in the spirit of good digital housekeeping but along with it went early examples of digital film promotion (Kelly, 2020).

So, what has this enquiry learnt about the first generations of film websites from this transmedia archaeological investigation of archives? Firstly, the temporality of a website is part of its meaning. Through the *Internet Archive* historical examples of film sites can be accessed first-hand and the archive charts the lifespan of campaigns, as well as the composition of film websites at various stages in its development. Next, the archives revealed the extraordinary diversity of film promotional sites including the emergence of stylistic conventions, genres and aesthetics which suggest this kind of film promotion is not just a mirror of the films it advertises, but can be considered as a form in its own right. The next point is that these archival examples illustrate the growing significance of promotion within the film industry and the changing relations of film to its promotion. But online promotion cannot be understood just in textural terms as the materiality of the medium is part of its meaning too. What was also evident from the search in the archives for *District 9*’s *D-9.com*, was a clear sense of the ongoing precarity of digital promotional materials.

Nowadays, of course, creative professionals and design agencies archive their work. Individual designers curate personal websites displaying their work with screen shots, walk-throughs, links to campaign coverage and commentary, highlighting their role in the creation of a campaign in ways that function like a curriculum vitae. While marketing agencies use recent work, awards, and accolades as online shop window displays to promote the agency. It is interesting to note that as film promotion budgets have risen, commentary has become increasingly elaborate, with agencies providing explanatory accounts of promotional campaigns that sometimes take the form of ‘making of’ reels in their own right.

Film promotion campaigns are displayed on platforms like *Behance* where other design professionals may view and show their appreciation. Take for example, *The Grand Budapest Hotel: Film Commission* on *Tumblr* site received a phenomenal 3.6k ‘likes’ and 42.2k views (as of 28th April 2022). Clearly *Behance* has become a significant locus of industry self-reflection (as well as self-congratulation). But as has been argued elsewhere, research into film promotion cannot just take producer’s accounts at face value (Gray, 2010:220). In order to get a better comprehension of the film promotion industry context, a deeper understanding is needed of the circumstances in which they are produced and this remains an outstanding challenge. But what research in these diverse archival settings has confirmed is the evolution of a new form of film promotion that plays an increasingly substantial role in film, beyond promotion. Film sites may provide exposition of the film’s fictional world, provide factual information, introduce different narrative modes – counter narratives, transmedia narratives, and narrative ‘experiences’- and take on aspects of storytelling process that were previously the preserve of the film, confirming their interest as a cultural artefact in their own right. So, having located where historical film websites have been collected, and how they have been archived, this chapter recounted encounters with the first generation of online archives and some of the insights these archives yielded about digital film promotion The next stage in the investigation is to consider more closely the artefacts at first hand.

1. There are 9 deposit libraries in the UK at present. Two British Library locations (one in London, St. Pancras and the other in Boston Spa, Yorkshire). There are further deposit libraries in the National Library of Wales (Aberystwyth and Cardiff), National Library of Scotland (Edinburgh and Glasgow), Bodleian Library (Oxford), Cambridge University Library, and one in Ireland at the Library of Trinity College, Dublin. [↑](#footnote-ref-1)
2. By placing robot txt files on their domain, a site can elect not to be archived and this signal is respected by the *Internet Archive*. [↑](#endnote-ref-1)
3. As of 4th August 2023 [↑](#footnote-ref-2)
4. A preliminary survey was undertaken of web design awards which indicated a considerable number had sprung up since the advent of the internet, illustrating the growing significance of the form. To qualify for the survey, awards had to meet three essential criteria: first they had to be conducted in English; secondly, the award had to include a film category; and lastly the award had to hold a publicly available record of winners. [↑](#footnote-ref-3)
5. Thilk currently writes a column in *The Hollywood Reporter online* on film marketing. [↑](#footnote-ref-4)
6. The term ‘digital dark age’ derives from the title of a presentation by Terry Kuny at the 63rd IFLA *(International Federation of Library Associations and Institutions) Council and General Conference* in 1997 and refers to the time between the widespread adoption of digital technologies and the production of digital-born artefacts, and the development of infrastructure and the wherewithal to preserve those artefacts (Stuart, 2012:554). [↑](#footnote-ref-5)