

404 – NOT FOUND: CRITIQUING THE DISAPPEARING WEB

Prof. Richard Vickers
Senior Lecturer New Media
Hull School of Art and Design,
University of Lincoln, UK
rvickers@lincoln.ac.uk

Introduction:

Within art & design education, critical analysis of existing and historically important work and artefacts has always played an important role. To study graphic design without studying the work of the people who shaped the art of graphic design, people such as Saul Bass & Paul Rand is unimaginable. To study typography without studying the inspirational work of Jan Tschichold would seem unthinkable, indeed the emerging Bauhaus movement in his native Germany had inspired the young Tschichold and propelled him on the trajectory that would help shape the art of typography in the 20th century. The artist Piet Mondrian wrote in the preface to Tschichold's 1928 book, *Die neue Typographie*: 'we are at a turning point of civilisation'.¹ We too are at a turning point of civilisation, the World Wide Web that was largely unknown to the majority of people only 10 years ago, now permeates everyday life in the modern world. How do we inspire and nurture the next generation of digital designers without viewing and critiquing existing or historical work and artefacts? Studying the work of influential graphic designers is usually done via the very medium that the designers themselves were working with, print. We can pick up a book and study Rand's work as it was intended, it may not be to scale, may lack some of its initial impact, lack political or contemporaneous context, however it can still be viewed in its intended form. New media products and artefacts have only

ever truly existed as digital work and interactive experiences that no book can ever truly reproduce.

404 not found, 410 gone: The digital dark age?

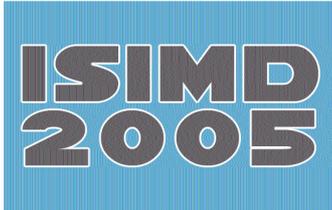
All regular users of the web have experienced the frustration of typing in a website address, only to have a 404 error displayed, the site has not been found, it is not available, that important URL that you saved to your favourites folder is simply no longer there. That innovative and unique site you wanted to share has simply ceased to exist; or at least isn't located where it was supposed to be, possibly gone forever. As educators in the field of new media design, it is essential that we constantly evaluate existing and chronological web based products and e-content with students. The critique is an essential part of new media design education, analysing what has come before, helps to shape what will come in the future. However when it comes to viewing web-based material, there is no certainty that it will still exist at the address it once did. The 404 in the title of this paper refers to the World Wide Web Consortium (W3C) Status Code Definitions for HTTP 1/1 (Hyper Text Transfer Protocol) that:

The server has not found anything matching the Request-URI. No indication is given of whether the condition is temporary or permanent. The 410 (Gone) status code SHOULD be used if the server knows, through some internally configurable mechanism, that an old resource is permanently unavailable and has no forwarding address. This status code is commonly used when the server does not wish to reveal exactly why the request has been refused, or when no other response is applicable.²

As the above paragraph states, the code 410 (Gone) should really be used if the site has been permanently removed or has no other known location. Of course there is no governing body to enforce that the correct error code is displayed when

¹ R. McLean, *Jan Tschichold: A life in typography*, Princeton Architectural Press, New York, 1997.

² R. Fielding et al, <http://www.w3.org/Protocols/rfc2616/rfc2616.html>



3RD INTERNATIONAL SYMPOSIUM OF INTERACTIVE MEDIA DESIGN

JANUARY 5 - 7, 2005

a site is permanently removed and the system relies upon self-governance and that owners or authors assume responsibility for managing the web based content that they consider been beyond its time. The World Wide Web consortium states:

The 410 response is primarily intended to assist the task of web maintenance by notifying the recipient that the resource is intentionally unavailable and that the server owners desire that remote links to that resource be removed. It is not necessary to mark all permanently unavailable resources as "gone" or to keep the mark for any length of time -- that is left to the discretion of the server owner.³

The Domain Name System (DNS) is used to translate the numerical computer address to a more user-friendly name, which is easier to remember and helps the user locate information more succinctly on the web. Domain names do indeed make navigating the web much easier, but they have a lifespan of only two years from the point of registration and require renewal every two years thereafter. So many domain names are not renewed and expire, ownership will eventually transfer to someone else with the new owners replacing the data with their own, contributing to an ever-growing problem of displaced or lost work and artefacts. Jakob Nielsen identifies broken links as a major problem, calling the phenomenon 'Linkrot'.⁴ He states:

Any URL that has ever been exposed to the Internet should live forever: never let any URL die since doing so means that other sites that link to you will experience linkrot.⁵

As early as 1997 'Linkrot' was identified as one of the biggest problems on the web by over 60% of users in a survey of that year. It has become such a regular occurrence, typing in a URL only to have a 404-error page displayed, that we just accept it, or have no option but to do so, as to whom can we complain anyway? The creative products and very fabric of web based new media design are disappearing before our eyes, or at least being superseded and displaced and/or replaced. There remains little or no evidence of those products, although hopefully they will still exist on somebody's hard drive somewhere, it is essential that these artefacts be saved for posterity before they are lost forever. Computer scientist Danny Mills suggests that historians will look back on this period and consider it to be 'the digital dark age'.⁶ Some of the worlds leading and most inspirational new media designers have taken the initiative to preserve and present their past work. Yugo Nakamura has archived his work dating back to 1998, facilitating access via his current site⁷ with absolute links to the individual pieces stored on his server space. Similarly new media artist and technologist Joshua Davis, has archived his personal and commercial work dating back to 1998, on his current site⁸. But these are exceptions rather than the norm and much work has disappeared and may even be lost forever.

Archiving the web.

Lyman & Varian's⁹ research of 2000 concluded that the web contained more than 4 billion public pages and an additional 550 billion connected documents on call in "the deep" web¹⁰. The web keeps growing at an exponential rate and some 7 million pages are added daily, whilst the average lifespan of a web page has shortened to only 44 days¹¹. The loss of

³ Ibid

⁴ Jakob Nielsen, Designing Web Usability, New Riders, USA, 2000. P.249.

⁵ Ibid

⁶ <http://www.longnow.org/10klibrary/library.htm>

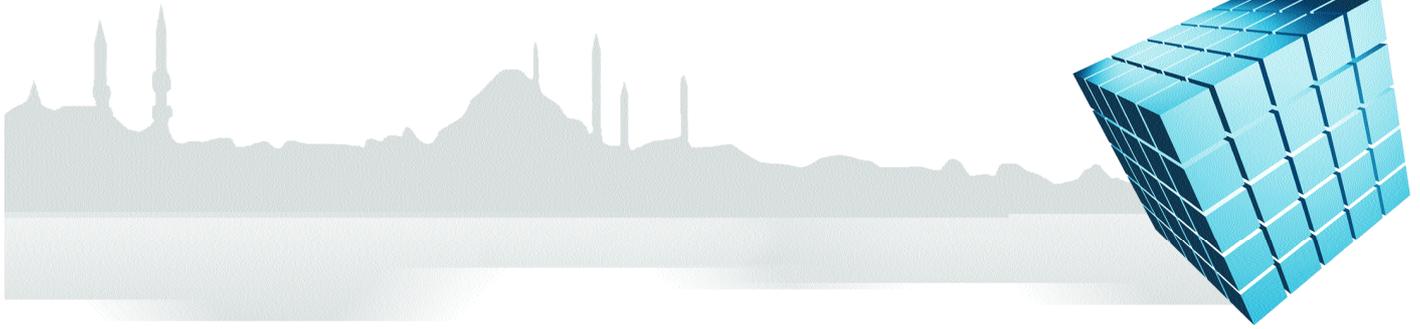
⁷ <http://www.yugop.com/>

⁸ <http://www.joshuadavis.com>

⁹ <http://www.sims.berkeley.edu/research/projects/how-much-info/internet/rawdata.xls>

¹⁰ <http://loop1.aiga.org/content.cfm?ContentID=99>

¹¹ Ibid



data on the web has been compared to the burning of the world's greatest archive of knowledge, the ancient library of Alexandria in Egypt and there are already initiatives to preserve collect or archive web-based material and artefacts. One such undertaking is the Internet Archive project that was established in 1996 to 'prevent the internet – a new medium with major historical significance – and other "born-digital" materials from disappearing into the past.'¹² An important undertaking that follows the more traditional approach of a library, indexing material for access by researchers, historians and scholars. Brewster Kahle founder of the Internet Archive identified that research found half of the web disappearing every year, stating that the "the Net has no memory"¹³. The archive is populated by Alexa Internet¹⁴, who have been crawling and creating snapshots of the web since 1996, donating a copy of each crawl to the Internet Archive. To fully utilise the potential of the Internet Archive the user needs to have basic UNIX programming skills, although it is hoped tools and methods will be developed to give the public easy and meaningful access. In 2001 the Internet Archive launched its 'Wayback Machine'¹⁵ a free service that allows the public to search and view the Internet Archive's enormous collection of web sites, dating back to 1996 and comprising over 10 billion web pages. The Internet Archive Wayback Machine now contains approximately 1 petabyte of data (equal to 1,024 terabytes or 1,048,576 gigabytes) and continues to grow at a rate of 20 terabytes per month.

One problem that I have experienced with using the Internet Archive is in its ability to store content rich websites, especially sites that contain Macromedia Shockwave & Flash movies. Some sites in the archive can have most or even all images missing if the images have not been successfully archived and are not available on the Internet Archive servers.

Also archived sites that use JavaScript rollovers can have the rollover state image missing, which detracts from the intended experience for the user. There are also problems with archiving dynamic pages, some of which are easily stored in an archive and some of which fall apart completely. When a dynamic page renders standard html, the archive works beautifully. When a dynamic page contains forms, JavaScript, or other elements that require interaction with the originating host, the archive will not contain the original site's functionality¹⁶, considering that the web is becoming ever more dynamic this could prove to be a major problem. It is these dynamic and content rich sites that usually represent the most innovative and creative work, the cultural artefacts for critical analysis and reflection as part of new media education. To lose these artefacts forever has serious implications for educators in the field of new media education and the design community at large.

To preserve, curate and display the creative web.

What initiatives could be established to preserve, curate and display the creative web for posterity? There is a real and pressing need to establish an international chronological collection of the most innovative, creative web based new media products, work and artefacts. How that collection is labelled is another debate, should it be a library, archive or museum, or something else? Whatever it is called it will need considerable organisation; an expert panel would need to be established to collate and curate the ever-growing collection as well as evaluating the historical importance of past work. For the collection to be truly representational work would need to be submitted internationally and the initial selection process for submission should be devolved to the participating countries. There already exist many suitable avenues from which to populate such a collection as most countries have national awards and competitions covering a broad spectrum of

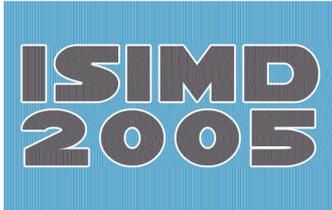
¹² <http://www.archive.org/about/about.php>

¹³ http://www.voxel.com.br/pt/knowledge/the_digital_information/pag07.aspx

¹⁴ <http://pages.alexa.com/company/technology.html>

¹⁵ http://www.archive.org/about/press_release.php

¹⁶ <http://www.archive.org/about/faqs.php>



3RD INTERNATIONAL SYMPOSIUM OF INTERACTIVE MEDIA DESIGN

JANUARY 5 - 7, 2005

commercial, personal and student work. In the United Kingdom there are two major national multimedia competitions, the British Interactive Media Association¹⁷ (BIMA) annual awards and the British Academy of Film & Television Arts¹⁸ (BAFTA) Interactive Awards. There are other national awards in the UK, for example the British Design & Art Direction¹⁹ (D&AD) awards which covers a broad spectrum of design work including interactive media and the Mando Group²⁰ Student Website Design awards. The BIMA and D&AD awards also have a student category and this year BAFTA have introduced a New Talent Award²¹, which intends to 'recognise notable creative innovation, technical excellence and production achievements across digital media for people with less than 3 years industry experience'. 2004 was the twentieth year of the BIMA awards, however only the 2003 and 2004 shortlisted nominees and winners are listed on the current BIMA website. The BAFTA interactive awards were established in 1997 but the BAFTA site does not list any previous nominees or winners. The D&AD website does list nominees and winners retrospectively from 1998 to the present day and also provides some basic information on the work with screengrabs possibly as a response to the time sensitivity of online work.

The leading pan European competition EUROPRIX²² was started in 1998 with a clear mission: to honour excellence in multimedia. The EUROPRIX Student Award was renamed the Top Talent Award in 2002 and the competition is dedicated to promoting young talent: those emerging producers who could become leading innovators in the future. Naturally the work and products selected and short-listed for these competitions and awards are of a very high standard,

exemplary in terms of creativity, innovation and execution. It must be remembered however that the submissions for these competitions are not wholly exhaustive in terms of representation of all existing material, and most require an entry fee or stipulate candidate eligibility criteria that may prohibit entry. There are many other excellent sources of 'best of' examples, such as Macromedias showcase site of the day²³ and the increasingly popular Favourite Website Awards²⁴. The Macromedia site of the day archive however only dates back to 2001 and for that year and 2002 the archive has been reduced to a collection of the ten 'best of' examples, which is mostly commercial work or work that serves to promote Macromedia products. The Favourite Website Awards (FWA) was established in 2000 to recognise the latest cutting edge website design and it is an industry recognised and respected award program. Awarded work is displayed with a screengrab and short annotation and as the FWA site relies on links to the actual work rather than an archive, the viewer is informed as to whether the site is still online. If it isn't the link is removed or it is stated that the current site at that address is no longer the site that won the award. Unfortunately this is not infallible and the user can still experience broken links and 404 errors.

A recent initiative is the World Summit Awards (WSA) established by the World Summit on the Information Society²⁵ (WSIS) whose mission statement is:

A truly Global Information Society is one where all persons, without distinction, are empowered freely to create, receive share and utilize information and knowledge for their economic, social, cultural and political development.²⁶

¹⁷ http://www.bima.co.uk/content_awards/about_awards.html

¹⁸ <http://www.bafta.org>

¹⁹ <http://www.dandad.org/>

²⁰ <http://www.mandogroup.com>

²¹ <http://www.bafta.org/interactive/index.htm>

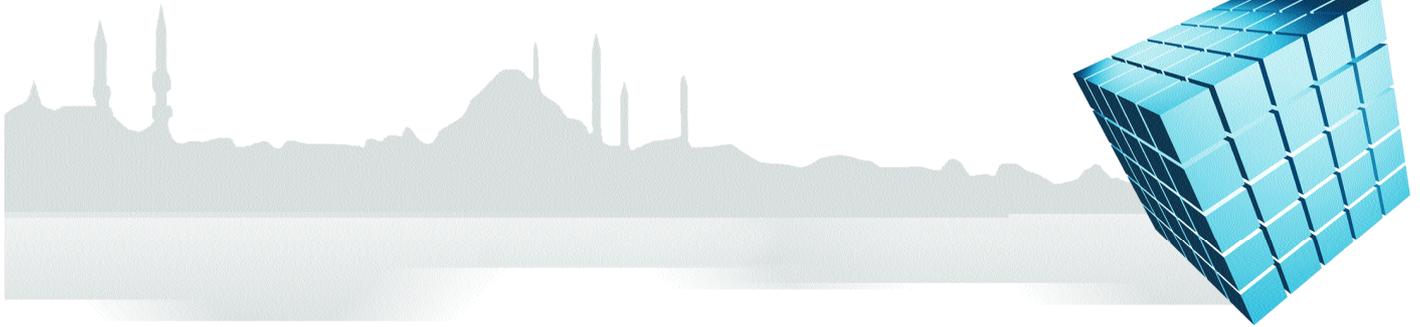
²² <http://www.toptalent.europrix.org/index.htm>

²³ <http://www.macromedia.com/cfusion/showcase/index.cfm>

²⁴ <http://www.favouritewebsiteawards.com/>

²⁵ <http://www.wsis-award.org/>

²⁶ Ibid



The WSIS began the World Summit Award in 2003 with the goal of making existing excellence visible, the website states that:

The World Summit Award, a global project, held in the framework of the WSIS, seeks to demonstrate the benefits of the Information Society in terms of the new qualities in content and applications, by selecting, presenting and promoting the best products from all over the world with a special emphasis on bridging the digital divide.²⁷

As Peter Bruck in his introduction to the 2003 WSA Worlds Best e-Contents publication states, 'Technology does not inspire. Contents do'.²⁸ The World Summit Award is a very positive and progressive initiative, sharing and developing the Information Society on a truly global scale. This is the spirit of the very web that Tim Berners-Lee had intended, a sharing network. For the 2003 WSA Awards 40 outstanding products were selected from over 800 global submissions by the WSA Grand Jury to exemplify the quality of today's multimedia market worldwide. The final 40 are presented via the WSA site with a short description and screengrab and when applicable a link to the actual work. Once again however these links will be susceptible to the time sensitivity of online work.

The above examples all follow a similar formula; a selection process is applied to select the best and most suitable work to be considered by an expert panel. The expert panel is usually made up of members selected from the design, media, communications and new media industries and education establishments, with a proven background and experience. The selection criteria for the work may include originality, innovation, creativity, and suitability of product for intended target audience and subject matter. Other criteria may include the quality and comprehensiveness of content, interface design,

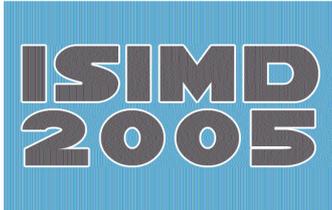
ease of use, functionality, navigation and orientation, the value added through use of interactivity, technical realisation and aesthetic value. These processes and procedures generate expert reviews, comments and annotations for each piece of work, forming the basis of critical analysis for the presentation of the work to a wider audience. Consideration must be given to how we preserve and display these cultural artefacts for future audiences, as we cannot continue to rely on links to view this work for any length of time. The many national and international awards are an ideal starting point for the creation and population of a creative web collection. For that collection to have real value it must be accessible from one virtual location ideally for perpetuity.

Conclusion:

The further and faster we hurtle into the future, the greater the risk that we may lose these artefacts forever. Digital data has become one of the most disposable commodities ever; we consume storage media at ever increasing rates. I personally have ten years worth of back-up and archive disks, but I still don't have copies of all the digital work that I have produced during that time, never mind all the innovative and stimulating creative work produced by others that I have seen, used and experienced during that period. We erase, delete or move data at will, making room for that next killer application, or 'groundbreaking' piece of work. In the rush to embrace the future, the recent past is only a click away from oblivion. We may have archived material on storage mediums that have been superseded and become obsolete. Many storage formats have come and gone in the last 20 years, for example the 5_-inch disk and now its successor the 3_-inch floppy disk is on the verge of relegation to the history books. The CD-Rom is almost certainly to be relegated to data storage history in the near future, as it is rapidly being superseded by the increased storage capabilities of the DVD. Indeed there exists doubt as

²⁷ Ibid

²⁸ Peter Bruck, Bridging the content gap - showcasing the World's Best, The World's Best e-Content publication 2003, WSA Office, ICNM, Austria.



3RD INTERNATIONAL SYMPOSIUM OF INTERACTIVE MEDIA DESIGN

JANUARY 5 - 7, 2005

to the longevity of the very media that we rely on to archive, with some experts suggesting that CD-R and DVD-R media may only have a life of between 5 – 15 years²⁹ before they degrade. Whilst the World Wide Web is unlikely to be superseded in the near future, its exponential growth is certain to continue and it is that very growth that causes the displacement and extinction of the some of the most creative online work and artefacts produced in its short history. The design community and educators must accept responsibility to archive, curate and preserve the creative web for reference and critical analysis for current and future generations. We must ensure that the creative web doesn't disappear forever and that rather than a digital dark age, this period may be remembered by historians as the age of digital enlightenment.

²⁹ Stewart Brand, <http://www.longnow.org/10klibrary/library.htm>