

The Deleted City: A Digital Archaeology

RICHARD VIJGEN

KEYWORDS *Data visualisation, digital archaeology, interactive, internet as a city, metaphors for the internet, net-art, net-culture*

DATE 2012

URL <http://www.deletedcity.net/>

ABSTRACT The Deleted City archives Geocities, a proprietary digital archaeology of the world wide web as it exploded into the 21st century. At that time the web was often described as an enormous digital library that you could visit, or contribute to, by building a home-page. The early citizens of the net (called “netizens”) took their netizenship seriously and built home-pages about themselves referenced to subjects they were experts in.

These pioneers found their brave new world through a free web-hosting provider that was modelled after a city and where you could get a free “piece of land” to build your digital home: Geocities. In such a digital world one’s neighbourhood was based on the subject of the homepage. “Heartland,” one such neighborhood, was—a neighbour-

hood for all things rural—by far the largest, but there were neighbourhoods for fashion, arts and far east related topics to name just a few.

Around the turn of the century, Geocities had tens of millions of “homesteaders,” as the digital tenants were called, and was bought by Yahoo! for three and a half billion dollars. Ten years later (2009) as other metaphors of the internet (such as the social network) had taken over, many of these homesteaders had left their properties vacant after migrating to Facebook, Geocities was shutdown and deleted. In an heroic effort to preserve 10 years of collaborative work by 35 million people, the Archive Team made a backup of the site just before it shut down. The resulting 650 Gigabyte bit-torrent file is the digital Pompeii that is the subject of an interactive excavation that allows you to wander through an episode of recent online history.

INTRODUCTION

In the fall of 2009, a message appeared on the homepage of Geocities.com. It stated the following: “On October 26, 2009, your GeoCities site will no longer appear on the Web, and you will no longer be able to access your GeoCities account and files”. With this message Yahoo! informed the world that it was about to shut down and delete the popular online community it had acquired ten years earlier for \$3.57 billion.

Geocities, founded by David Bohnett and John Rezner in 1994, was an online community modelled after a city, which would become one of the first large online communities on the internet. For many who were new to the web

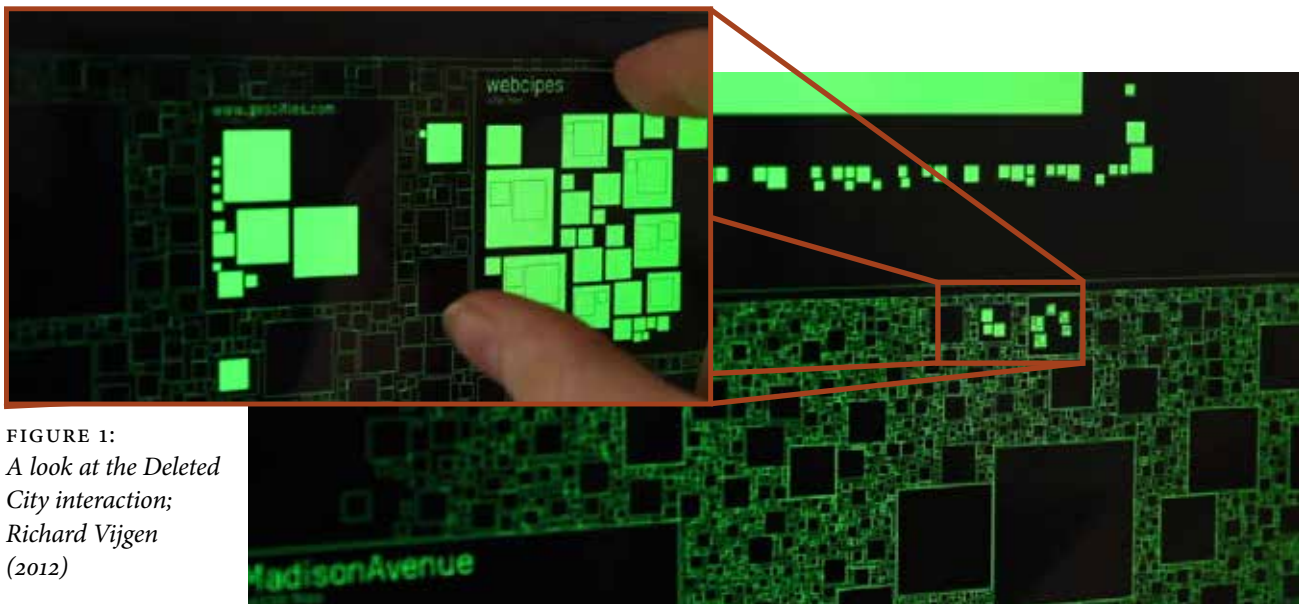


FIGURE 1:
*A look at the Deleted
City interaction;
Richard Vijgen
(2012)*

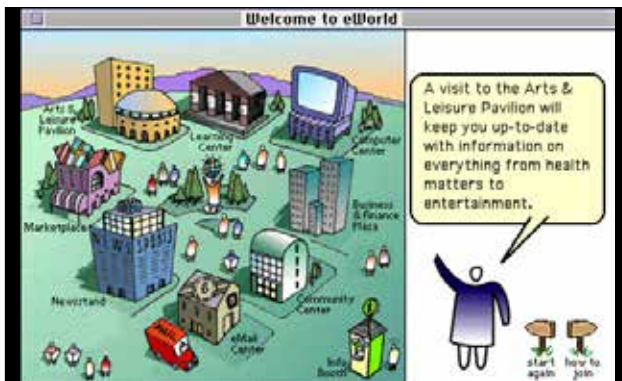


FIGURE 2: *The internet as a city.*
eWorld, an online service by Apple, 1994.



FIGURE 3: *Geocities' representation of a block of home pages.*
Wayback Machine, July 4, 1998.



FIGURE 4: *User made emblem by which residents identified themselves with their neighborhood.*

(including myself), Geocities provided a very attractive model to get involved with this new medium. In the early and mid-1990s, it was not uncommon to describe the internet using metaphors like “digital city” or “global library” (FIGURE 2). Geocities was in a way a little of both: it followed the analogy of a city by arranging users’ home pages in thematic neighborhoods, while at the same time providing (in a playful manner) a framework for contributing to an ever-expanding body of knowledge.

New users, referred to by Geocities as homesteaders (or more broadly netizens, a portmanteau word composed of “citizen” and “internet”), would select a neighborhood based on the subject of their homepage. These neighborhoods had descriptive names often corresponding with places that exist in the real world; Hollywood for home pages about movies, and Athens for home pages about historical topics. Once having selected a neighborhood, you could then specify your subject further by choosing a “block” or sub-neighborhood. Athens/Acropolis, for example, could be a place for home pages about ancient Greek mythology. Within this specific area you would get a free piece of land (one megabyte initially) to build a homepage.

Although Geocities never enforced the correspondence between neighborhood and content, the model of a city provided a mental and spatial arrangement of information. Visually, however, Geocities was represented as a folder structure (FIGURE 3). By linking your online identity to a homepage located in a specific neighborhood (FIGURE 4), the Geocities model invited users to add information to the internet, rather than just extract information from it. Once you had obtained the necessary land, you could then start with the construction of your homepage. Interestingly, the concept of a homepage was not very strictly defined. Geocities provided a page wizard that allowed you to get started quickly using configurable templates, but many decided to build their home pages by hand using html. The lack of a fixed definition of a home page, combined with the flexibility of a large number of people experimenting (oftentimes struggling) with the medium, led to what you might call in retrospect a “home page culture” (FIGURE 5). This culture developed in tandem with the medium itself, incorporating new technical possibilities (audio, moving images) as they became available. Thus, new conventions and idioms, enabled by these emerging technologies were explored.

A well-known example is the “under construction” message. Originally an apology for a homepage that was not quite finished (perhaps because the author was

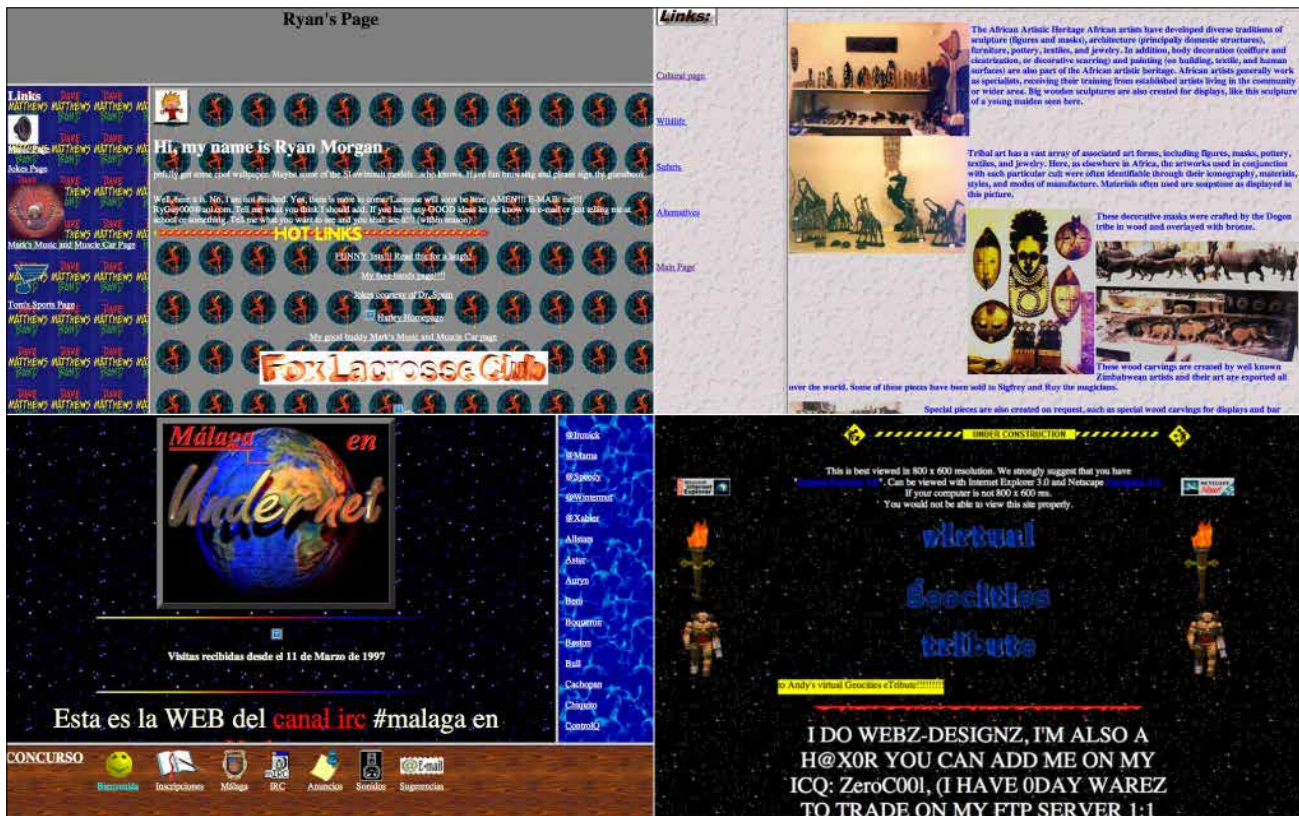


FIGURE 5: Selection of typical Geocities home pages: repeating backgrounds, animated gifs, html tables and frames.

trying to master HTML code), it became a widely used phrase, found on most home pages. Eventually, it became a concept synonymous for the DIY mentality of home page culture, and was celebrated to an extent that thousands of different visual variations were made (FIGURE 6). Another expression that has become a part of popular culture as a result of the experimentation by Geocities users, is the



FIGURE 6: Collection of “under construction” signs. Found in the Archive Team backup.

animated gif. At a time when streaming online video was still mostly a promise, the animated gif allowed authors to create very short animations consisting of just several frames. Animated gifs became a hallmark of the 1990s home page, leading to a retro revival in the 2010s. Recently, the animated gif has also become an accepted artistic medium that has found its way into galleries and museums.

The sound of the home page was defined largely by the use of MIDI files. Originally developed as a digital interface between musical instruments and the computer, the note-based file format was used by home page builders to synthesize instrumental versions of popular songs. The use of MIDI files as background music for a html document, reached its high point in the home pages of Geocities, but has rarely been used since.

THE RISE AND FALL

By 1999, at the height of the dot-com bubble, Geocities had become the third most visited website on the internet and had gone public a year earlier. Shortly thereafter, it was bought by Yahoo! for \$3.5 billion in stock. In an

attempt to monetize on its investment, Yahoo! tried—in what many of the homesteaders saw as a virtual coup-d'état—to take ownership of all copyrights that, until then, had always belonged to the authors themselves. This clash of cultures—helped by the emergence of new metaphors of the internet, such as the social network—gradually led to a mass migration of users away from Geocities to newly founded communities like Myspace, Friendster, and later, Facebook. This left Yahoo! with what could be described as a digital ghost town. On October 26, 2009 they decided to pull the plug. With 30 days notice, all 38 million home pages would be deleted.

THE 30-DAY ARCHIVE

This 30-day time frame was used by archivists around the world to save what could be saved. In a coordinated attempt to download as much of Geocities as time permitted, several groups were able to copy a large (but unknown percentage) of Geocities as it then existed in October of 2009.

One of these groups is the Archive Team. After a month of collaborative downloading, they compiled a 650 gigabyte (compressed) archive, containing most of the tens of millions of home pages that made up Geocities. The resulting file was made available on BitTorrent in order to distribute it to as many computers as possible and save it for posterity.

THE DELETED CITY

The 650 gigabyte BitTorrent file was the starting point for the project *The Deleted City, A Digital Archaeology* (FIGURE 7–10). The Deleted City is an interactive visualization which I call a “digital Pompeii”; a virtual city frozen in time and preserved in the Archive Team backup. The installation visualizes the millions of home pages as a virtual city map. Monochrome lines, reminiscent of old phosphorescent computer screens, represent the thematic neighborhoods and sub-neighborhoods within them. After Yahoo!’s acquisition of Geocities in 1999, users were also allowed to build home pages outside the neighborhoods, allowing for so-called “vanity URLs” like `geocities.com/~myHomePage` (Geocities was case-sensitive). By allowing this, Yahoo! largely departed from the internet-as-a-city concept that was epitomized by Geocities. Because in the installation, the layout of the map is generated by a computer program that mimics the growth of an actual city, these small independent home pages are densely clustered near the center of the map (FIGURE 7) and slowly give way to large “suburbs” as you move away from the center towards the edges of the city.



FIGURE 7: Close up of the Athens, surrounded by “vanity URLs.” Richard Vijgen, *The Deleted City* (2012).



FIGURE 8: Close up of the “Catwalk” block located within the “Fashion Avenue” neighborhood; Richard Vijgen, *The Deleted City* (2012).

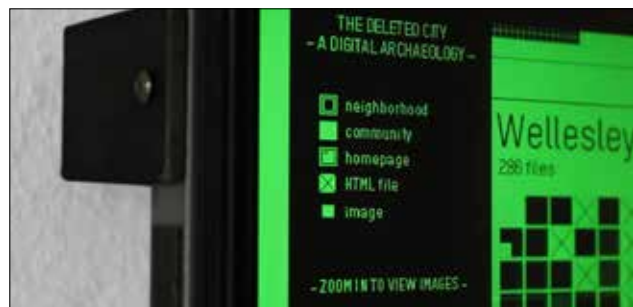


FIGURE 9: Close up of the screen showing the legend; Richard Vijgen, *The Deleted City* (2012).



FIGURE 10: Close up of the screen showing individual images contained within a home page; Richard Vijgen, *The Deleted City* (2012).

By using drag and pinch gestures on a multi-touch screen, the user can navigate around the city in a manner similar to Google Maps. Starting with a zoomed-out view, showing almost the entire city, you can zoom in on a neighborhood, thereby slowly revealing the sub-neighborhoods, blocks and individual home pages (FIGURE 8). Zooming in even further reveals the individual files (FIGURE 10), images and texts that a homepage consists of. MIDI files become audible when they are in the vicinity, and animated gif files start playing once they are near. Zoom out a little to get an overview of the area, and zoom in again on a neighboring home page about the same subject. Exploring a neighborhood like this shows the strength of the concept of the internet-as-a-city, where individual users become neighbors based on the subject of their work. But it also shows the weakness and lack of enforcement of the connectivity between content and neighborhood; since amidst a group of pages about archaeology, you might find a page with a family's vacation photos.

After exploring The Deleted City for a while, you might feel as if you have opened a digital time capsule. The gesture-based interaction with the touch screen is inherently 21st century, but the contents of the installation look and feel as if they come from a different era.

PERSPECTIVE

You might say that the Geocities backup is in fact a relic from a different era. Even if from a historical perspective (let alone an archaeological one), it is just from the recent

past. However, digital time moves fast, and even though it is often said that the internet does not forget, once a particular technology has been replaced by a newer version, it seldom survives. Geocities was one of the first successful user-contributed communities on the internet, and as it turns out, one of the first to be deleted. In a time when the internet has moved beyond the experimental phase and has become accessible to a truly global audience, few of its users still have to—or want to—build their own home pages. Even the concept of the home page itself has been replaced by that of the profile; the dominant mode of representation in today's social internet. Still many of the concepts and stylistic elements developed during the era of the internet-as-a-city, have found their way into the templates and frameworks that define the internet we use today.

You might argue that the spirit of the global library, where all contribute to a body of knowledge, lives on in Wikipedia, and the personal tone of voice found in many home pages can still be seen in Tumblr blogs, where authors can express themselves beyond the templates and frameworks used by Facebook and others.

Today, however, the somewhat utopian concepts of the internet as a virtual city or a global library have largely been replaced by more utilitarian notions of the internet as a social network or cloud. Where Geocities invited you to contribute, modern day networks invite you to share. Where hand-coded home pages led to a gradual evolution of conventions, change in modern networks such as Facebook is incremental and centralized. The netizens

```
String[] neighbourhooddata;
JSONArray objects;

neighbourhood (float xPositie, float yPositie, float objectSize, int loopval, String naam, int i) {
    xPos = xPositie;
    yPos = yPositie;
    boxSize = objectSize;
    boxHeight = boxSize;
    boxWidth = boxSize;
    loopnummer = loopval;
    buurtNaam = naam;
    bestanden = numFiles;

    if (loadStrings("neighbour1/"+naam+".csv")!=null) {
        neighbourhoodData= loadStrings("neighbour1/"+naam+".csv");
    }

    if (loopnummer>0) {
        for (int i=0; i<loopnummer-1; i++) {

            if (rectRectIntersect(xPos, yPos, xPos+boxWidth, yPos+boxHeight, neighbourhoodArray[i].xPos, n

                //println("overlap");

            if (looplap > 0) {
                looplap = 2;
                //println("geen overlap meer");
                break;
            }
        }
    }
}
```

FIGURE 11: Sample of the file manifest; Richard Vijgen (2011)

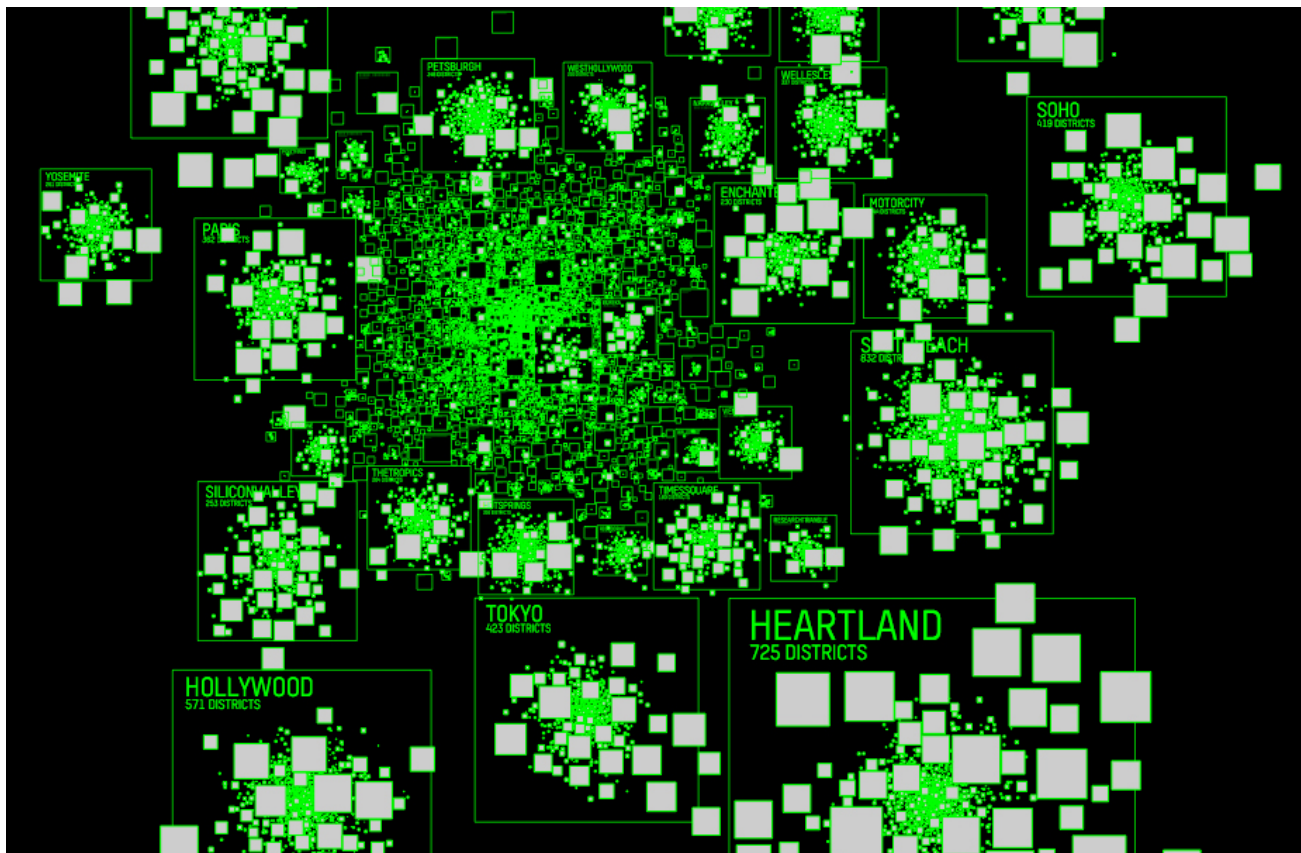


FIGURE 12: *Sample of the file manifest; Richard Vijgen (2011)*

have become users.

The Deleted City aims to revisit a previous incarnation of the internet and allows you to see how the technology, the esthetics, the metaphors, and the values that underpin it have changed. In some manner for better, in others, for the worse.

TECHNICAL BRIEF

The Deleted City is an interactive visualization of the 650 gigabyte compressed backup of Geocities.com made by the Archive Team. Because Yahoo! had declined to cooperate and share information about the information structure of Geocities with the Archive Team, they used the GNU get command to find connections within the site in order to duplicate its structure. As a result, only files that were linked to could be “seen” and downloaded. This means that, although the backup most likely contains a significant percentage (the actual size is unknown), it is probably not complete.

Downloading the file took several months. For BitTorrent to work, at least some users have to “seed” a file

in its entirety and be willing to share. For a 650 gigabyte file this proved to be difficult. After most downloaders reported being stuck at 44% for several weeks, a lively discussions on the Archive Teams IRC channel and Archive Team’s foreman Jason Scott’s website resulted in people donating bandwidth and sending physical hard disks containing the files across continents to restart the BitTorrent transfer.

After downloading the file, it turned out that Geocities file system was case-sensitive. Two files that have the same name, but different capitalization, can exist separately in this file system. Downloading these files to a standard FAT32 or Mac OS Journaled file system will overwrite one file with the other, resulting in an incomplete and corrupt download. Reformatting the disk to Mac OS Journaled Case Sensitive and re-downloading the missing files solved this problem.

Extracting the compressed files in a terminal in “verbose mode” again took several days, but allowed me to use the output as a manifest describing all files and folders in the archive (FIGURE 11). The extraction manifest,

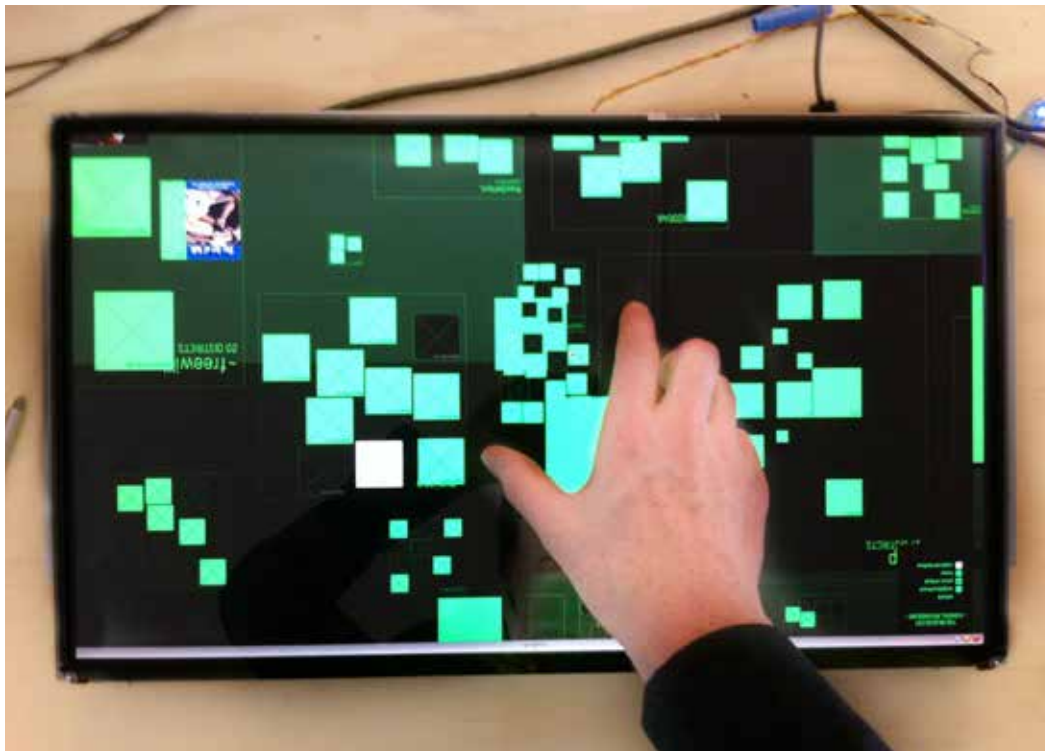


FIGURE 13: *Sample of the file manifest;* Richard Vijgen (2011)

containing about 36 million files, was uploaded into a database for further analysis. Apart from the uppercase/lowercase issue, there was another problem surrounding sym-links, or symbolic links. These shortcut URLs, which point to another location in the archive, could sometimes lead to as much as three different URLs pointing to the same file. Using the file database I created earlier, these URLs were grouped together in order to create a cleaned-up, uniform description of the archive.

The next step was to generate a visual representation of the database containing the description of all files. Using the scripting language PHP, I designed a set of rules that result in a map where each neighborhood is given a size depending on the number of home pages it contains. Neighborhoods are then placed in a two-dimensional grid by an algorithm that looks for the most efficient placement, while preventing the neighborhoods from overlapping each other. This results in small neighborhoods being placed in the unused areas between larger neighborhoods, mimicking a city that tries to use the land efficiently. Several algorithms were explored to find a balance between efficient use of space and visual expression (FIGURE 12).

Once all neighborhoods were positioned on the map, the process repeated itself for sub-neighborhoods, blocks,

home pages, and files; until all the files in the database were given a coordinate and a size. The new database that resulted from this process, described the layout of the map. It was then exported to an optimized SQLite file of about 5 gigabyte that could be visualized.

The visual aspect of the installation is made using Processing. A touch screen (FIGURE 13) sends user interaction to the software using the TUIO protocol, after which it queries the SQLite map database to return only the part of the map that falls within the boundaries of the screen. Because the visual representation is real time and vector-based (rather than tile-based), a zoomed-out view of the map might result in millions of neighborhoods being drawn on the screen, which compromises the fluidity of the application. To prevent this, the map reveals more detail as you zoom in, while keeping the level of detail as high as the performance of the computer permits.

BIOGRAPHY

Richard Vijgen (1982) is a information designer/artist living in the Netherlands. His Studio for Object Oriented Information Design & Research initiates projects that visually explore contemporary data culture.