Project "The Digital City Revives"
A Case Study of Web Archaeology

Tjarda de Haan
Amsterdam Museum
Kalverstraat 92
1012 PH Amsterdam
+31 6 47132758
T.deHaan@amsterdammuseum.nl

ABSTRACT
Twenty-two years ago a city emerged from computers, modems and telephone cables. On 15 January 1994 De Digitale Stad (DDS; The Digital City) opened its virtual gates in Amsterdam. DDS, the first virtual city in the world, and made the internet (free) accessible for the first time to the general public in the Netherlands. But like many other cities in the world history, this city disappeared. In 2001 The Digital City, the website, was taken offline and perished as a virtual Atlantis. Although the digital (r)evolution has reshaped our lives dramatically in the last decades, our digital heritage, and especially the digital memory of the early web, is at risk of being lost. Or worse already gone. Time for the Amsterdam Museum and partners to act and start to safeguard our digital heritage. But, how to excavate The Digital City, a virtual Atlantis, and reconstruct it into a virtual Pompeii? In the case study of web archaeology we will try to answer the questions: how to excavate, reconstruct, present, preserve and sustainably store born-digital heritage and make it accessible to the future generations? [1]

Keywords
The Digital City, web archaeology, digital heritage, digital preservation, collaboration

1. INTRODUCTION
De Digitale Stad (DDS; The Digital City) is the oldest Dutch virtual community and played an important role in the internet history of Amsterdam and the Netherlands. For the first time internet was (free) accessible to general public in the Netherlands. DDS is an important historical source for the early years of the internet culture in the Netherlands. The virtual city and its inhabitants produced objects, ideas and traditions in new digital forms such as web pages, newsgroups, chat, audio and video. DDS was a testing ground, and operated at the cutting edge of creativity, information and communication technology and science. It was not only an experiment with computers, but an experiment with questions, problems and challenges posed by the emerging information and communication technology.

And things were moving fast on the electronic frontier. DDS followed the developments closely which resulted in several interfaces (cityscapes):

1. **DDS 1.0**: 15 January 1994; all information and communication was offered in the form of a text-based environment (command-line interface; MS-DOS, UNIX) in Bulletin Board System technology. The so called 'Free-Nets' in the United States and Canada where a major source of inspiration for the founders. Free-Nets were 'community networks', or 'virtual communities' developed and implemented by representatives from civil society ('grassroots movement'). The metaphor of the city was reflected in the organization of the interface. There was a post office (for email), public forums to meet other visitors, a town hall and a central station (the gateway to the internet).
2. **DDS 2.0**: 15 October 1994; entry to the World Wide Web with the first DDS website with a graphical interface and hyperlinks.
3. **DDS 3.0**: 10 June 1995; introduction of the interactive 'squares' interface, the basic framework of the city's structure. Each square had its own theme and character, and served as a meeting place for people interested in that particular theme. Visitors could find information and exchange ideas with each other. 'Inhabitants' could build their own 'house' (a web page), send and receive emails (worldwide!), participate in discussion groups, chat in cafes, take part in the 'Metro', vote etc.

The Dutch social network DDS proved to be very successful. During the first weeks in 1994 all modems were sold out in Amsterdam. In ten weeks' time 12.000 residents subscribed. There was 'congestion' at the digital gates. Over the years the DDS user base of 'inhabitants' was growing: in 1994 there were 12.000 users, in 1995: 33.000, 1997: 60.000, 1998: 80.000 and in 2000: 140.000. DDS attracted international interest for the design it had chosen: DDS used the metaphor of a city to structure the still relatively unknown internet and made the users into 'inhabitants' of the city.

But in 2001 The Digital City was taken offline and perished as a virtual Atlantis. Ten years later, in 2011, the Amsterdam Museum started the project re:DDS, the reconstruction of DDS. Not only to tell and show the story of this unique internet-historical monument of Amsterdam, but also –and more important- to raise awareness about the risk of the loss of our digital heritage. This was the beginning of our case study in web archaeology: how to excavate, reconstruct, preserve and sustainably store born-digital data to make it accessible to the future generations'?

2. CHALLENGES

The Digital City is a digital treasury from the early days of the web in the Netherlands. Our digital heritage is now at risk of being lost and with this we risk losing the early years of the internet in the Netherlands. We face many challenges on different levels.

2.1 Our Digital Heritage is Getting Lost!

"The world's digital heritage is at risk of being lost", the UNESCO wrote more than a decade ago in 2003, and "its preservation is an urgent issue of worldwide concern" [2]. The UNESCO acknowledged the historic value of our 'born digital' past and described it as "unique resources of human knowledge and expression". Only one year ago in 2015, Google's Vint Cerf warned, again, for a 'digital Dark Age': "Humanity's first steps into the digital world could be lost to future historians. We face a forgotten generation, or even a forgotten century. Our life, our memories (…) increasingly exist as bits of information - on our hard drives or in the cloud. But as technology moves on, they risk being lost in the wake of an accelerating digital revolution". [3]

2.2 Out of the Box

The acquiring and preservation of digitally created expressions of culture have different demands than the acquiring and preservation of physical objects. This is a new area for the heritage field in general and the Amsterdam Museum in particular. The museum has to cross boundaries, and get out of its comfort-zone to break new ground in dealing with digital heritage. To seek out new technologies, and new disciplines. To boldly dig what the museum has not dug before. How to dig up the lost hardware, software and data? And how to reconstruct a virtual city and create a representative version in which people can 'wander' through the different periods of DDS and experience the evolution of this unique city? The challenge is: can we –and how?- excavate and reconstruct The Digital City, from a virtual Atlantis to a virtual Pompeii?

2.3 Complexity of Born-digital Material

In crossing the boundaries and dealing with new (born-digital) material the museum encountered the following challenges:

- **Material:** born-digital material is complex and vulnerable and has various problems. Due to the rapid obsolescence of hardware and software and the vulnerability of digital files, data could be lost or become inaccessible. Another problem has to do with the authenticity. With born-digital objects it is no longer clear what belongs to the original object, and what has been added later. DDS is a complex information system with different applications. DDS was built on SUN systems, open source applications and self-written programs. How to preserve digital objects that are interactive, networked, process-oriented and context-dependent? And finally, important issues regarding to privacy, copyright and licensing form major questions.
- **Methods**: there is a difference between the (well known) web-harvesting and (relatively new) digital archaeology. Web-harvesting is the equivalent of taking a snapshot of a live object, while in our project we aim to recreate the object itself (or at least to create a representative version for people to access) from the 'dead' web. Since the data is no longer online, we first had to 'excavate' the digital artefacts. Fortunately, there are a some great internationally projects and initiatives that inspire us. With web-harvesting projects, such as The Wayback Machine and GeoCities, current data are harvested (scraped/mirrored) and displayed or visualized. In recent web archaeological projects, such as the restoration of the first website ever, info.cern.ch by CERN, and the project 'Digital Archaeology' of curator Jim Boulton, the original data and software are found and reconstructed, and shown on the original hardware or through emulation techniques.
- **Division of tasks:** who will take which responsibilities to retrieve, reconstruct, preserve and store born-digital heritage and make it accessible to the public? There is currently no central repository for tools (for example to read obsolete media), no comprehensive software library (for example to archive the old software, including Solaris, Windows, and MacOS), no central sustainable e-depot and infrastructure and there is a lack of web archaeological tools.
- **Historical (re)presentations and preservation strategies**: what are the (realistic) approaches into (re)presenting of historical data: 'historical true' or a 'quick and dirty'? How to preserve (and represent) historical digital-born data (migration, conversion, emulation, virtualization)?
- **Approach:** At present there is an alarming lack of awareness in the heritage field of the urgency (or funding) that our digital heritage is getting lost. We decided just to do it and act... (with lots of trial and error) and to start developing a roadmap to safeguard and preserve our born-digital past.

3. THE START

So in 2011 the Amsterdam Museum initiated the project re:DDS, the reconstruction of DDS, and started defining the objectives. To consider the project as a success the museum aimed to achieve the following goals:

- To give this unique (digital) heritage the place it deserves in the history of Amsterdam and to tell the story of DDS.
- To create a representative version of the internet-historical monument DDS in the Amsterdam Museum for people to visit and experience.
- To start a pilot in digital archaeology and share knowledge.
- To safeguard DDS in the collections of the heritage institutions for sustainable preservation.

To start the project the museum laid out the 're:DDS Roadmap':

1. Launch of the open history laboratory and a living virtual museum: http://hart.amsterdammuseum.nl/re-dds.
2. Bring Out Your Hardware & Finding Lost Data: crowdsourcing the archaeological remains (with The Grave Diggers Party as a kick-off event) and collect stories and memories.
3. The Rise of the Zombies: analyze and reconstruction.
4. Flight of the Zombies: presentation of the Lost & Found and a reconstruction.
5. Enlightenment: Let the Bytes Free!: Conclusions and evaluation.

Let's us take you back in time and share our first steps in web archaeology.

3.1 Crowdsourcing

First step was: let's find the data! There was no DDS archive, so the only chance of finding ánd piecing the DDS data together lay with people: old inhabitants, former DDS employees and volunteers. Web archaeologists Tjarda de Haan (guest e-curator of the Amsterdam Museum) and Paul Vogel (volunteer) had worked for DDS and had connections with former residents and (ex) DDS employees. To reach out to the crowd, calls and invitations were sent out through mailing lists, blogs and social media. Especially Twitter and Facebook proved to be indispensable tools to get for example old hardware which is no longer easily available.

3.2 Grave Diggers Party

On Friday, 13 May 2011 the Amsterdam Museum organized the 'Grave Diggers Party' with the Waag Society in Amsterdam. A party with a cause. A party to re-unite people and collect lost memories, both personal and digital.

We set up 'The Archaeological Site re:DDS', with a 'Working Space', with workstations to be able to collect the data on central storage, the 'Historical (e-)Depot'. Participants were able to dig in the Wayback Machine and store their excavations in the Historical (e-) Depot. Computers were used as excavators. Storage was used as buckets. UNIX commands and mice functioned as pades, pick-axe and trowels, scripts as metal detectors. Metadata were written down on 'find cards', so all lost and found artefacts (analog or digital) were documented: who brought in the material, were did it come from, how was it original used, what is the current state and who could we contact for more information. As well we set up a 'Museum Space', with 'Tourist Tours' in the 'Cabinet of Curiosities', where we show the lost and found artefacts (like old DDS servers, terminals, modem banks, tape robots, screenshots of images, manuals etc.).

So, we invited former residents, former employees and to DDS kindred souls to the Grave Diggers Party: "Help us dig up this unique city and be part of the first excavation work of the re:DDS!,". "Look at your attic and/or hard drives and bring all servers, modem banks, VT100 terminals, freezes, disks, scripts, zips, disks, floppies, tapes, backups, log files, videos, photos, screenshots and bring all your memories and stories you can find!".

Fifty enthusiastic (some international) cybernauts came along with full bags, hard drives and USB sticks to kick off the archaeological excavating. And after the party the Waag Society served for three weeks as an interactive archaeological site. In the 'Working Space' digital excavations were done, and the temporary exhibition was growing day by day.

During the 'Tourist Tours' people were interviewed and stimulated to share their stories and memories.

After three weeks of digging we found some great artefacts. De Waag Society found and donated two public terminals. The terminals were designed by Studio Stallinga in 1995. Residents of the city of Amsterdam who did not have a computer with a modem at home, could make free use of these public terminals in various public places in Amsterdam. The terminals were located among others in De Balie, the Amsterdam Museum, the public library and the city hall.

Former system administrators brought in discarded servers they rescued from the trash. We excavated servers with exotic names such as Alibaba, Shaman, Sarah and Alladin. Their pitiful status: cannibalized (robbed of components, the costs were very high so everything was always reused) or broken or the hard drives were wiped and reformatted. A former resident donated one of the first modem banks. Another former resident sent a specially made radio play for DDS in 1994, 'Station Het Oor' ('Station The Ear'). The play was made during the first six weeks after the opening of DDS. It was based on discussions and contributions of the first digital city dwellers. And we excavated thirty gigabytes of raw data, including backups of the squares, houses, projects. Furthermore we collected a huge amount of physical objects, like various manuals, magazines, photographs, videotapes and an original DDS mouse pad.

3.3 Freeze!

As cherry on the cake we excavated the most important and unique artefacts, namely the three DLT tapes, titled: 'Alibaba freeze', 'Shaman (FREEZ)' and 'dds freeze'. Together they form the 'freeze' of DDS of 1996. On 15 January 1996 DDS, the 'ten week experiment that got out of hand', existed for exactly two years. For the second anniversary of DDS, the pioneers of the social network sent a digital message in a bottle: "In the past two years, the Digital City has been focused on the future. This anniversary is a good time to take a moment to reflect on what happened the last two years in the city. High-profile discussions, large-scale urban expansion, people coming and going, friendships, loves and quarrels, the Digital City already has a long history. Three versions of the Digital City have been launched in two years. People come and go. Trends come and go. Houses are rigged and decorated. But where are all digital data kept? Who knows 5 years from now how DDS 3.0 looked like? The Digital City will be 'frozen' on Monday, January 15th at 18:00 o'clock. A snapshot with everything the city has to offer will be stored on tapes and will be hermetically sealed. The tapes with the data, along with a complete description of the programs and machines the city is run with and upon, will be deposited in an archive to study for archaeologists in a distant future". [4] The tapes however were never deposited, there was never any documentation made, and the tapes were more or less forgotten. Fortunately for us they were rediscovered a few weeks after the Grave Diggers Party.

The tapes came (of course) without the matching tape reader. After a frantic search an 'antique' tape reader was found in the National Library of the Netherlands in Den Hague, where it was used as ... a footstool. The Library, partner of the project, donated the tape reader to the project. In big excitement we started digging. But after two nights of reverse engineering the tape reader and the tape, we only found 70MB where we hoped to find 3GB. We were pretty disappointed. All sounds coming from the tape reader (tok, grrrrrr, beeeep) gave us the impression that containers of data were being flushed, and we thought the city had evaporated. We were puzzled. Had there ever been more data on the tape? Is this the 'freeze' we had hoped to find? What was broken, the tape reader or the tapes. Or both?

After our failed attempts to read the tapes and our bloodcurdling reverse engineering experiences (dismantling the reader and manually rewinding a tape), we called the Computer Museum of the University of Amsterdam for help. Would they be able to read

the historical data of DDS of the tapes? Engineer Henk Peek started to excavate immediately. Only a few weeks later he mailed us a status report:

> I've excavated about 11 Gigabytes of data tapes.

> It looks very positive!

3.4 'The 23 Things of Web Archaeology'

We made such an enormous progress in, what we now call, our 'slow data project'. It was time for the next step: the reconstruction, bit by bit, byte for byte. But given the bias in our knowledge and the complexity, nature and scope of the matter, we needed the expertise of specialist institutions. Together with our allied partners we started to explore the (im)possibilities of the reconstruction of born-digital data. We started to document our findings in the form of '23 Things'. [5] [6] In every 'Thing' we aim to explain our bottlenecks, choices and solutions. In addition, each partner will describe the state of affairs in the field of its expertise, and will illustrate this, where possible, with recent and relevant examples. In this way we share our joint experience and knowledge and in doing so we hope to lower the threshold for future web archaeological projects.

Data are the new clay, scripts are the new shovels and the web is the youngest layer of clay that we mine. Web archaeology is a new direction in e-culture in which we excavate relatively new (born-digital) material, that has only recently been lost, with relatively new (digital) tools. Both matter and methods to excavate and reconstruct our digital past are very young and still developing.

In 'The 23 Things of Web Archaeology' we research the following issues:

- **Born-digital material**. What is born digital heritage? How does it work? How it is stored and used?
- **Excavate and reconstruction.** What are the current methods and techniques how to excavate and reconstruct born-digital material (the physical and digital remains and the context in which they are found)?
- **Make accessible and presentation**. Finally, we look at how we can interpret the remains and the context in which they are found and make it accessible.

4. TOOLS AND METHODS

To start with the reconstruction of the lost and found DDS data the Amsterdam Museum partnered up with the University of Amsterdam. We brought together Gerard Alberts, teacher and author of history of computing, the students and former DDS administrators and programmers. The university provided the domain specific knowledge, student power and … made the first effort of reconstructing the DDS applications and interfaces. In the next section we describe the work broadly, from dealing with the lost and found data of the Grave Diggers Party to the first reconstruction. A more detailed elaboration will be written in the (short) future.

4.1 Physical Media

As stated before getting the raw data of the lost and found media turned out to be quite a challenge. Old hard discs connectors were no longer in use (for example SCSI's). In the end we managed to find (crowdsourcing!) an old system, a SUN server that was still fully operational, which we used to extract the raw data from the different storage systems that we recovered.

The tapes proved to be a little harder. The earlier mentioned specialist Henk Peek had to take apart another tape drive to fix it. After manually cleaning the drive he managed to read out large sections of the tapes that could be reconstructed with SUN restore. In the end he managed to read 11 GB of data of the tapes.

4.2 eForensics

In extracting the data from the file images we used eForensic methodologies. This mostly comes down to using low level block copy tools such as 'dd' (a UNIX command) and extracting the files from the resulting file system images. Multiple readout passes of the low level were done. And the resulting images were check summed to ensure that low level media errors were not made. Outliers were discarded.

4.3 DeNISTing

An often used eDiscovery tool is deNISTing, from the National Institute of Standards and Technology (NIST). We planned to use deNISTing not only as a way to identify data files but also to use the hash set to automatically identify OS type, applications and the locations of directories in the total data set. However it turned out to be not as useful as expected. Since the NIST hash set is geared towards Windows.

4.4 Virtual Machines

One of the chosen methods of the project was to emulate all the systems in the original working state, through a combination of virtual machines and software defined networking. We planned to emulate the entire cluster. Sparc emulation however proved to be painfully slow and therefor unworkable.

4.5 Reconstructing

Compiling the C code, most of DDS code was written in C, proved to be an easier road than emulation. Even though include files and libraries had changed through time. The students managed to compile most of the programs in the end.

4.6 First Results

In the end the students, with support of former DDS employees, were able to A. reconstruct DDS3.0, the third and most known DDS interface, and B. build a replica using emulation, DDS4.0. An enormous achievement, and a huge step into the new (scientific) discipline web archaeology! Next step is to incorporate these results in the museum. DDS had already been included in the permanent collection of the museum, in a minimal version. The audience can see our first web archaeological excavations, the DDS 'avatars', take place behind the original DDS public terminal and watch television clips about DDS. In the nearby future we aim to enable people to interact with DDS, by taking a walk through the historical digital city, and 'experience' how the internet was at the beginning and how did it look like in the 20th century.

5. NEXT LEVEL

To enter the next level of our project we teamed up with the University of Amsterdam, Dutch Institute for Sound and Vision, Waag Society and started the project "The Digital City revives". With the joint forces of museums, innovators, creative industries,

archives and scientists we will face our last major challenges: how to open up and sustainable store the DDS data into e-depots, how to contribute to a hands-on jurisprudence of privacy and copyright for future web archaeological projects, how to share our knowledge and lower the threshold for future web archaeological projects and how to present the DDS born-digital heritage in a museum context for future generations?

Our goals of our project "The Digital City revives" are:

- Reconstruct and preserve DDS.
- Provide insight into the (existing and new) processes, techniques and methods for born-digital material and the context in which they are found, to excavate and reconstruct.
- Ask attention to the danger of 'digital amnesia'.
- To provide museums and organizations with specialized knowledge about the reconstruction of born-digital heritage and lower the threshold for future web archaeological projects. Disseminating knowledge about new standards for archives on the storage of digital-born heritage in 'The 23 Things of Web Archaeology' and a 'DIY Handbook of Web Archaeology'.
- Make DDS data 'future-proof' in making the digital cultural heritage:
    - Visible (content): promoting (re)use of DDS.
    - Usable (connection): improving (re)use of DDS collection by making it available by linking and enriching data.
    - Preservable (services): maintain DDS sustainable and keep it accessible.

To be continued!

6. ACKNOWLEDGMENTS

We would like to thank our partners for their enormous contributions to our project. Our partners are (till now): Amsterdam City Archives, Creative Industries Fund NL, Digital Heritage Netherlands, Dutch Computer Heritage Foundation, Dutch Cultural Coalition for Digital Durability, International Institute of Social History, Karin Spaink, independent researcher, LIMA, National Coalition on Digital Preservation, National Library of the Netherlands, Netherlands Institute for Sound and Vision, old inhabitants, (ex) DDS employees, especially Paul Vogel and Michaël van Eeden, and DDS affiliated web archaeologists, University of Amsterdam, University of Amsterdam Computer Museum and Waag Society.

7. REFERENCES

[1] This article is partly based on an earlier article published in "Data Drift. Archiving Media and Data Art in the 21st Century" (October 2015), published by RIXC, LiepU MPLab, http://rixc.org/en/acoustsicspace/all/, by Tjarda de Haan and Paul Vogel.

[2] UNESCO published the Charter on the Preservation of Digital Heritage (October 2003). http://portal.unesco.org/en/ev.php-URL_ID=17721&URL_DO=DO_TOPIC&URL_SECTION=201.html

[3] http://www.theguardian.com/technology/2015/feb/13/google-boss-warns-forgotten-century-email-photos-vint-cerf

[4] http://web.archive.org/web/20100830120819/http://www.almedia.nl/DDS/Nieuws/freeze.html

[5] The 23 Things of Web archaeology is based on the Learning 2.0 – 23 Things, http://plcmcl2-things.blogspot.nl/, a program of Helene Blowers and inspired by the many sites of the 23 Things that followed.

[6] http://hart.amsterdammuseum.nl/23dingenwebarcheologie