THE SECOND DIGITAL PRESERVATION CHALLENGE

July 2008

Juan-José Boté Vericad
Universitat Oberta de Catalunya
jbotev@uoc.edu

INDEX

INTRODUCTION

The Second Digital Preservation Challenge is an initiative of the Digital Preservation Europe (DPE) consortium that invites individual participants from any area with an interest in computer science to propose new ideas for digital preservation that overcome the barriers hindering access to five digital objects, each accompanied by a scenario based on real-life situations: Master back-up on a tape, computer gaming platform, obsolete database, electronic art and web archiving.

This report is organised in tasks. Those task were Legacy application file, Images from a Legacy Computer Gaming Platform, Obsolete Database, Electronic Art and Web Archiving.

At the end the tasks there is an appendix relative to the software used on the challenge.

SCENARIO 1- Legacy Application File

Your company archivist discovered an old tape in a store-room. The content is not known but the label "Master Backup" suggests that is highly valuable.
There were four types of files on it, one type of text documents, one type of graphics and two unknown file types.
You are asked to identify the unknown file types and display the content of the given sample files in such a way that they may be used in a different application.
You are also asked to design an appropriate preservation strategy that will facilitate access to such records, and that can be applied, as far as possible, in an automatic manner. Moreover, you are asked to estimate the cost/effort required to deploy the strategies you propose.

Task

a) Identify the type of content of the unknown files.
b) Propose one or more suitable preservation strategies and provide a thorough description that highlights their advantages and disadvantages.
c) Implement a preservation strategy capable of mass handling of files of this type, giving an estimate of the cost/effort for deploying the strategies.
d) Apply the preservation strategy to the objects and display the files.
e) Analyse the benefits and the shortcomings of the preservation strategy.

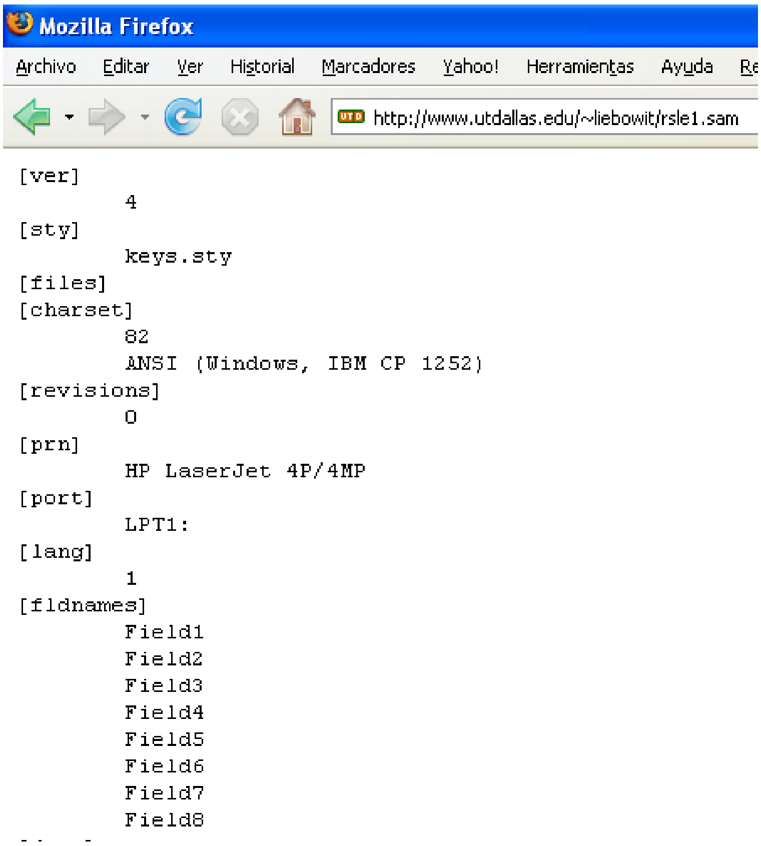
Task 1a –Identify the type of content of the unknown files._

Unknown files were in two folders. Folder unknown 1 and unknown 2.

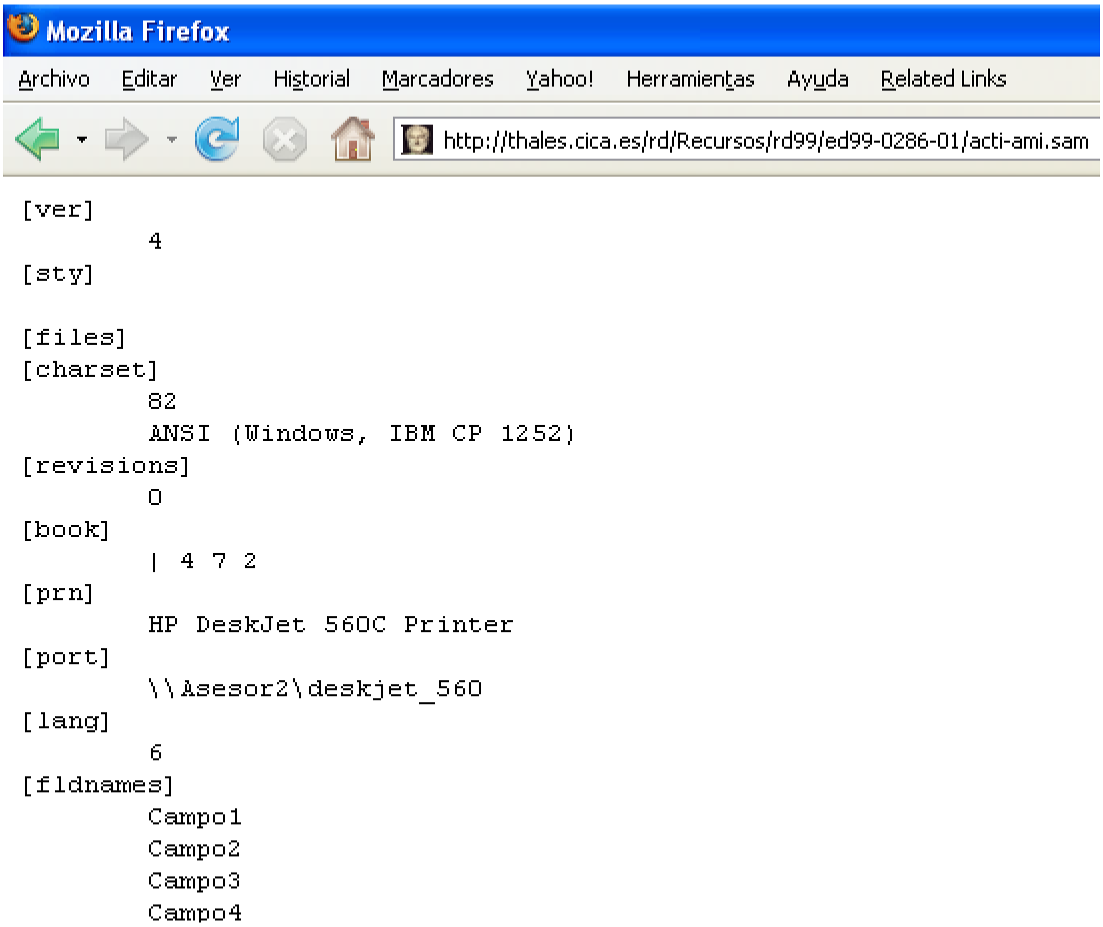
FOLDER UNKNOWN1
Folder unknown 1 contained one file called NEWS.
The file NEWS had no extension associated. The first step was to examine the file. With a text-plain editor, Notepad the file had been opened to see what could be the content.
A part of information of the file is shown,

```
[ver]
     4
[sty]

[files]
[charset]
     82
     ANSI (Windows, IBM CP 1252)
[revisions]
     0
[prn]
     Microsoft Office Document Image Writer
[port]
     Ne01:
[lang]
     18
[fldnames]
     Field1
     Field2
     Field3
     Field4
     Field5
     Field6
     Field7
     Field8
```

Some similar files to the file NEWS were found[1] on the internet[2] as it is seen on the images below.


Mozilla Firefox
Archivo Editar Ver Historial Marcadores Yahoo! Herramientas Ayuda Re
http://www.utdallas.edu/~liebowit/rsle1.sam

```
[ver]
        4
[sty]
        keys.sty
[files]
[charset]
        82
        ANSI (Windows, IBM CP 1252)
[revisions]
        0
[prn]
        HP LaserJet 4P/4MP
[port]
        LPT1:
[lang]
        1
[fldnames]
        Field1
        Field2
        Field3
        Field4
        Field5
        Field6
        Field7
        Field8
```


Mozilla Firefox
Archivo Editar Ver Historial Marcadores Yahoo! Herramientas Ayuda Related Links
http://thales.cica.es/rd/Recursos/rd99/ed99-0286-01/acti-ami.sam

```
[ver]
        4
[sty]

[files]
[charset]
        82
        ANSI (Windows, IBM CP 1252)
[revisions]
        0
[book]
        | 4 7 2
[prn]
        HP DeskJet 560C Printer
[port]
        \\Asesor2\deskjet_560
[lang]
        6
[fldnames]
        Campo1
        Campo2
        Campo3
        Campo4
```

[1] Plan de Acción 1998-2000.(2000) Sociedad Matemática Thales. Consejeria de Educación y Ciencia. Junta de Andalucia. 1998. http://thales.cica.es/rd/Recursos/rd99/ed99-0286-01/acti-ami.sam

[2] Stan Liebowtiz Home Page. (July, 2006.) University of Texas at Dallas http://www.utdallas.edu/~liebowit/rsle1.sam;

Having found similar files on the internet,.sam extension belongs to Lotus 1-2-3 Ami Pro. File NEWS had been opened with Lotus 1-2-3 SmartSuite.

Ami Pro data file[3].

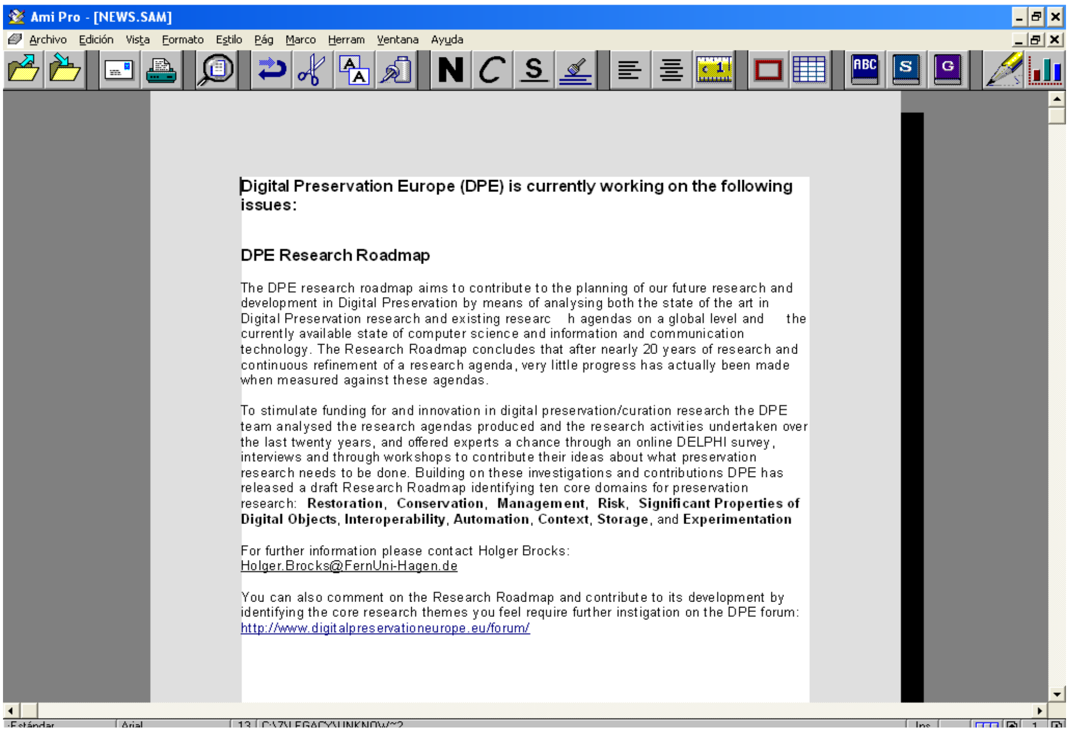
Ami Pro is a software from Lotus Development Corporation and a part of Lotus SmartSuite.

> "Lotus[4] was founded in 1982 by partners Mitch Kapor and Jonathan Sachs with backing from Ben Rosen. Lotus' first product was presentation software for the Apple II known as Lotus Executive Briefing System. Kapor founded Lotus after leaving his post as head of development at VisiCorp (the makers of the Visicalc spreadsheet) and selling all his rights to the VisiPlot and VisiTrend products to VisiCorp.
>
> Kapor and Sachs produced an integrated spreadsheet and graphics program, Lotus who was a clearly superior product than Visicalc. Lotus released Lotus 1-2-3 in January 1983. The name referred to the three ways the product could be used, as a spreadsheet, graphics package, and database manager. In fact, the latter two functions were less often used, nonetheless 1-2-3 was the most powerful spreadsheet program available.
>
> **Lotus Software** (called **Lotus Development Corporation** before its acquisition by IBM) is a software company with headquarters in Westford, Massachusetts. Lotus is most commonly known for the **Lotus 1-2-3 spreadsheet** application, by far the first feature-rich, user friendly, highly reliable and WYSIWYG-enabled product to become widely available in the early days of the IBM PC, when there was no Graphical user interface.
>
> In the 1990s, to compete with Microsoft's Windows applications, Lotus had to buy in products such as Freelance Graphics, Ami Pro (word processor), Approach (database), and Threadz, which became Lotus Organizer. Several of the these (1-2-3, Freelance Graphics, Ami Pro, Approach, and Lotus Organizer) were bundled together under the name Lotus SmartSuite."

Actually Lotus belongs to IBM Corporation[5].

On the images below, there is part of the information visualization of the files with AmiPro which is a part of the Lotus 1-2-3 SmartSuite. The document contained 13 pages. The two first pages are shown as a screenshot.

VISUALIZATION OF THE INFORMATION



Digital Preservation Europe (DPE) is currently working on the following issues:

DPE Research Roadmap

The DPE research roadmap aims to contribute to the planning of our future research and development in Digital Preservation by means of analysing both the state of the art in Digital Preservation research and existing researc h agendas on a global level and the currently available state of computer science and information and communication technology. The Research Roadmap concludes that after nearly 20 years of research and continuous refinement of a research agenda, very little progress has actually been made when measured against these agendas.

To stimulate funding for and innovation in digital preservation/curation research the DPE team analysed the research agendas produced and the research activities undertaken over the last twenty years, and offered experts a chance through an online DELPHI survey, interviews and through workshops to contribute their ideas about what preservation research needs to be done. Building on these investigations and contributions DPE has released a draft Research Roadmap identifying ten core domains for preservation research: **Restoration**, **Conservation**, **Management**, **Risk**, **Significant Properties of Digital Objects**, **Interoperability**, **Automation**, **Context**, **Storage**, and **Experimentation**

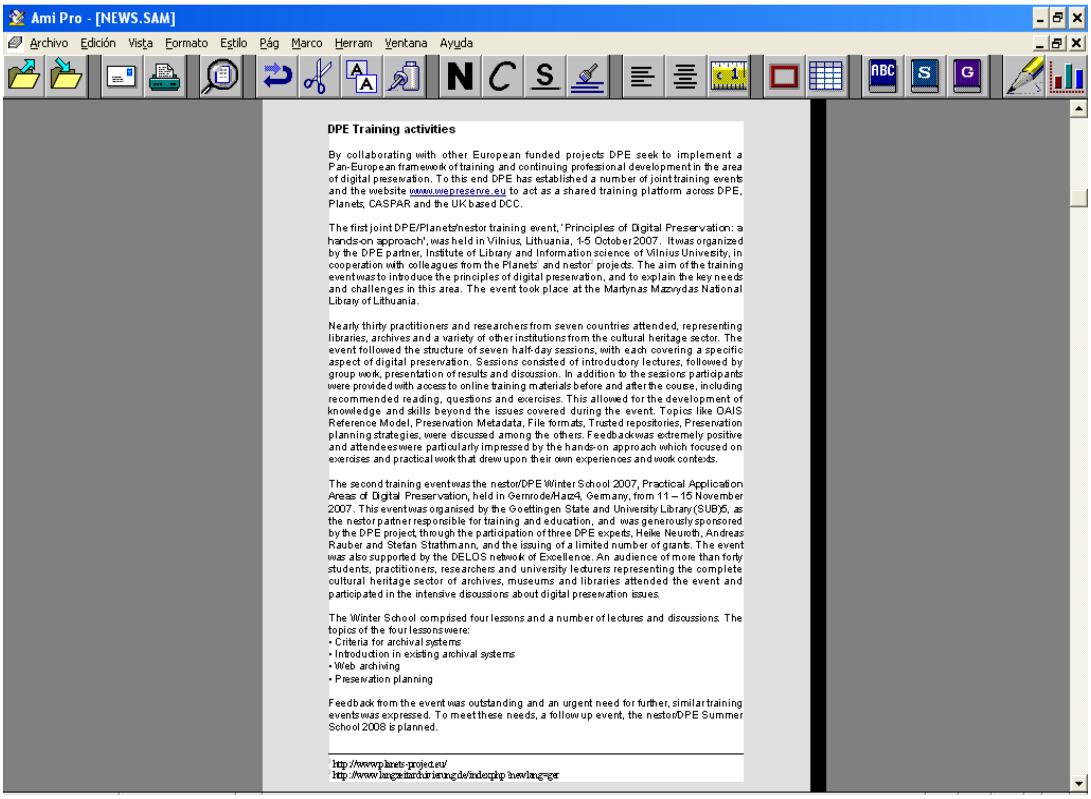
For further information please contact Holger Brocks:
Holger.Brocks@FernUni-Hagen.de

You can also comment on the Research Roadmap and contribute to its development by identifying the core research themes you feel require further instigation on the DPE forum:
http://www.digitalpreservationeurope.eu/forum/

[3] http://file-extension.net/seeker/file_extension_sam

[4] Adapted from: http://en.wikipedia.org/wiki/Lotus_Software

[5] http://www-306.ibm.com/software/lotus/


Ami Pro - [NEWS.SAM]
Archivo Edición Vista Formato Estilo Pág Marco Herram Ventana Ayuda

DPE Training activities

By collaborating with other European funded projects DPE seek to implement a Pan-European framework of training and continuing professional development in the area of digital preservation. To this end DPE has established a number of joint training events and the website www.wepreserve.eu to act as a shared training platform across DPE, Planets, CASPAR and the UK based DCC.

The first joint DPE/Planets/nestor training event, 'Principles of Digital Preservation: a hands-on approach', was held in Vilnius, Lithuania, 1-5 October 2007. It was organized by the DPE partner, Institute of Library and Information science of Vilnius University, in cooperation with colleagues from the Planets[1] and nestor[2] projects. The aim of the training event was to introduce the principles of digital preservation, and to explain the key needs and challenges in this area. The event took place at the Martynas Mazvydas National Library of Lithuania.

Nearly thirty practitioners and researchers from seven countries attended, representing libraries, archives and a variety of other institutions from the cultural heritage sector. The event followed the structure of seven half-day sessions, with each covering a specific aspect of digital preservation. Sessions consisted of introductory lectures, followed by group work, presentation of results and discussion. In addition to the sessions participants were provided with access to online training materials before and after the couse, including recommended reading, questions and exercises. This allowed for the development of knowledge and skills beyond the issues covered during the event. Topics like OAIS Reference Model, Preservation Metadata, File formats, Trusted repositories, Preservation planning strategies, were discussed among the others. Feedback was extremely positive and attendees were particularly impressed by the hands-on approach which focused on exercises and practical work that drew upon their own experiences and work contexts.

The second training event was the nestor/DPE Winter School 2007, Practical Application Areas of Digital Preservation, held in Gernrode/Harz4, Germany, from 11 – 15 November 2007. This event was organised by the Goettingen State and University Library (SUB)5, as the nestor partner responsible for training and education, and was generously sponsored by the DPE project, through the participation of three DPE experts, Heike Neuroth, Andreas Rauber and Stefan Strathmann, and the issuing of a limited number of grants. The event was also supported by the DELOS network of Excellence. An audience of more than forty students, practitioners, researchers and university lecturers representing the complete cultural heritage sector of archives, museums and libraries attended the event and participated in the intensive discussions about digital preservation issues.

The Winter School comprised four lessons and a number of lectures and discussions. The topics of the four lessons were:

- Criteria for archival systems
- Introduction in existing archival systems
- Web archiving
- Preservation planning

Feedback from the event was outstanding and an urgent need for further, similar training events was expressed. To meet these needs, a follow up event, the nestor/DPE Summer School 2008 is planned.

[1] http://www.planets-project.eu/
[2] http://www.langzeitarchivierung.de/index.php?newlang=ger

FOLDER UNKNOWN2
This folder contained 3 files called:

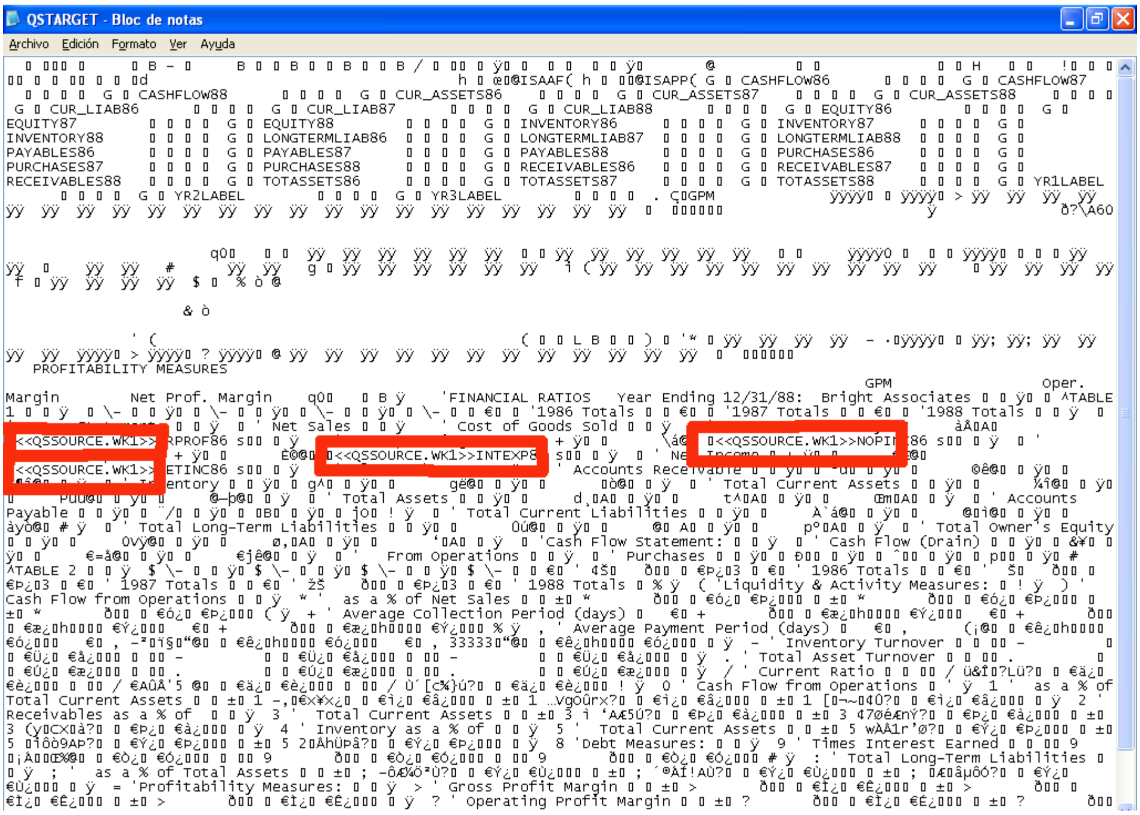
LONDON
QSSOURCE
QSSTARGET

With no extension associated.

Identification: the three unknown files were spreadsheets that belonged also to Lotus 1-2-3 SmartSuite.
Those spreadsheets were done in Lotus1-2-3 with the extension .wk1[6]
The first step to read those files has been to open them with Notepad. Normally open non-identified type of files with a text editor may give us clues about the kind of file were going to treat. This is due because at the beginning of the files some headers are often placed and may be read and identified

QSSTARGET file, had some references to a file **QSSOURCE.WK1** as it is seen in the image below.

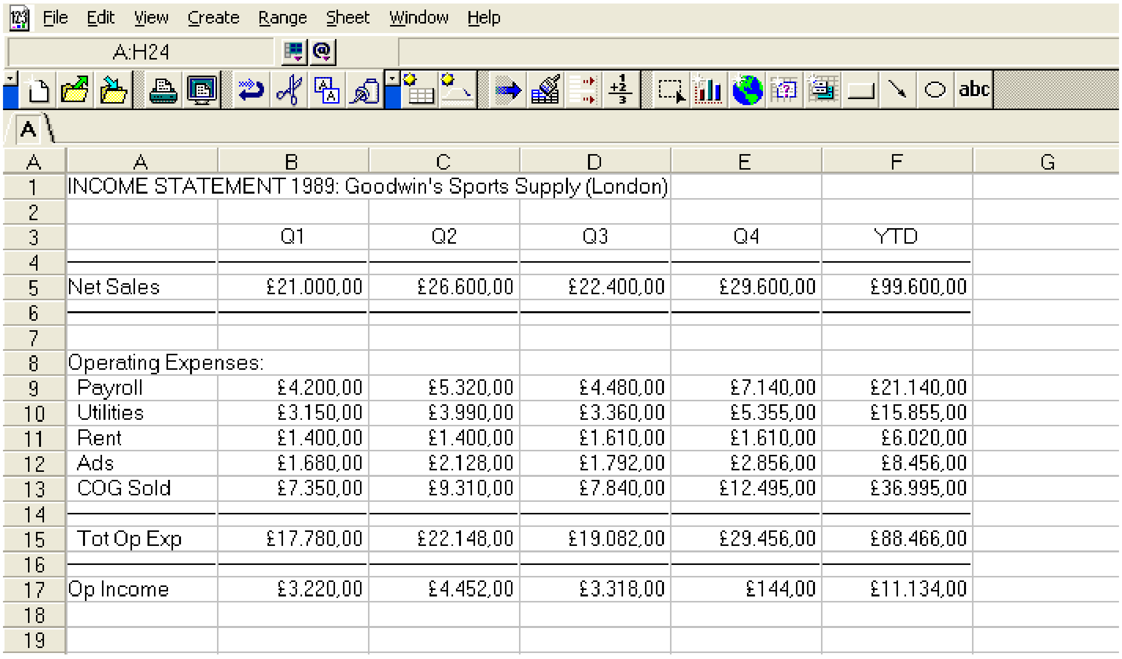
[6] http://filext.com/file-extension/WK1



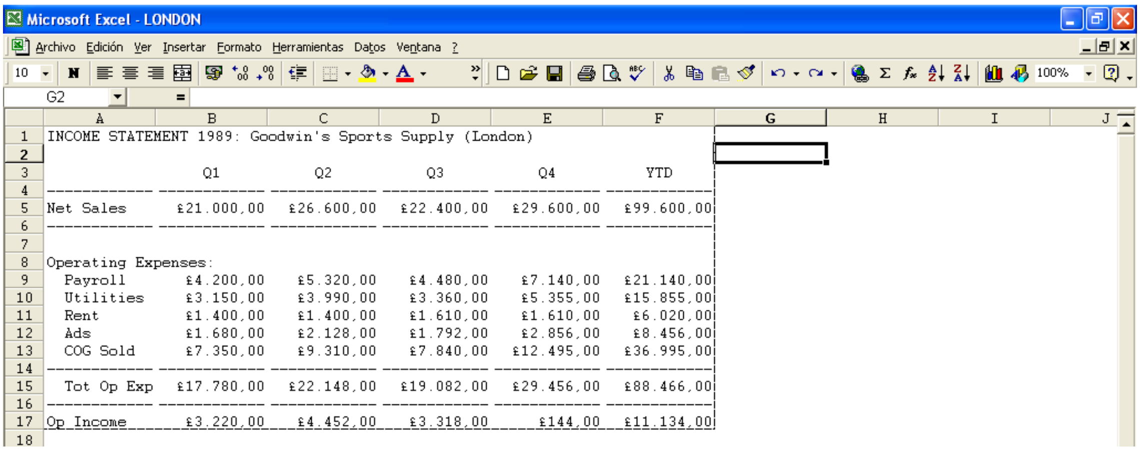
```
 ▯ ▯▯▯ ▯      ▯ B – ▯      B ▯ ▯ B ▯ ▯ B ▯ ▯ B / ▯ ▯▯ ▯ ÿ▯ ▯   ▯ ▯   ▯ ▯ ÿ▯       @          ▯ ▯           ▯ ▯ H   ▯ ▯   !▯ ▯ ▯
▯▯ ▯ ▯ ▯▯ ▯ ▯ ▯d                                   h ▯ œ▯@ISAAF( h ▯ ▯▯@ISAPP( G ▯ CASHFLOW86     ▯ ▯ ▯ ▯  G ▯ CASHFLOW87
 ▯ ▯ ▯ ▯  G ▯ CASHFLOW88      ▯ ▯ ▯ ▯  G ▯ CUR_ASSETS86     ▯ ▯ ▯ ▯  G ▯ CUR_ASSETS87     ▯ ▯ ▯ ▯  G ▯ CUR_ASSETS88     ▯ ▯ ▯ ▯
 G ▯ CUR_LIAB86      ▯ ▯ ▯ ▯  G ▯ CUR_LIAB87      ▯ ▯ ▯ ▯  G ▯ CUR_LIAB88      ▯ ▯ ▯ ▯  G ▯ EQUITY86      ▯ ▯ ▯ ▯  G ▯
EQUITY87        ▯ ▯ ▯ ▯  G ▯ EQUITY88        ▯ ▯ ▯ ▯  G ▯ INVENTORY86     ▯ ▯ ▯ ▯  G ▯ INVENTORY87     ▯ ▯ ▯ ▯  G ▯
INVENTORY88     ▯ ▯ ▯ ▯  G ▯ LONGTERMLIAB86  ▯ ▯ ▯ ▯  G ▯ LONGTERMLIAB87  ▯ ▯ ▯ ▯  G ▯ LONGTERMLIAB88  ▯ ▯ ▯ ▯  G ▯
PAYABLES86      ▯ ▯ ▯ ▯  G ▯ PAYABLES87      ▯ ▯ ▯ ▯  G ▯ PAYABLES88      ▯ ▯ ▯ ▯  G ▯ PURCHASES86     ▯ ▯ ▯ ▯  G ▯
PURCHASES87     ▯ ▯ ▯ ▯  G ▯ PURCHASES88     ▯ ▯ ▯ ▯  G ▯ RECEIVABLES86   ▯ ▯ ▯ ▯  G ▯ RECEIVABLES87   ▯ ▯ ▯ ▯  G ▯
RECEIVABLES88   ▯ ▯ ▯ ▯  G ▯ TOTASSETS86     ▯ ▯ ▯ ▯  G ▯ TOTASSETS87     ▯ ▯ ▯ ▯  G ▯ TOTASSETS88     ▯ ▯ ▯ ▯  G ▯ YR1LABEL
     ▯ ▯ ▯ ▯  G ▯ YR2LABEL        ▯ ▯ ▯ ▯  G ▯ YR3LABEL        ▯ ▯ ▯ ▯  . Ç▯GPM              ÿÿÿÿ▯ ▯ ÿÿÿÿ▯ > ÿÿ  ÿÿ  ÿÿ  ÿÿ
ÿÿ  ÿÿ  ÿÿ  ÿÿ  ÿÿ  ÿÿ  ÿÿ  ÿÿ  ÿÿ  ÿÿ  ÿÿ  ÿÿ  ÿÿ  ÿÿ  ÿÿ  ÿÿ  ÿÿ  ÿÿ  ▯  ▯▯▯▯▯▯                    ÿ                ð?\A60


                   q0▯   ▯ ▯  ÿÿ  ÿÿ  ÿÿ  ÿÿ  ÿÿ  ÿÿ  ▯ ▯ ÿÿ  ÿÿ  ÿÿ  ÿÿ  ÿÿ  ÿÿ   ▯ ▯     ÿÿÿÿ0 ▯   ▯ ▯ ÿÿÿÿ▯ ▯ ▯ ▯ ÿÿ
ÿÿ  ▯     ÿÿ  ÿÿ   #     ÿÿ  ÿÿ   g ▯ ÿÿ  ÿÿ  ÿÿ  ÿÿ  ÿÿ  ÿÿ    ì ( ÿÿ  ÿÿ  ÿÿ  ÿÿ  ÿÿ  ÿÿ  ÿÿ  ÿÿ  ÿÿ  ÿÿ    ▯ ÿÿ  ÿÿ  ÿÿ  ÿÿ
 f ▯ ÿÿ  ÿÿ  ÿÿ  ÿÿ   $ ▯  % ò @

                 & ò

               ' (                                        ( ▯ ▯ L B ▯ ▯ ) ▯ '* ▯ ÿÿ  ÿÿ  ÿÿ  ÿÿ  – ·▯ÿÿÿÿ▯ ▯ ÿÿ; ÿÿ; ÿÿ  ÿÿ
ÿÿ  ÿÿ  ÿÿÿÿ▯ > ÿÿÿÿ▯ ? ÿÿÿÿ▯ @ ÿÿ  ÿÿ  ÿÿ  ÿÿ  ÿÿ  ÿÿ  ÿÿ  ÿÿ  ÿÿ  ÿÿ  ÿÿ  ÿÿ  ▯  ▯▯▯▯▯▯
   PROFITABILITY MEASURES
                                                                                    GPM                 Oper.
Margin       Net Prof. Margin    q0▯   ▯ B ÿ    'FINANCIAL RATIOS   Year Ending 12/31/88:  Bright Associates ▯ ▯ ÿ▯ ▯ ^TABLE
1 ▯ ▯ ÿ  ▯ \- ▯ ▯ ÿ▯ ▯ \- ▯ ▯ ÿ▯ ▯ \- ▯ ▯ ÿ▯ ▯ \- ▯ ▯ €▯ ▯ '1986 Totals ▯ ▯ €▯ ▯ '1987 Totals ▯ ▯ €▯ ▯ '1988 Totals ▯ ▯ ÿ  ▯
                  ▯ ▯ ÿ  ▯ ' Net Sales ▯ ▯ ÿ    ' Cost of Goods Sold ▯ ▯ ÿ                             àÂ▯A▯
<<QSSOURCE.WK1>>RPROF86 s▯▯ ▯ ÿ                     + ÿ▯ ▯       \á@  ▯<<QSSOURCE.WK1>>NOPI 86 s▯▯ ▯ ÿ  ▯ '
                  + ÿ▯ ▯      È©@▯ ▯<<QSSOURCE.WK1>>INTEXP8 s▯▯ ▯ ÿ  ▯ ' Ne                            @▯
<<QSSOURCE.WK1>>ETINC86 s▯▯ ▯ ÿ                      ' Accounts Receivable ▯ ▯ ÿ▯ ▯  ▯▯ ▯ ÿ▯ ▯      ©ê@▯ ▯ ÿ▯ ▯
                entory ▯ ▯ ÿ▯ ▯ g^▯ ▯ ÿ▯ ▯      gë@▯ ▯ ÿ▯ ▯      ▯ò@▯ ▯ ÿ  ▯ ' Total Current Assets ▯ ▯ ÿ▯ ▯      ¼î@▯ ▯ ÿ▯
▯     Puü@▯ ▯ ÿ▯ ▯      @–þ@▯ ▯ ÿ  ▯ ' Total Assets ▯ ▯ ÿ▯ ▯     d¸▯A▯ ▯ ÿ▯ ▯     t^▯A▯ ▯ ÿ▯ ▯     Œm▯A▯ ▯ ÿ  ▯ ' Accounts
Payable ▯ ▯ ÿ▯ ▯ ¨/▯ ▯ ÿ▯ ▯ ▯B▯ ▯ ÿ▯ ▯ jo▯ ! ÿ  ▯ ' Total Current Liabilities ▯ ▯ ÿ▯ ▯     À`á@▯ ▯ ÿ▯ ▯     @▯i@▯ ▯ ÿ▯ ▯
àyó@▯ # ÿ  ▯ ' Total Long-Term Liabilities ▯ ▯ ÿ▯ ▯     Óú@▯ ▯ ÿ▯ ▯     @▯ A▯ ▯ ÿ▯ ▯     pº▯A▯ ▯ ÿ  ▯ ' Total Owner's Equity
▯ ▯ ÿ▯ ▯     Ový@▯ ▯ ÿ▯ ▯     ø,▯A▯ ▯ ÿ▯ ▯     '▯A▯ ▯ ÿ  ▯ 'Cash Flow Statement: ▯ ▯ ÿ  ▯ ' Cash Flow (Drain) ▯ ▯ ÿ▯ ▯ &¥▯ ▯
ÿ▯ ▯     €=â@▯ ▯ ÿ▯ ▯     €jê@▯ ▯ ÿ  ▯ '   From Operations ▯ ▯ ÿ  ▯ ' Purchases ▯ ▯ ÿ▯ ▯ Ð▯▯ ▯ ÿ▯ ▯ ˆ▯▯ ▯ ÿ▯ ▯ p▯▯ ▯ ÿ▯ #
^TABLE 2 ▯ ▯ ÿ  $ \- ▯ ▯ ÿ▯ $ \- ▯ ▯ ÿ▯ $ \- ▯ ▯ ÿ▯ $ \- ▯ ▯ €▯ ' ¢Š▯   ð▯▯ ▯ €Þ¿▯3 ▯ €▯ ' 1986 Totals ▯ ▯ €▯ '  Š▯   ð▯▯ ▯
€Þ¿▯3 ▯ €▯ ' 1987 Totals ▯ ▯ €▯ ' žŠ    ð▯▯ ▯ €Þ¿▯3 ▯ €▯ ' 1988 Totals ▯ % ÿ  ( 'Liquidity & Activity Measures: ▯ ! ÿ  ) '
Cash Flow from Operations ▯ ▯ ÿ  * '  as a % of Net Sales ▯ ▯ ±▯ *       ð▯▯ ▯ €ó¿▯ €Þ¿▯▯▯ ▯ ±▯ *       ð▯▯ ▯ €ó¿▯ €Þ¿▯▯▯ ▯
±▯ *       ð▯▯ ▯ €ó¿▯ €Þ¿▯▯▯ ( ÿ  + ' Average Collection Period (days) ▯   €▯ +       ð▯▯ ▯ €æ¿▯h▯▯▯▯ €Ÿ¿▯▯▯   €▯ +       ð▯▯
▯ €æ¿▯h▯▯▯▯ €Ÿ¿▯▯▯   €▯ +      ð▯▯ ▯ €æ¿▯h▯▯▯▯ €Ÿ¿▯▯▯ % ÿ  , ' Average Payment Period (days) ▯   €▯ ,      (¡@▯ ▯ €ê¿▯h▯▯▯▯
€ó¿▯▯▯   €▯ , –²▯ï§▯"@▯ ▯ €ê¿▯h▯▯▯▯ €ó¿▯▯▯   €▯ , 33333▯"@▯ ▯ €ê¿▯h▯▯▯▯ €ó¿▯▯▯ ▯ ÿ  – ' Inventory Turnover ▯ ▯ ▯▯ –       ▯
▯ €Ü¿▯ €å¿▯▯▯ ▯ ▯▯ –         ▯ ▯ €Ü¿▯ €å¿▯▯▯ ▯ ▯▯ –         ▯ ▯ €Ü¿▯ €å¿▯▯▯ ▯ ÿ  . ' Total Asset Turnover ▯ ▯ ▯▯ .        ▯
▯ €Ú¿▯ €æ¿▯▯▯ ▯ ▯▯ .         ▯ ▯ €Ú¿▯ €æ¿▯▯▯ ▯ ▯▯ .         ▯ ▯ €Ú¿▯ €æ¿▯▯▯ ▯ ÿ  / ' Current Ratio ▯ ▯ ▯▯ / ü&Î▯?Lü?▯ ▯ €ä¿▯
€è¿▯▯▯ ▯ ▯▯ / €AÛÂ'5 @▯ ▯ €ä¿▯ €è¿▯▯▯ ▯ ▯▯ / Ù´[c¾}ú?▯ ▯ €ä¿▯ €è¿▯▯▯ ! ÿ  0 ' Cash Flow from Operations ▯ ' ÿ  1 '  as a % of
Total Current Assets ▯ ▯ ±▯ 1 –,▯€x¥x¿▯ ▯ €ì¿▯ €â¿▯▯▯ ▯ ±▯ 1 …Vgoûrx?▯ ▯ €ì¿▯ €â¿▯▯▯ ▯ ±▯ 1 [▯¬~▯¢Ò?▯ ▯ €ì¿▯ €â¿▯▯▯ ▯ ÿ  2 '
Receivables as a % of  ▯ ▯ ÿ  3 '  Total Current Assets ▯ ▯ ±▯ 3 ì 'AÆ5Ú?▯ ▯ €Þ¿▯ €à¿▯▯▯ ▯ ±▯ 3 ¢7øéÆnÝ?▯ ▯ €Þ¿▯ €à¿▯▯▯ ▯ ±▯
3 (y▯Cx▯à?▯ ▯ €Þ¿▯ €à¿▯▯▯ ▯ ÿ  4 ' Inventory as a % of ▯ ▯ ÿ  5 '  Total Current Assets ▯ ▯ ±▯ 5 wÀÂ1r'ø?▯ ▯ €Ÿ¿▯ €Þ¿▯▯▯ ▯ ±▯
5 ▯îòò9AP?▯ ▯ €Ÿ¿▯ €Þ¿▯▯▯ ▯ ±▯ 5 2▯ÂhÜPâ?▯ ▯ €Ÿ¿▯ €Þ¿▯▯▯ ▯ ÿ  8 'Debt Measures: ▯ ▯ ÿ  9 ' Times Interest Earned ▯ ▯ ▯▯ 9
▯¡À▯▯Œ%@▯ ▯ €Ò¿▯ €Ó¿▯▯▯ ▯ ▯▯ 9       ð▯▯ ▯ €Ò¿▯ €Ó¿▯▯▯ ▯ ▯▯ 9       ð▯▯ ▯ €Ò¿▯ €Ó¿▯▯▯ # ÿ  : ' Total Long-Term Liabilities ▯
▯ ÿ  ; '  as a % of Total Assets ▯ ▯ ±▯ ; –ôÆ¼Ö²Ù?▯ ▯ €Ÿ¿▯ €Ù¿▯▯▯ ▯ ±▯ ; ´®ÀÍ!AÙ?▯ ▯ €Ÿ¿▯ €Ù¿▯▯▯ ▯ ±▯ ; ▯Æ▯âµôÓ?▯ ▯ €Ÿ¿▯
€Ù¿▯▯▯ ▯ ÿ  = 'Profitability Measures: ▯ ▯ ÿ  > ' Gross Profit Margin ▯ ▯ ±▯ >       ð▯▯ ▯ €Ì¿▯ €Ê¿▯▯▯ ▯ ±▯ >       ð▯▯ ▯
€Ì¿▯ €Ê¿▯▯▯ ▯ ±▯ >       ð▯▯ ▯ €Ì¿▯ €Ê¿▯▯▯ ▯ ÿ  ? ' Operating Profit Margin ▯ ▯ ±▯ ?       ð▯▯ ▯ €Ì¿▯ €É¿▯▯▯ ▯ ±▯ ?       ð▯▯
```

Afterwards, looking for information over the internet about the extension .wk1, it was clear that there were spreadsheet files done with Lotus 1-2-3.

Once renamed the files with WK1 extension, **QSSTARGET.WK1**, **QSSOURCE.WK1** and **LONDON.WK1** and openend with Lotus 1-2-3, there were some elements on the spreadsheets that were missed as the currency. As operations were supposed to be effective in London it was supposed that the currency were British Pounds (£)
The data from the **LONDON** file was as it is shown on the images below. It was also possible to open it with Excel® spreadsheet.

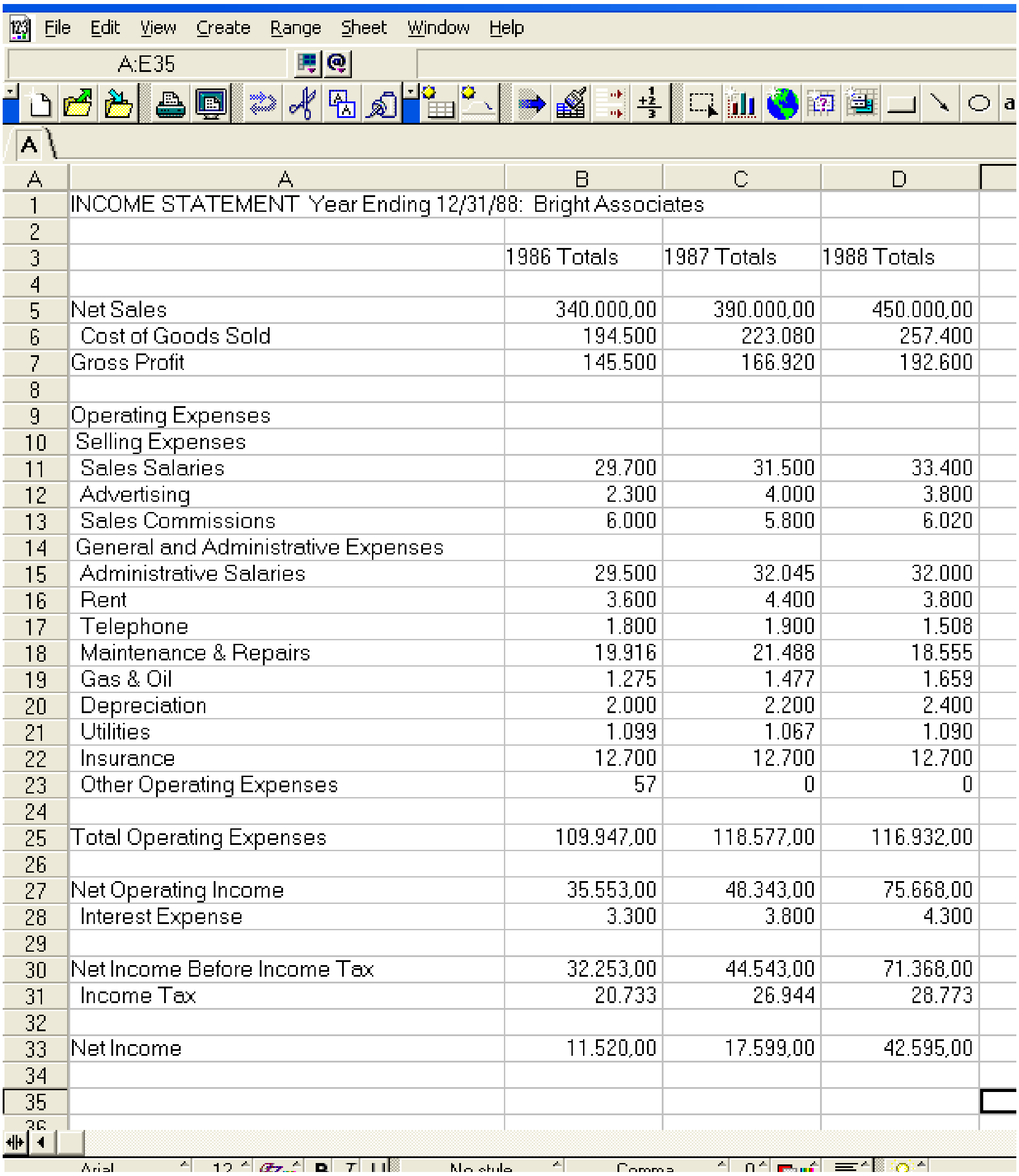


A	A	B	C	D	E	F	G
1	INCOME STATEMENT 1989: Goodwin's Sports Supply (London)						
2							
3		Q1	Q2	Q3	Q4	YTD	
4							
5	Net Sales	£21.000,00	£26.600,00	£22.400,00	£29.600,00	£99.600,00	
6							
7							
8	Operating Expenses:						
9	Payroll	£4.200,00	£5.320,00	£4.480,00	£7.140,00	£21.140,00	
10	Utilities	£3.150,00	£3.990,00	£3.360,00	£5.355,00	£15.855,00	
11	Rent	£1.400,00	£1.400,00	£1.610,00	£1.610,00	£6.020,00	
12	Ads	£1.680,00	£2.128,00	£1.792,00	£2.856,00	£8.456,00	
13	COG Sold	£7.350,00	£9.310,00	£7.840,00	£12.495,00	£36.995,00	
14							
15	Tot Op Exp	£17.780,00	£22.148,00	£19.082,00	£29.456,00	£88.466,00	
16							
17	Op Income	£3.220,00	£4.452,00	£3.318,00	£144,00	£11.134,00	
18							
19							


Microsoft Excel - LONDON
Archivo Edición Ver Insertar Formato Herramientas Datos Ventana ?
10
N
100%
G2
=

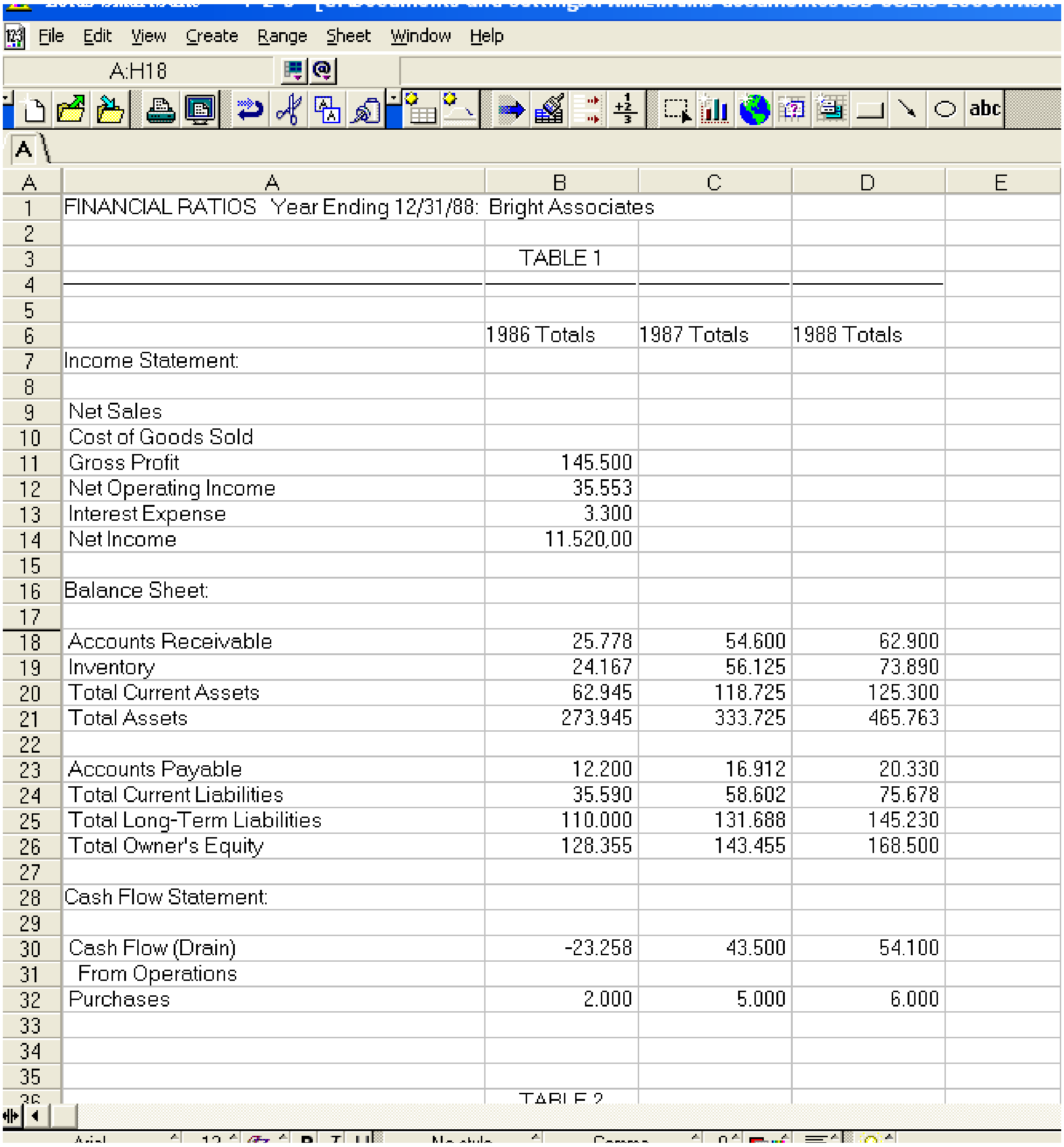
	A	B	C	D	E	F
1	INCOME STATEMENT 1989: Goodwin's Sports Supply (London)					
2						
3		Q1	Q2	Q3	Q4	YTD
4	------------	------------	------------	------------	------------	------------
5	Net Sales	£21.000,00	£26.600,00	£22.400,00	£29.600,00	£99.600,00
6	------------	------------	------------	------------	------------	------------
7						
8	Operating Expenses:					
9	Payroll	£4.200,00	£5.320,00	£4.480,00	£7.140,00	£21.140,00
10	Utilities	£3.150,00	£3.990,00	£3.360,00	£5.355,00	£15.855,00
11	Rent	£1.400,00	£1.400,00	£1.610,00	£1.610,00	£6.020,00
12	Ads	£1.680,00	£2.128,00	£1.792,00	£2.856,00	£8.456,00
13	COG Sold	£7.350,00	£9.310,00	£7.840,00	£12.495,00	£36.995,00
14	------------	------------	------------	------------	------------	------------
15	Tot Op Exp	£17.780,00	£22.148,00	£19.082,00	£29.456,00	£88.466,00
16	------------	------------	------------	------------	------------	------------
17	Op Income	£3.220,00	£4.452,00	£3.318,00	£144,00	£11.134,00
18						

Also, data from **QSSOURCE.WK1** were

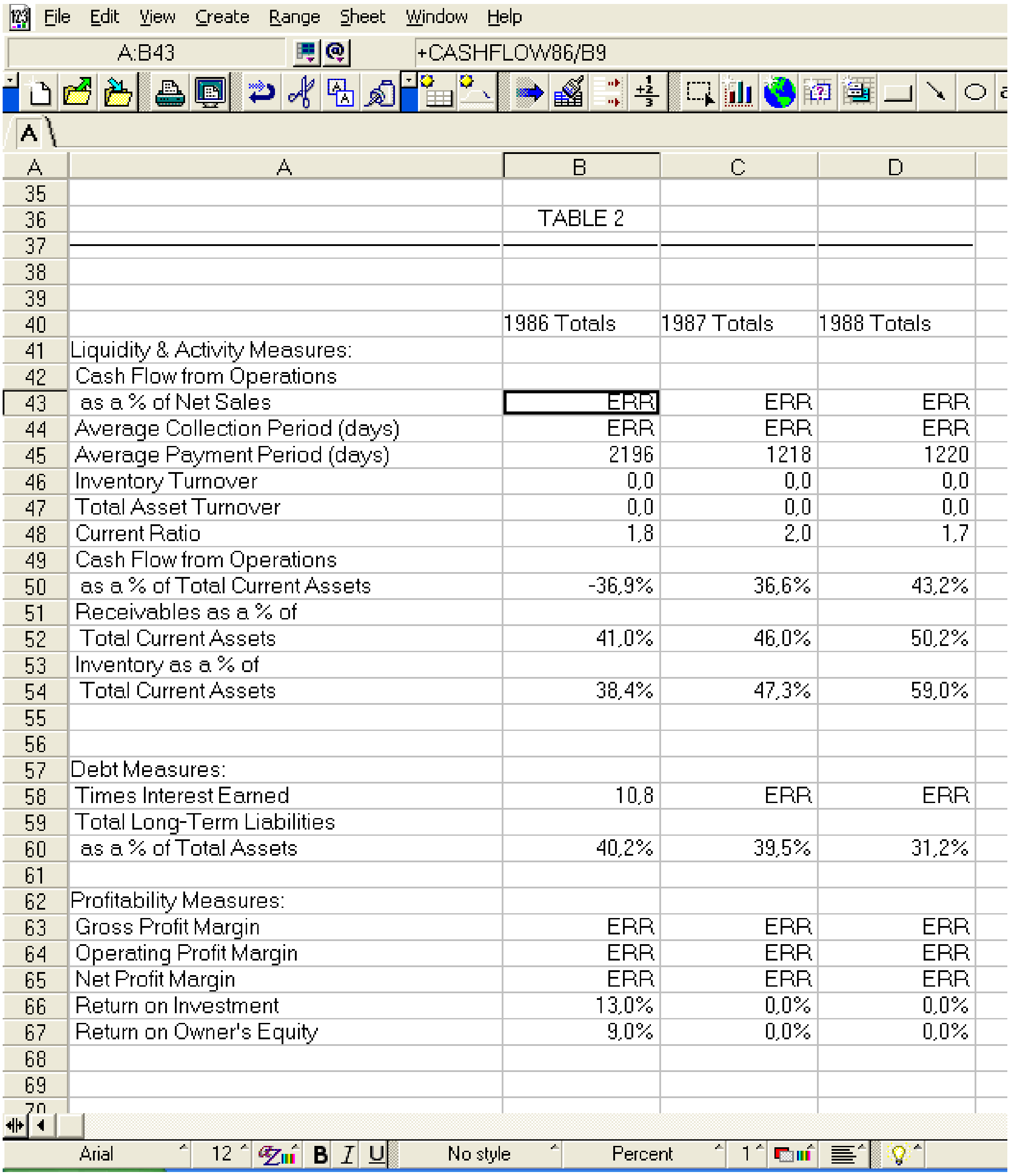


	A	B	C	D
1	INCOME STATEMENT Year Ending 12/31/88: Bright Associates			
2				
3		1986 Totals	1987 Totals	1988 Totals
4				
5	Net Sales	340.000,00	390.000,00	450.000,00
6	Cost of Goods Sold	194.500	223.080	257.400
7	Gross Profit	145.500	166.920	192.600
8				
9	Operating Expenses			
10	Selling Expenses			
11	Sales Salaries	29.700	31.500	33.400
12	Advertising	2.300	4.000	3.800
13	Sales Commissions	6.000	5.800	6.020
14	General and Administrative Expenses			
15	Administrative Salaries	29.500	32.045	32.000
16	Rent	3.600	4.400	3.800
17	Telephone	1.800	1.900	1.508
18	Maintenance & Repairs	19.916	21.488	18.555
19	Gas & Oil	1.275	1.477	1.659
20	Depreciation	2.000	2.200	2.400
21	Utilities	1.099	1.067	1.090
22	Insurance	12.700	12.700	12.700
23	Other Operating Expenses	57	0	0
24				
25	Total Operating Expenses	109.947,00	118.577,00	116.932,00
26				
27	Net Operating Income	35.553,00	48.343,00	75.668,00
28	Interest Expense	3.300	3.800	4.300
29				
30	Net Income Before Income Tax	32.253,00	44.543,00	71.368,00
31	Income Tax	20.733	26.944	28.773
32				
33	Net Income	11.520,00	17.599,00	42.595,00
34				
35				

The data from **QSSTARGET.WK1** was:



A	A	B	C	D	E
1	FINANCIAL RATIOS Year Ending 12/31/88: Bright Associates				
2					
3		TABLE 1			
4					
5					
6		1986 Totals	1987 Totals	1988 Totals	
7	Income Statement:				
8					
9	Net Sales				
10	Cost of Goods Sold				
11	Gross Profit	145.500			
12	Net Operating Income	35.553			
13	Interest Expense	3.300			
14	Net Income	11.520,00			
15					
16	Balance Sheet:				
17					
18	Accounts Receivable	25.778	54.600	62.900	
19	Inventory	24.167	56.125	73.890	
20	Total Current Assets	62.945	118.725	125.300	
21	Total Assets	273.945	333.725	465.763	
22					
23	Accounts Payable	12.200	16.912	20.330	
24	Total Current Liabilities	35.590	58.602	75.678	
25	Total Long-Term Liabilities	110.000	131.688	145.230	
26	Total Owner's Equity	128.355	143.455	168.500	
27					
28	Cash Flow Statement:				
29					
30	Cash Flow (Drain)	-23.258	43.500	54.100	
31	From Operations				
32	Purchases	2.000	5.000	6.000	
33					
34					
35					
36		TABLE 2			



A:B43 +CASHFLOW86/B9

A	A	B	C	D
35				
36		TABLE 2		
37				
38				
39				
40		1986 Totals	1987 Totals	1988 Totals
41	Liquidity & Activity Measures:			
42	Cash Flow from Operations			
43	as a % of Net Sales	ERR	ERR	ERR
44	Average Collection Period (days)	ERR	ERR	ERR
45	Average Payment Period (days)	2196	1218	1220
46	Inventory Turnover	0,0	0,0	0,0
47	Total Asset Turnover	0,0	0,0	0,0
48	Current Ratio	1,8	2,0	1,7
49	Cash Flow from Operations			
50	as a % of Total Current Assets	-36,9%	36,6%	43,2%
51	Receivables as a % of			
52	Total Current Assets	41,0%	46,0%	50,2%
53	Inventory as a % of			
54	Total Current Assets	38,4%	47,3%	59,0%
55				
56				
57	Debt Measures:			
58	Times Interest Earned	10,8	ERR	ERR
59	Total Long-Term Liabilities			
60	as a % of Total Assets	40,2%	39,5%	31,2%
61				
62	Profitability Measures:			
63	Gross Profit Margin	ERR	ERR	ERR
64	Operating Profit Margin	ERR	ERR	ERR
65	Net Profit Margin	ERR	ERR	ERR
66	Return on Investment	13,0%	0,0%	0,0%
67	Return on Owner's Equity	9,0%	0,0%	0,0%
68				
69				
70				

Arial 12 No style Percent 1

Some links from QSSOURCE.WK1 were associated to the file **QSSTARGET**.KW1. In order to reach the data, **QSSOURCE** should be openned first. In this way the data could be recuperated.

Screenshots associated to **QSSTARGET** are divided in two parts due to it had two tables. They may be seen in this document.

Some data had the message **ERR** This is due because in the upper table, TABLE1 no data were filled in the moment of creating the spreadsheet.

It was also possible to recover those files into an Excel® spreadsheet.

Task 1b - Propose one or more suitable preservation strategies and provide a thorough description that highlights their advantages and disadvantages.

FILE NEWS (text document)

Strategy 1

In case of NEWS files the easiest way was to translate the file through RTF[7]. format because future versions of software support this format.

An automatic software could be programmed in order to import file from other programs and translated it to RTF.

From Lotus 1-2-3 AmiPro it was possible to save as a RTF format. After this file is possible to recuperated it on Microsoft Word®.

The **Rich Text Format** (often abbreviated **RTF**) is a free document file format developed by Microsoft in 1987 for cross-platform document interchange. Most word processors are able to read and write RTF documents.

Most documents can be read from RTF format but it is not know if this will be valid on the future. Actually is documented by Microsoft®[8].

Advantages of the strategy

-Any document from 8-bit format on RTF format can actually be read by most of programs. Actual version of Microsoft Office allow to read these files.
-Most text editor programs running under WINDOWS, MAC o LINUX between other platforms may read those file with independency of the platform.
-RTF is a human readable file from any text editor whether it supports RTF formats.
-It is not necessary to produce any change on the original document, reducing cost preservation. It is a function that has the same program.

Disadvantages of the strategy

-RTF is a propietary format and it is not known whether this format can be read in the future and mantained by Microsoft.
-It would not be possible to automatize that all documents to preserve can be converted to RTF format. Some manual intervention could be needed increasing operation cost and effort. Some ancient word processors did not stand RTF format.

Strategy 2

In case of the document called NEWS could be migrated to a XML marked-up or its variant RDF[9] or OWL[10] structure in order to give the document the adequate properties such as data format as bold, quotation marks, citations, tables, and being identified for future software and represented on the screen as it is was done on the original form.

An automatic software could be programmed in order to import file from other programs and translated it to XML, OWL o RDF format.

RDF and OWL allow to give properties to any document through element properties. With the adequate software the document may be composed and visualized again.

Advantages of the strategy

-A OWL or RDF file allow to give elements properties of any element, so it seems to be easy to recover the text and the properties of the text and represent it on a screen.

- XML seems to be an standard readable from almost any software from any platform.

[7] http://en.wikipedia.org/wiki/Rich_Text_Format
[8] http://msdn.microsoft.com/en-us/library/aa140277.aspx
[9] http://www.w3.org/RDF/
[10] http://www.w3.org/2004/OWL/

- XML is not a propietary language and it will be always possible migrate and XML marked-up file to a new XML marked-up language.

-An automatic sofware in mass handling will probably enhance the cost/effort converting to XML format than not RTF format, because reading files from ancient word processors the output on XML format would always be the same and probably it would not need manual intervention.

Disadvantges of the strategy

-Defining properties on a XML document to represented properties of a document could not easy whether this strategy is applied in mass handling.
The reason maybe produced because structure of different word processors files are quite different and it may be would not be possible to have all the information about the marks on the original file depending on the word processor.

-There should be an agreement about the composition of any file in XML identifying marks of the original file and where the file comes from and how to be retrieved later.

-Depending on the word processor some properties on the XML file should be redefined manually

-It is probably also to have manual intervention depending on the file conversion. This refers to ancient 8-bit text editor programs.

-XML is not easy readable file and a presentation layer is needed as part of the software.

FILES LONDON, QSSOURCE and QSSTARGET (spreadsheets)

In order to follow a preservation strategy on spreadsheets LONDON QSSTARGET and QSSOURCE, some steps must have been counted on:

Formulae, cells format (fonts, character attributes and cell appearance[11]), macros and links to ther pages.

Strategy 1

1.-Screenshots of the whole pages in a image format i.e: PNG[12] (***P****ortable* ***N****etwork* ***G****raphics*) provides a patent-free replacement for GIF and can also replace many common uses of TIFF; including graphics relative to the spreadsheet.
If spreadsheet is to big i.e: 64MB or more it may be convinient to make an screenvideo i.e: MPEG-2 or JPEG2000[13] of the whole spreadsheet.

2.-Migrate files to a new standard platform (like XML Math[14]) who stands at least the formula and links. If document have graphics, may be recovered or reconstructed with the whole data recovered.
Actually XML- math may support formulas from spreadsheets.

3.-If spreadsheet has macros in a XML text maybe conserved.

4.-Build a record description about all the formulas done on the spreadsheet in case it would not possible to recuperate some formulas.

Advantatges of this strategy:

-It is possible in this way to access to all the information, formulas, graphics, links, macros and data format.

-Access to graphics and links avoid lost of information.

-XML seems to be actually a standard in order to export and import files due to the metadata.

Disadvantges of this strategy:

[11] http://ahds.ac.uk/preservation/spreadsheets-preservation-handbook.pdf
[12] http://www.w3.org/Graphics/PNG/
[13] http://www.esds.ac.uk/aandp/create/data.asp
[14] http://nssdc.gsfc.nasa.gov/nssdc_news/june01/xdf.html

-Graphics and images should probably be reconstructed.

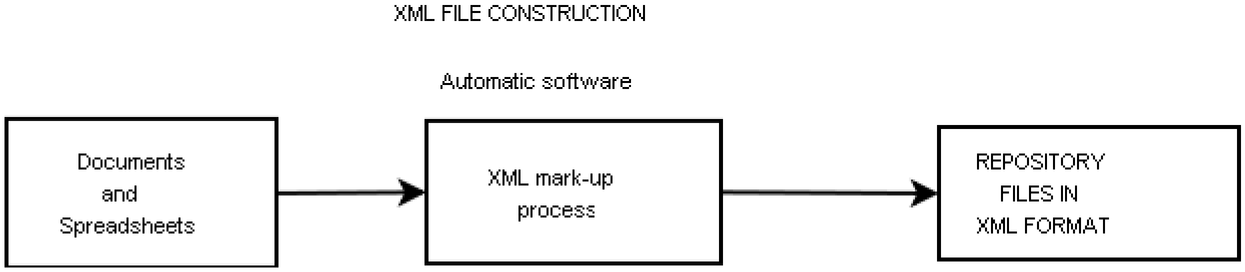
-If spreadsheet files are to big i.e: 64 MB it will be lots of screenshots and it may be difficult to have all the spreasheet on screenshot that should be also preserved later.

-Having screenshots will force later to preserve not only the spreadsheet information. It will also force to preserve screenshot images or the screenvideo on on MPEG2 or JPEG2000 format.

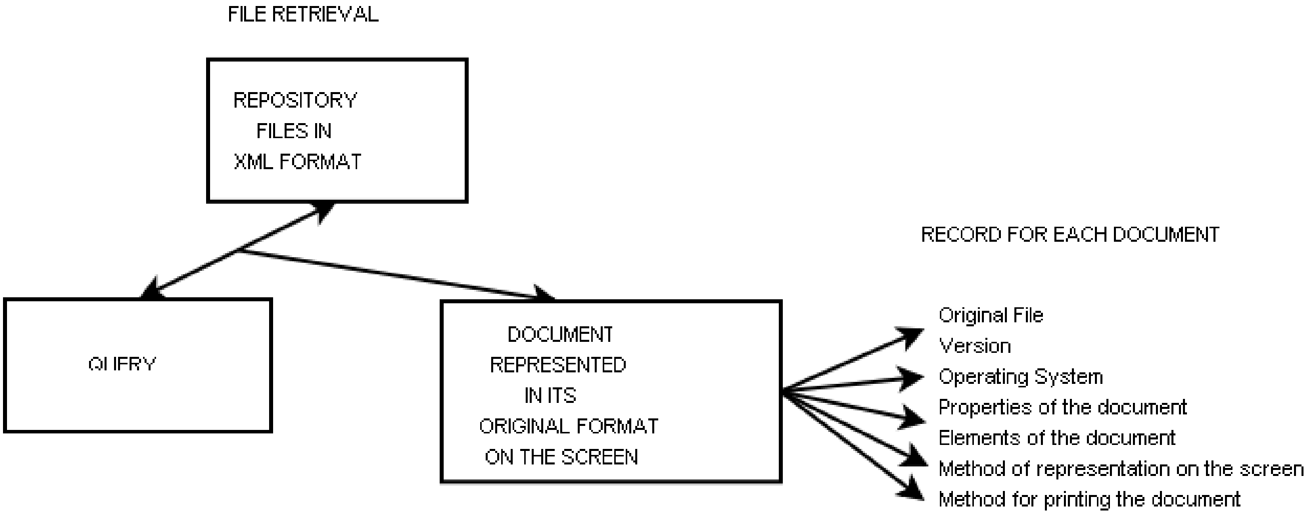
TASK 1c - Implement a preservation strategy capable of mass handling of files of this type, giving an estimate of the cost/effort for deploying the strategies.

Preservation strategy capable of mass handling could be cheap if conversion from original documents is simple. On the other hand if manual intervention is needed could be very expensive.
To work on a very few formats in a log-term, but with metadata inside is desirable. This metadata could serve to say how to reconstruct the file. This is the reason of XML proposal.
If document had images they should be treated by a software and documented also by the program with no manual intervention.

Processes to the strategy:



An automatic sofware could read files. For each file the would be a XML mark-up process upon the characteristics on the document. An spreadsheet has different format than a text document. After this XML-mark-up process, this file will feed a repository on a XML format. For each file the will be some records as Original File, Version of the software, Operating System where it run, Properties of the document, Elements and any descritpion to reconstruct correctly the file onto the screen and in a paper. A document that could be printed should also be printable.


FILE RETRIEVAL
REPOSITORY FILES IN XML FORMAT
QUERY
DOCUMENT REPRESENTED IN ITS ORIGINAL FORMAT ON THE SCREEN
RECORD FOR EACH DOCUMENT
Original File
Version
Operating System
Properties of the document
Elements of the document
Method of representation on the screen
Method for printing the document

To reconstruct file on the screen some Query may be done to the repository. The response to the query will be the document on to the screen but, according to the metadata of the file, file is reconstructed onto the screen.

The big cost/effort on the process would be to build records for each document and tag them it. A big art of analysis for each document should be done first. Also elaborate the software to reconstruct the document will not be ans easy task. Formulas, graphics, links or format document will be also a considerable effort.

TASK 1d - Apply the preservation strategy to the objects and display the files.

According to the former strategy a part of file LONDON has been trasnformed to XML as a sample. The same process would be done with the rest of the files NEWS.SAM, QSSTARGET.WK1, QSSOURCE.WK1

```
<?xml version="1.0" encoding="windows-1252" standalone="yes"?>
<Originalfile>london.wk1</Originafile>
<Version>Lotus 1-2-3</Version>
<Operatingsystem>Window 3.1</Operatingsystem>

.
                      .OTHER TAGS TO RECONSTRUCT THE FILE
.
.
<Rows>
  <Row>
    <Columns
      A="INCOME STATEMENT 1989: Goodwin&apos;s Sports Supply (London)"
    />
  </Row>
  <Row>
    <Columns
      B="Q1"
      C="Q2"
      D="Q3"
      E="Q4"
      F="YTD"
    />
  </Row>
  <Row>
    <Columns
      A="-"
      B="-"
      C="-"
      D="-"
      E="-"
      F="-"
    />
  </Row>
  <Row>
    <Columns
      A="Net Sales"
      B="21000"
      C="26600"
      D="22400"
      E="29600"
      F="99600"
    />
  </Row>
  <Row>
    <Columns
      A="-"
      B="-"
      C="-"
      D="-"
      E="-"
      F="-"
    />
  </Row>
  <Row>
```

```
    <Columns
      A="Operating Expenses:"
    />
  </Row>
  <Row>
    <Columns
      A="  Payroll"
      B="4200"
      C="5320"
      D="4480"
      E="7140"
      F="21140"
    />
  </Row>
  <Row>
    <Columns
      A="  Utilities"
      B="3150"
      C="3990"
      D="3360"
      E="5355"
      F="15855"
    />
  </Row>
  <Row>
    <Columns
      A="  Rent"
      B="1400"
      C="1400"
      D="1610"
      E="1610"
      F="6020"
    />
  </Row>
  <Row>
    <Columns
      A="  Ads"
      B="1680"
      C="2128"
      D="1792"
      E="2856"
      F="8456"
    />
  </Row>
</Rows>
```

TASK 1e - Analyse the benefits and the shortcomings of the preservation strategy.

Benefits

Benefits of this strategy are the facility to retrieve the data without lossing integrity and formats. But this facility should be evaluated and reviewed in mid term due to the technologies and platform's changes.

Normally migration from platforms is a very complex task. Automatic programs for this strategy should do big efforts on translating files to XML standards. Lots of properties of the files have not easy conversion to XML.

From now to the future once the files are on XML standard transformed, it will be easy to migrate to an enhanced format in order to preserve information and data.

Shortcomings

Shortcomings of this strategy is that according XML standard a presentationlayer should be constructed to each file.

Some intregrity of data maybe loose if graphics are not saved before and documented.

To each file different entries for the XML file should be written because each document has different functions than the rest.

If XML transformation is not well done lots of information would be lost. It will be very important the validity of the data.

In order to retrieve documents from the XML repository, it is also important that document would be well tagged in order to avoid documental noise.

SCENARIO 2 - Images from a Legacy Computer Gaming Platform

An image archive received a donation from an artist representing working material from his early years. While ingesting the data into the image archive repository, the system failed to identify some of the file formats. The artists cannot remember the name of the particular application or the computer platform. He also found a related file for one type of the images, but he does not know what the file is. Can you display the images? Include the images in your report in an appropriate form.

Task
Your task is to:

1. Identify the application and computer platform that the files come from.
2. Display the images in an appropriate form.
3. Propose valid preservation alternatives pointing out their advantages and disadvantages.
4. Implement one or more of the preservation strategies (for example emulation or migration).
5. Evaluate and compare their performance.

TASK 2a - Identification of the computer platform

After revewing the files, there are two folders.

Folder1 contained three files **SKY**, **GRID** and XLART.XFD.
Folder2 contained two files **EINSTEIN** and **CAR** with no extension

FOLDER1
Information about XFD extension[15] was found. It was related[16] and used by ATARI XL/XE emulators[17].
The file XLART.XFD, corresponds to an ATARI, XL/XE 8-bit platform.

What's ATARI?
ATARI was a company founded in 1972 as Atari INC and its industry is consumers electronics and video console games. Actually Atari is Atari Inc., formerly Inforgrames Inc./GT Interactive.[18]

What's an emulator?
An emulator is software that allows imitate the behaviour of a computer software outdated by the technology into other computer with a diferent platform.

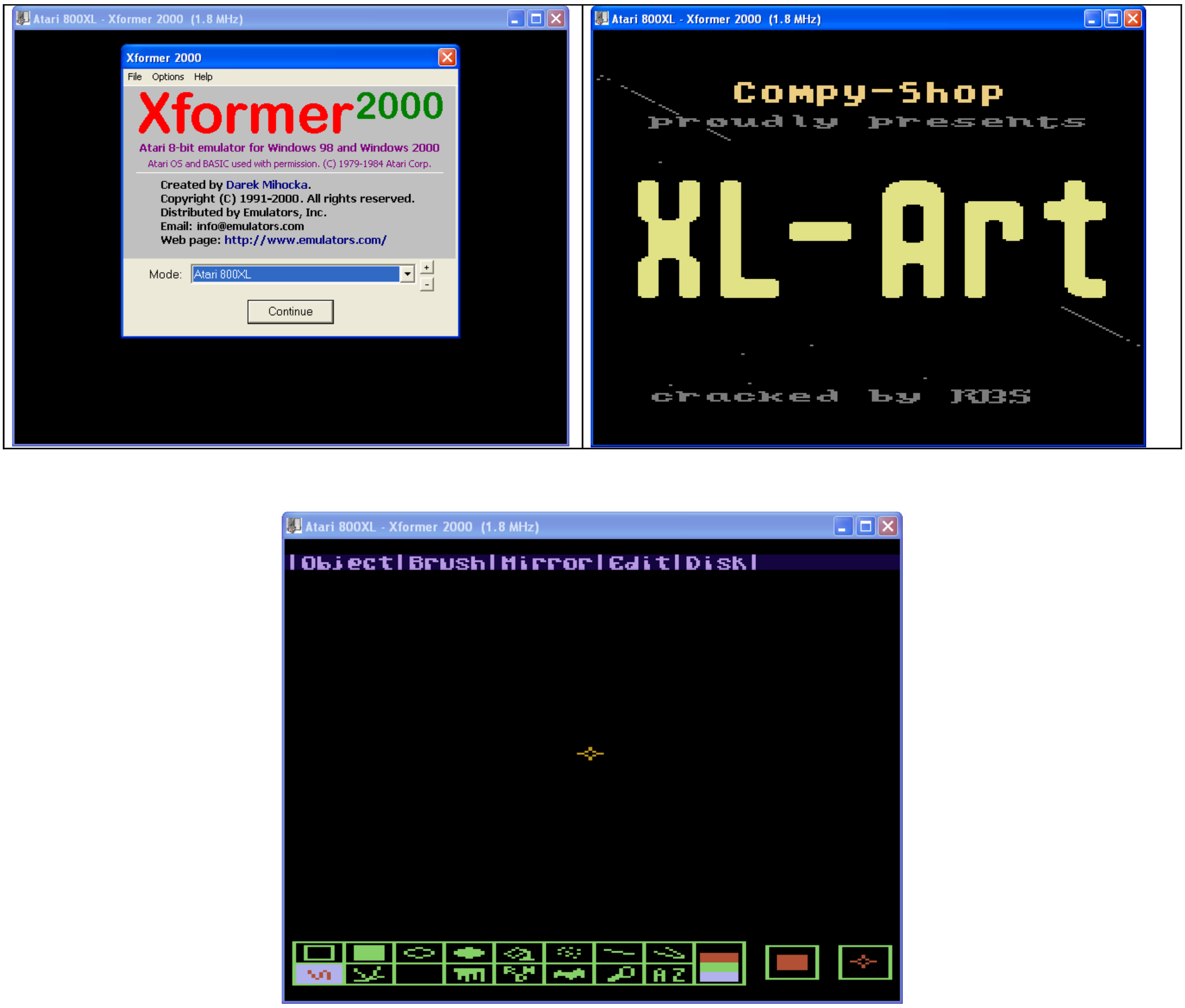
In order to run those files it is necessary to introduce about these kind of files and emulator's concept.

In case of ATARI XL/XE it is possible to find some emulators. Those emulators are:

XFORMER2000 from Emulators Inc

"Xformer 2000 is the Atari 8-bit emulator for Windows, made by the same people that brought you the ST Xformer Atari 8-bit emulator for GEM, and the PC Xformer Atari 8-bit emulator for MS-DOS. Gone is the command line of the MS-DOS based PC Xformer. "[19].

[15] http://filext.com/file-extension/XFD
[16] http://www.atarimax.com/ape/docs/DiskImageFAQ/
[17] http://www.emulators.com/
[18] http://en.wikipedia.org/wiki/ATARI
[19] http://www.emulators.com/



Atari 800XL - Xformer 2000 (1.8 MHz)

Xformer 2000

File Options Help

Xformer2000

Atari 8-bit emulator for Windows 98 and Windows 2000

Atari OS and BASIC used with permission. (C) 1979-1984 Atari Corp.

Created by Darek Mihocka.
Copyright (C) 1991-2000. All rights reserved.
Distributed by Emulators, Inc.
Email: info@emulators.com
Web page: http://www.emulators.com/

Mode: Atari 800XL

Continue

Atari 800XL - Xformer 2000 (1.8 MHz)
Compy-Shop
proudly presents
XL-Art
cracked by KBS
Atari 800XL - Xformer 2000 (1.8 MHz)
|Object|Brush|Mirror|Edit|Disk|
A Z

Screenshots from the XFORMER2000 emulator and run of the XLART.XFD on the emulator

ATARI800WIN PLUS 4.0 Beta 7 [20].

This emulator is a free software; under the terms of the GNU General Public and done by
Atari800 © 1995-1998 David Firth
Atari800 © 1998-2005 Atari800 development team
Atari800Win © 1998-2000 Richard Lawrence
Atari800Win PLus © 2000-2003 Tomasz Szymankowski
Atari800Win PLus © 2004-2005 Marcin Lewandowski
R: handler over TCP/IP © 2000 Daniel Noguerol
Help file, some routines © 2000-2003 Piotr Fusik
libpng 1.2.8 © 1998-2004 Glenn Randers-Pehrson
zlib 1.2.3 © 1995-2005 Jean-loup Gailly and Mark Adler

There are four machine type available:
400/800 OS-A, 400/800 OS-B, XL/XE and 5200.

400/800 was first family of the Atari home computers. There were two versions of the operating system for 400/800 – rev. A and rev. B. The latter one fixes many bugs in the former, however some old software requires rev. A.

Newer Atari home computers series, XL and XE, had more memory, better operating system (16 KB) and a built-in BASIC language interpreter (8 KB).

[20] http://atari800.sourceforge.net/download.html

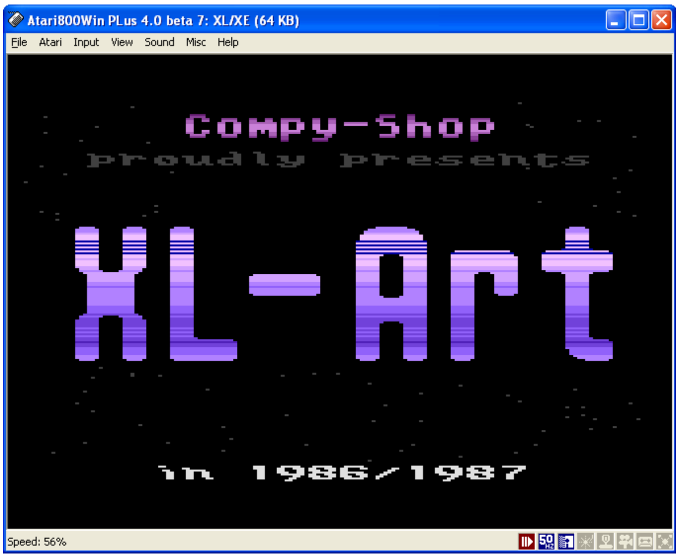
Atari 400, 800, 600 XL, 800 XL, 130 XE and 5200. 130 XE compatible memory expansions: 320 KB, 576 KB and 1088 KB.

Emulation of a cassette player. Supported are Cassette images (CAS) and raw files. The cassette emulation uses the SIO patch, so you won't wait a dozen or so minutes for a program to load.

Emulation of 4 virtual hard disks. See H: device handler (virtual hard disks).

Printer emulation. See P: device handler (printer).

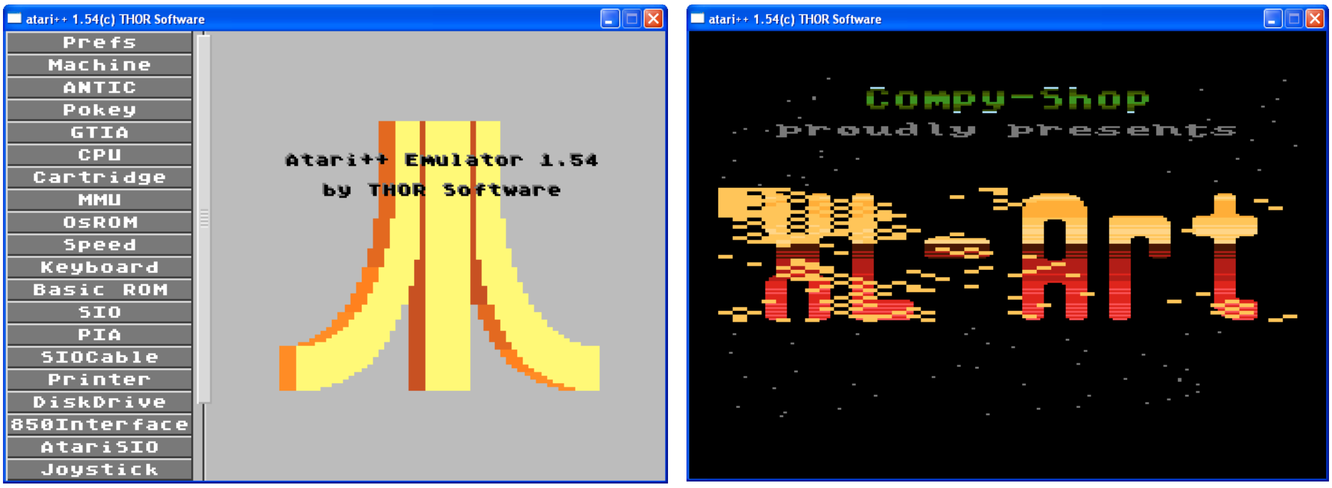
Stereo (two POKEYs) emulation.


Atari800Win PLus 4.0 beta 7: XL/XE (64 KB)
File Atari Input View Sound Misc Help
Compy-Shop
proudly presents
XL-Art
in 1986/1987
Speed: 56%

Screenshot from the ATARI800WIN PLUS emulator and the execution of the XLART.XFD on the emulator

ATARI++ emulator[21].

The Atari++ Emulator is a Unix based emulator of the Atari 8 bit computers, namely the Atari 400 and 800, the Atari 400XL, 800XL and 130XE, and the Atari 5200 game console. The emulator is auto-configurable and will compile on a variety of systems (Linux, Solaris, Irix)
atari++ is an emulator for (now rather aged) Atari 8 bit computers. It emulates the Atari800, Atari400, the 800XL and 600XL, the 65XE and 130XE and the Atari 5200 Game Console. The emulation is cycle-precise, that is "on the fly" modifications of chip registers will be visible on the screen immediately, emulating even programs using horizontal kernel tricks correctly.



Screenshots from the ATARI++ emulator and the execution of the XLART.XFD on the emulator.

[21] http://www.math.tu-berlin.de/~thor/atari++/

Files supported:
All these emulators support XFD and ART files. Those kind of files are

XFD

The simplest format of an Atari disk image is XFD (Xformer disk, Xformer is another Atari emulator). In this format the file contains only raw dumps of all sectors, thus in some cases the structure of the disk may be not recognized correctly. Therefore this format is not recommended.

XFD-Image format invented by Emulators Inc, for their ST Xformer emulator. Identical to .ATR except without the 16 byte header.

ATR

A bit more complicated, but much better and more common, is ATR, which additionally has 16-byte header containing some information about disk size, density and write protection.

XFD and ATR files extensions, are ROM files which allow the emulator to run the software. It is possible to convert one format to another and vicecersa through the ARDF program[22]. It is also possible to design ROM files in ATR or in XFD formats.

Task 2B - Display the images in an appropriate form.

The images who came with file with were relatives to the drawing software XL-ART. SKY and GRID

Those images were with PIC extension and belonged to the ROM file XLART.XFD.

This images could be displayed through the program Graph2Font[23].

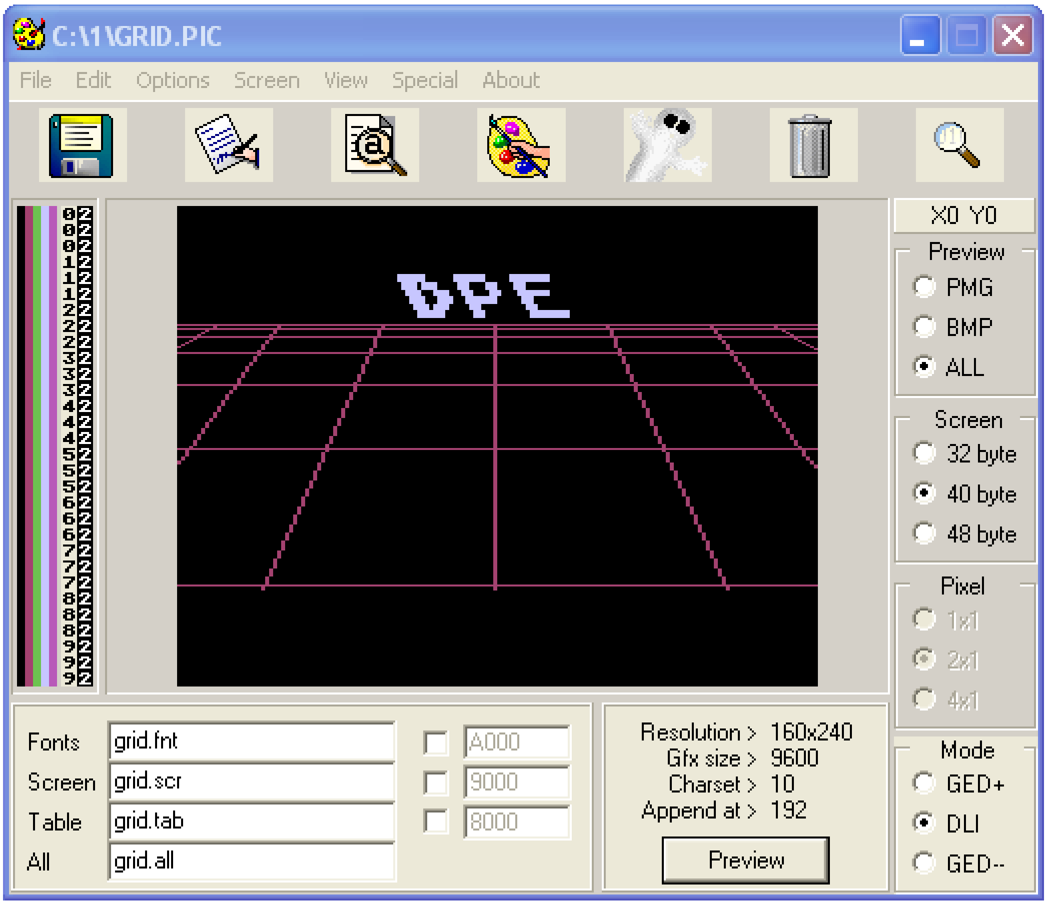
The Graph2Font (G2F) program treats the image as fields with dimensions of 4x8, 2x8, 8x8 pixels. These fields are charset(s) (fonts) and if only some field is repeated then it's exchanged with adequate sign (font). This is some kind of loseless image compression. At the beginning this program had been used only to convert graphics to chars but now it also allows you to do multicolor graphics for the 8-bit Atari.

On the images belows all images are show on the program Graphfont2.

.

[22] http://pvb.free.fr/Atari/index.php

[23] http://g2f.atari8.info/



C:\1\GRID.PIC

X0 Y0

Preview: ○ PMG ○ BMP ● ALL

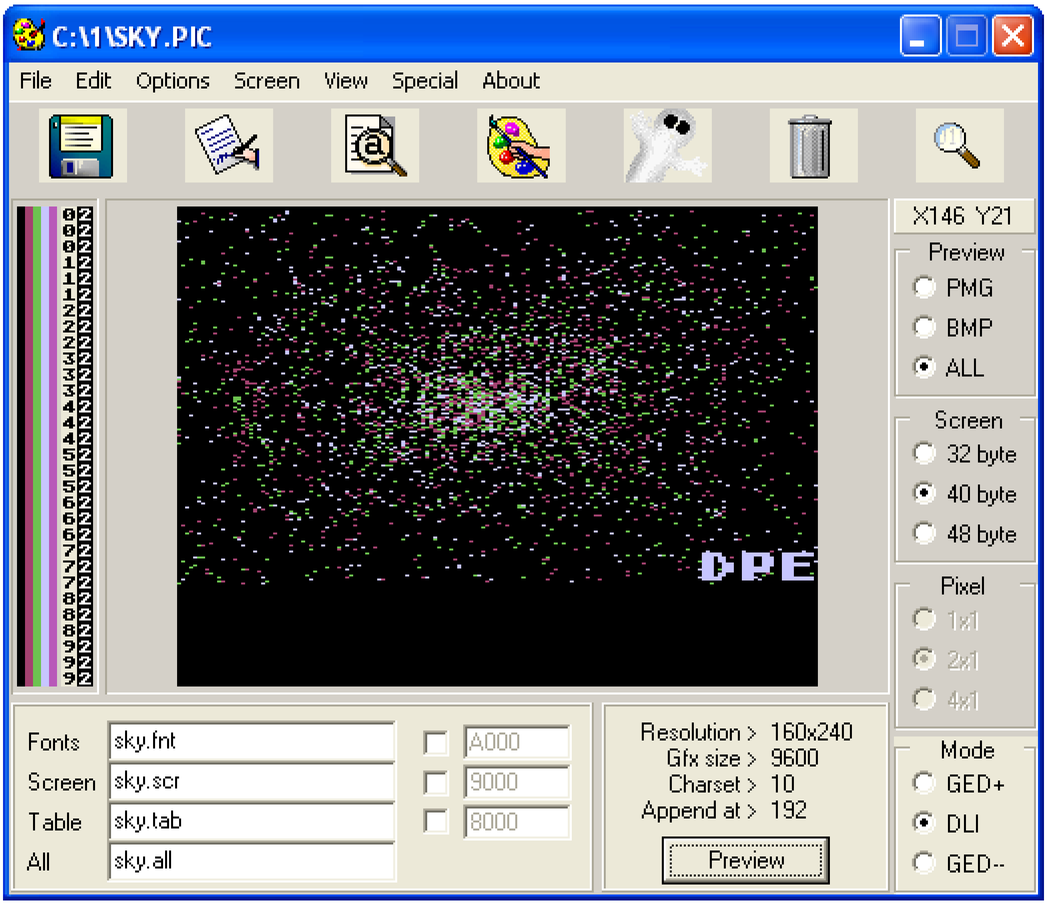
Screen: ○ 32 byte ● 40 byte ○ 48 byte

Pixel: ○ 1x1 ● 2x1 ○ 4x1

Mode: ○ GED+ ● DLI ○ GED--

Fonts grid.fnt □ A000
Screen grid.scr □ 9000
Table grid.tab □ 8000
All grid.all

Resolution > 160x240
Gfx size > 9600
Charset > 10
Append at > 192



C:\1\SKY.PIC

X146 Y21

Preview: ○ PMG ○ BMP ● ALL

Screen: ○ 32 byte ● 40 byte ○ 48 byte

Pixel: ○ 1x1 ● 2x1 ○ 4x1

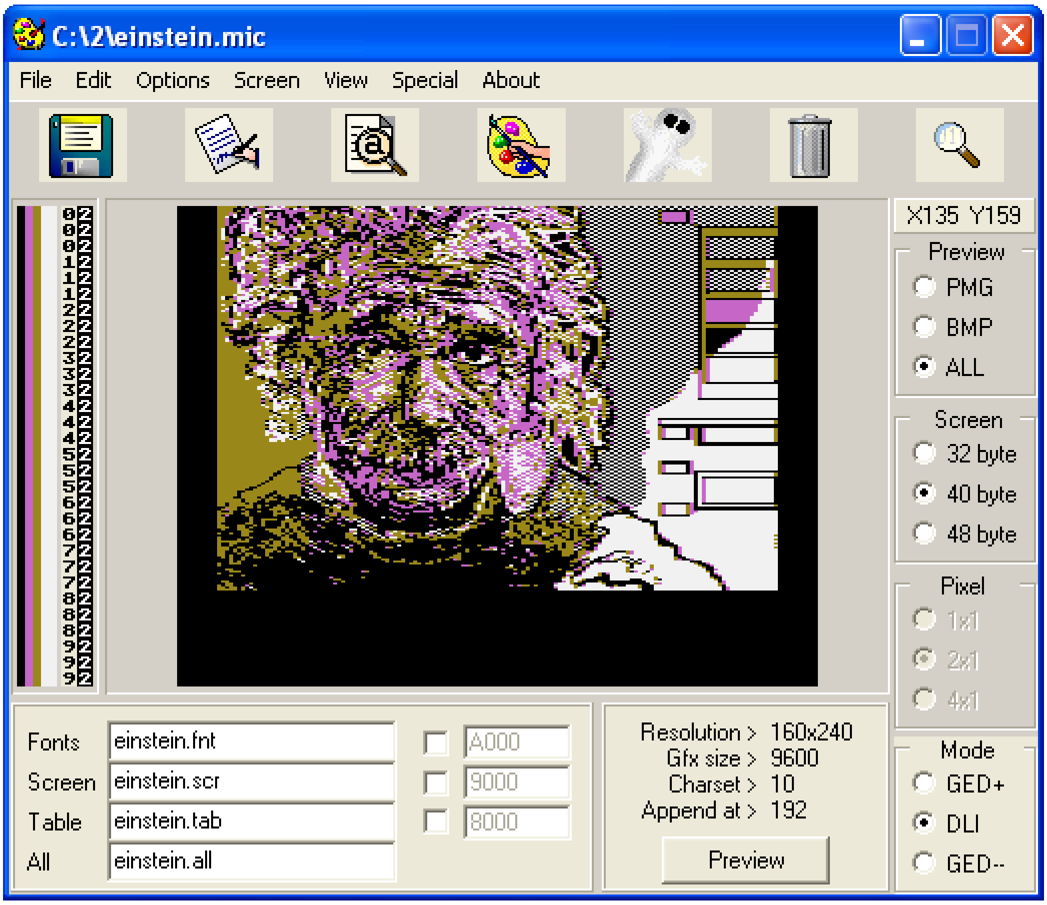
Mode: ○ GED+ ● DLI ○ GED--

Fonts sky.fnt □ A000
Screen sky.scr □ 9000
Table sky.tab □ 8000
All sky.all

Resolution > 160x240
Gfx size > 9600
Charset > 10
Append at > 192

FOLDER2

The other images placed in folder 2 were images in **mic** format. .mic - Microsoft Image Composer data file extension Those images were also opened with Graphfont2.



C:\2\einstein.mic

X135 Y159

Preview: ○ PMG ○ BMP ● ALL

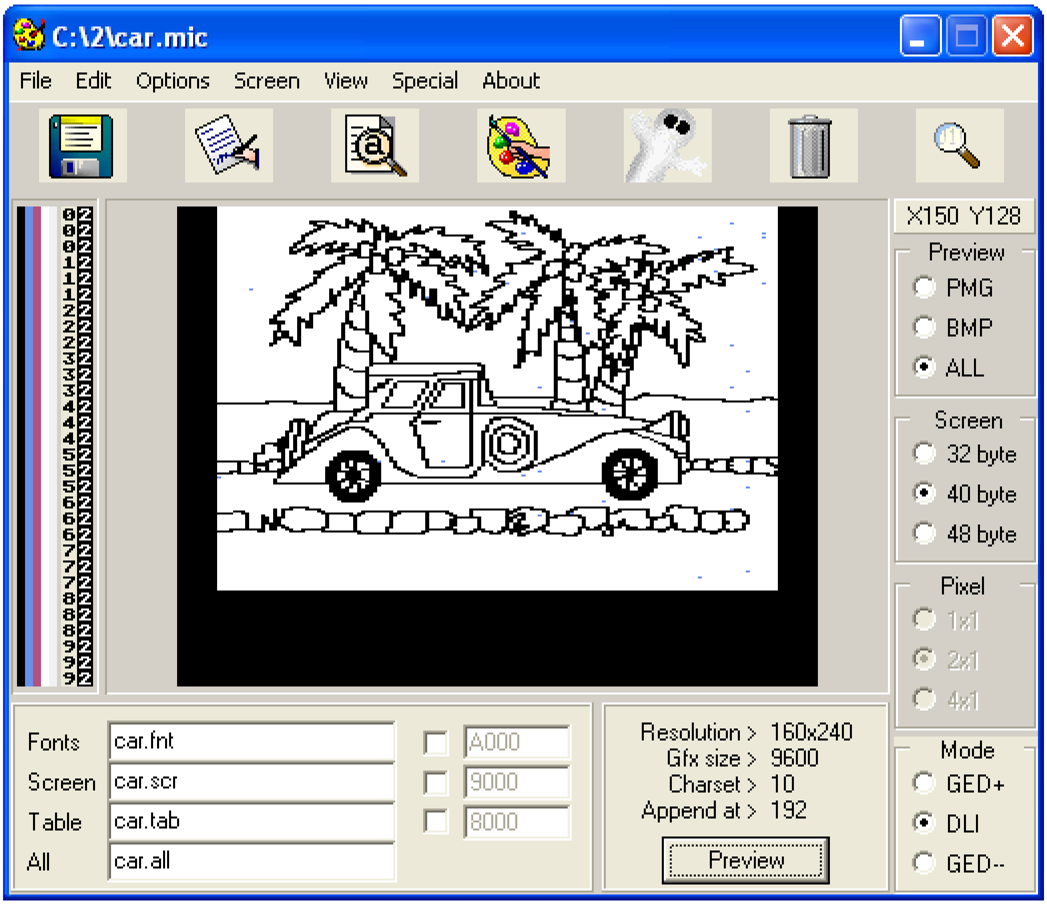
Screen: ○ 32 byte ● 40 byte ○ 48 byte

Pixel: ○ 1x1 ● 2x1 ○ 4x1

Mode: ○ GED+ ● DLI ○ GED--

Fonts	einstein.fnt	□	A000
Screen	einstein.scr	□	9000
Table	einstein.tab	□	8000
All	einstein.all		

Resolution > 160x240
Gfx size > 9600
Charset > 10
Append at > 192



C:\2\car.mic

X150 Y128

Preview: ○ PMG ○ BMP ● ALL

Screen: ○ 32 byte ● 40 byte ○ 48 byte

Pixel: ○ 1x1 ● 2x1 ○ 4x1

Mode: ○ GED+ ● DLI ○ GED--

Fonts	car.fnt	□	A000
Screen	car.scr	□	9000
Table	car.tab	□	8000
All	car.all		

Resolution > 160x240
Gfx size > 9600
Charset > 10
Append at > 192

TASK 2c - Propose valid preservation alternatives pointing out their advantages and disadvantages.

Emulation seems to be the best option in this case. This software was designed to run on 8-bit platform and by companies that probably actually most of them do not exist.

Advantages:

Emulation will allow running any software on this platform.

It's not necessary to migrate any sofware, just to ellaborate new one who simulate software with the properties of the old one.

Disadvantges:

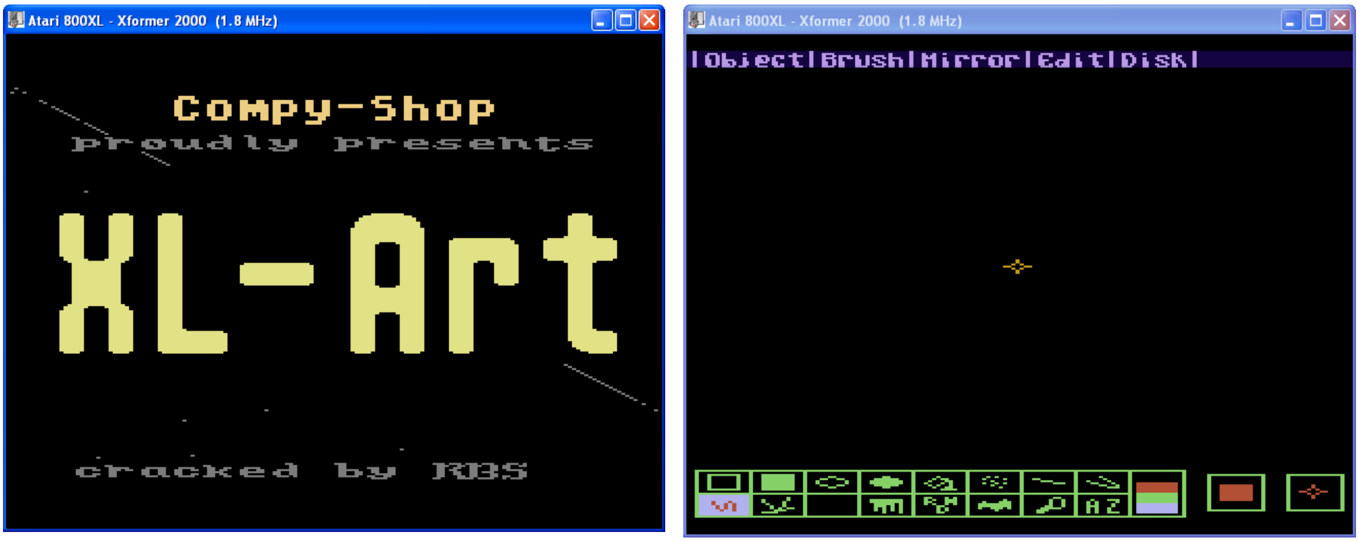
Emulators should be programmed for every platform where the software must run. This means to create emulators for Windows, MAC Os Linux and the knowledge of different platforms and software who must be well documented.

In case of emulations some gadgets as joystisck, keyboard may not work properly.

Adjusting the speed of new CPU's with new ones to run programs properly must be done very acurately.

TASK 2d – Implement one or more of the preservation strategies (for example emulation or migration)

In our case, emulation has been proved in this task. Looking for emulators to run the program and looking for software to wiew the images. As it is said before emulator as XFORMER have been proved.


Atari 800XL - Xformer 2000 (1.8 MHz)
Compy-Shop
proudly presents
XL-Art
cracked by KBS
Atari 800XL - Xformer 2000 (1.8 MHz)
|Object|Brush|Mirror|Edit|Disk|

Screenshots on the emulator XFORMER

TASK 2e – Evaluate and compare their performance

In case of the program performance is mesaured in terms of speed. The program was designed to run under 8-bit platform and it should be like that. Probably, construction of new emulators will supose a big effort in order to update the systems, but it will be a bigger effort to migrate each game o program to a new platform rather than constructing an emlator. IT's not the same in terms of cost to construct o adapt an emulator to a new platform than reconstruct all software preserved to a new platform.

Something to take care is the documentation of the software who has to be done. This means, that for each emulator that has to be migrated, there should exist more acurately documentation to run old software. Otherwise, there's a risk on running emulator and old software and consequently losing of software due to not having documented the emulator's migration.

SCENARIO 3 - Obsolete Database

In the beginning of 2003, the Porto Regional Archive (ADP) initiated a project called DigitArq. The goal of the project was to bring together its various finding aids, previously scattered throughout the archive in many different forms and formats, into a single centralised repository based on international standards such as ISAD(G) and EAD. The planned repository would enable the standardisation of all archival procedures and the development of new data services such as search mechanisms and description tools.

However, in late 2007 a fire broke out in the server room destroying the server that held all the information produced over the last 25 years.

In addition to this all the backup tapes that were kept in a cabinet in the adjoining room were destroyed. Around 80% of the information had been synchronised with a similar repository at the National Archives in Lisbon, and this information was easily recovered.

The other 20% had been migrated from an old database that was still kept at the archive but had not been used since 1990. The ADP staff were unable to use the database, so they decided to hire a digital preservation expert to do the job.

Task

a) Identify the database system that is necessary to interpret the provided data files.
b) Provide access to the data of the database.
c) Propose an adequate procedure to migrate the records from the old system to the new one (please provide output files, scripts or standards to support your answer).
d) Propose a disaster recovery plan so that if a similar incident takes place the disruption will be kept to a minimum.

Task 3a - Identify the database system that is necessary to interpret the provided data files.

Files ares from CD/ISIS database in WINSIS[24].

WINISIS comes from CD/ISIS that ran on an IBM mainframe and was designed in the mid's 1970s under Mr Giampaolo Del Bigio for UNESCO'S Computeried Documentation System. It was based on the internal ISIS (Integrated Set of Information Systems) at the International Labour Office in Geneva.

At the International Conference in Bogota held in 1995 to mark 10 years of CDS/ISIS the first result of the migration to Microsoft Windows® were presented and Abel Packer from the Regional Library of Medicice (BIREME), Sao Paulo, presented CISIS and ISIS_DLL. ISIS_DLL, develped jointly with UNESCO, permits applications written in compatible Windows Language, e.g. Visual Basic, to interact with CDS/ISIS databases.

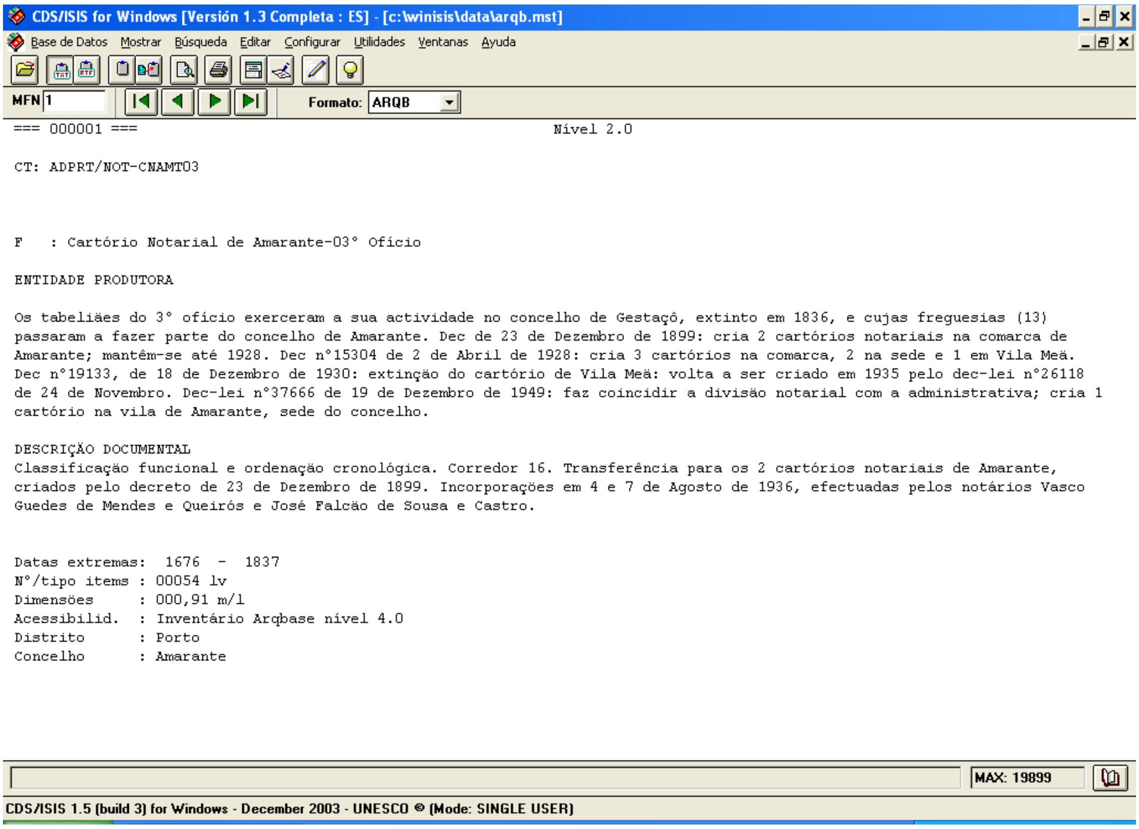
Actually it is possible to get a WINISIS licence free from BIREME[25]. It is also possible through the national distributors. In case of Spain it is possible to ask for a Licence to CINDOC[26], Instituto de Información y Documentación Cientifica Spain, 28002 Madrid

WINSIS is a database system used in library's that retrieve files to text complet is possible. Actually there are some versions that a published catalog online is allowed, like WXIS, or the new system JISIS not yet disposal.

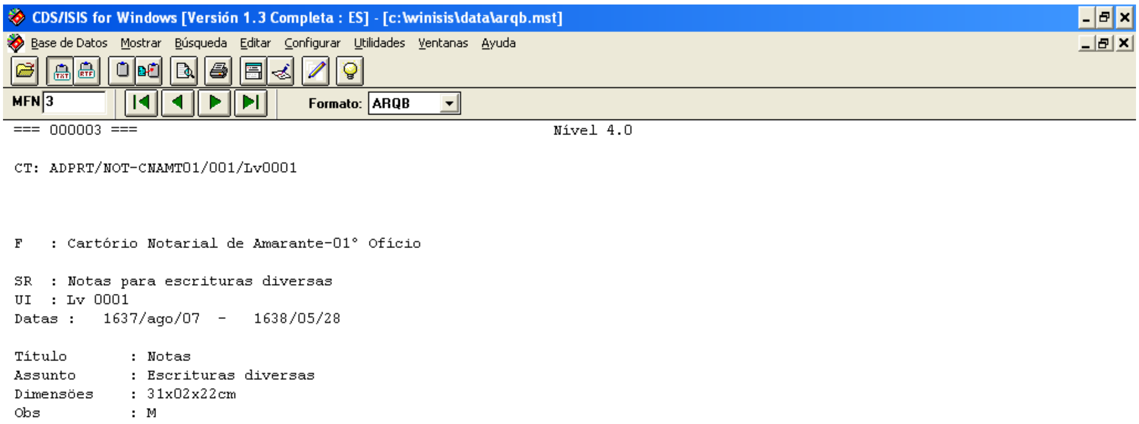
[24] http://www.unesco.org/isis/files/winisis/windows/doc/spanish
[25] http://productos.bvsalud.org/product.php?id=winisis&lang=es
[26] http://www.cindoc.csic.es/

Task 3b - Provide access to the data of the database.

Once identified the files, next step is to visualize the database. Here are some screenshots to the database. Those screenshots are from some records. This database has 19899 records.
Winisis worked initially on Windows 3.1® but it works also on Windows XP Professional®


CDS/ISIS for Windows [Versión 1.3 Completa : ES] - [c:\winisis\data\arqb.mst]
Base de Datos Mostrar Búsqueda Editar Configurar Utilidades Ventanas Ayuda
MFN 1
Formato: ARQB
MAX: 19899
CDS/ISIS 1.5 (build 3) for Windows - December 2003 - UNESCO © (Mode: SINGLE USER)

```
=== 000001 ===                                            Nível 2.0

CT: ADPRT/NOT-CNAMT03



F   : Cartório Notarial de Amarante-03° Ofício

ENTIDADE PRODUTORA

Os tabeliães do 3° ofício exerceram a sua actividade no concelho de Gestaçô, extinto em 1836, e cujas freguesias (13)
passaram a fazer parte do concelho de Amarante. Dec de 23 de Dezembro de 1899: cria 2 cartórios notariais na comarca de
Amarante; mantêm-se até 1928. Dec n°15304 de 2 de Abril de 1928: cria 3 cartórios na comarca, 2 na sede e 1 em Vila Meä.
Dec n°19133, de 18 de Dezembro de 1930: extinção do cartório de Vila Meä: volta a ser criado em 1935 pelo dec-lei n°26118
de 24 de Novembro. Dec-lei n°37666 de 19 de Dezembro de 1949: faz coincidir a divisão notarial com a administrativa; cria 1
cartório na vila de Amarante, sede do concelho.

DESCRIÇÃO DOCUMENTAL
Classificação funcional e ordenação cronológica. Corredor 16. Transferência para os 2 cartórios notariais de Amarante,
criados pelo decreto de 23 de Dezembro de 1899. Incorporações em 4 e 7 de Agosto de 1936, efectuadas pelos notários Vasco
Guedes de Mendes e Queirós e José Falcão de Sousa e Castro.


Datas extremas:  1676  -  1837
N°/tipo items : 00054 lv
Dimensões     : 000,91 m/l
Acessibilid.  : Inventário Arqbase nível 4.0
Distrito      : Porto
Concelho      : Amarante
```


CDS/ISIS for Windows [Versión 1.3 Completa : ES] - [c:\winisis\data\arqb.mst]
Base de Datos Mostrar Búsqueda Editar Configurar Utilidades Ventanas Ayuda
MFN 3
Formato: ARQB

```
=== 000003 ===                                            Nível 4.0

CT: ADPRT/NOT-CNAMT01/001/Lv0001



F   : Cartório Notarial de Amarante-01° Ofício

SR  : Notas para escrituras diversas
UI  : Lv 0001
Datas :    1637/ago/07  -    1638/05/28

Título        : Notas
Assunto       : Escrituras diversas
Dimensões     : 31x02x22cm
Obs           : M
```

Task 3c - Propose an adequate procedure to migrate the records from the old system to the new one (please provide output files, scripts or standards to support your answer).

WINISIS allows to export files to ISO 2709 amb XML. In order to conservate standards I've choosen to export the database to ISO 2709. In this way later with a program that may read ISO 2709 O MARC2 [27]may be read this file.
MARC is an acronym for Machine-Readable Catalogue or Cataloguing. It is not, however, a kind of catalogue nor a method of cataloguing but a system by which data elements within bibliographic records are uniquely labeled for

[27] http://www.loc.gov/marc/

computer handling. MARC is an implementation of the international standard "Information and documentation - Format for information exchange". (ISO 2709-1996)[28].

```
01337000000000169000450030206590000030303030065931000180096230400050098030600050
09853120011009903150013010013140083010140990004010971010008011012010006011093010
05201115-^aOs tabeli„es do 3§ of¡cio exerceram a sua actividade no concelho de <
Gesta‡“>, extinto em 1836, e cujas freguesias (13) passaram a fazer parte do con
celho de Amarante. Dec de 23 de Dezembro de 1899: cria 2 cart¢rios notariais na
comarca de Amarante; mantˆm-se at‚ 1928. Dec n§15304 de 2 de Abril de 1928: cria
 3 cart¢rios na comarca, 2 na sede e 1 em <Vila Me„>. Dec n§19133, de 18 de Deze
mbro de 1930: extin‡„o do cart¢rio de Vila Me„: volta a ser criado em 1935 pelo
dec-lei n§26118 de 24 de Novembro. Dec-lei n§37666 de 19 de Dezembro de 1949: fa
z coincidir a divis„o notarial com a administrativa; cria 1 cart¢rio na vila de
Amarante, sede do concelho.-^aClassifica‡„o funcional e ordena‡„o cronol¢gica. C
orredor 16. Transferˆncia para os 2 cart¢rios notariais de Amarante, criados pel
o decreto de 23 de Dezembro de 1899. Incorpora‡”es em 4 e 7 de Agosto de 1936, e
fectuadas pelos not rios Vasco Guedes de Mendes e Queir¢s e Jos‚ Falc„o de Sousa
 e Castro.-^aPorto^bAmarante-1676-1837-^a00054 lv-^a000,91 m/l-^aInvent rio Arqb
ase n¡vel 4.0^b‹ndice alfab‚tico de not rios. ‹ndice cronol¢gico.-2.0-^aADPRT-^a
NOT-^aCNAMT03^bCart¢rio Notarial de Amarante-03§ Of¡cio-
01433000000000181000450009900040000010100080000420100060001230100520001830207070
00703110037007773030303008143100018011173040005011353060005011403120011011453150
013011563140082011 69-2.0-^aADPRT-^aNOT-^aCNAMT04^bCart¢rio Notarial de Amarante-
04§ Of¡cio-^aOs tabeli„es do 4§ of¡cio exerceram a sua actividade na vila de Ama
rante, constitu¡da por uma freguesia - <S„o Gon‡alo de Amarante> - at‚ 1836, ano
 em que passou a ser sede do concelho (decreto de 6 de Novembro de 1836). Dec de
```

One of the best ways to migrates records from the old seems to be, to translate the database to XML.
There are at least two ways for translating ISIS database to XML. From a freeware application ISIS2XML[29], developed by Mr. Pierre Chabert (France), or with WINSIS utilities XML translation. An XML file is provided.

Task 3d - Propose a disaster recovery plan so that if a similar incident takes place the disruption will be kept to a minimum.

Disaster recovery plan:

When a disaster recovery plan is asked to do it, some questions must be count on.

-If the data are critical
-How many users are going to read this data
-Diferents platforms where data should be read

First of all a backup strategy must be done.

Three backup files should exist.

One of them locally in case of hardware failure
Another copy out from the building in a safety place
The third backup may be place on p2p network archive where other information of the organization is placed.

[28] http://www.ibiblio.org/msmckoy/marc2.html
[29] http://portal.unesco.org/ci/en/ev.php-URL_ID=5335&URL_DO=DO_TOPIC&URL_SECTION=201.html

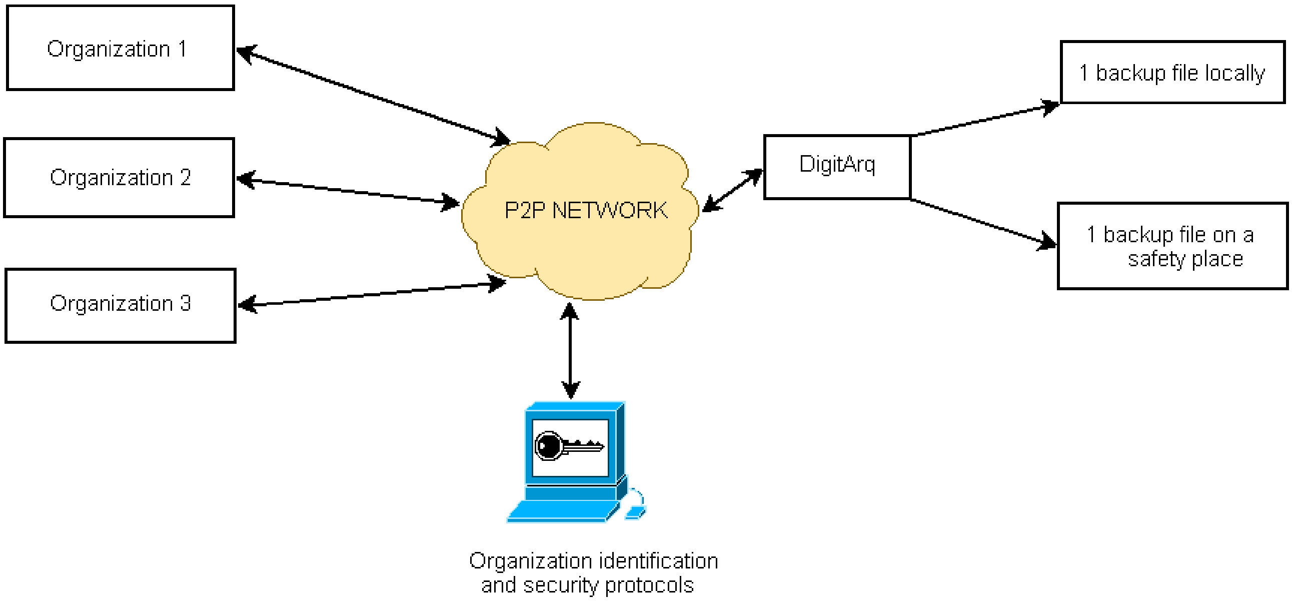
This p2p network archive servers are out also from the building, and cost may be shared by other archival organizations who are also interested on having backup files on a safety place.

In this case, server should have a policy of whichs documents organizations are allowed to retrieve due to share it with other organizations.

In this case the data are distributed in many computers and ther will be always some copies to retreive from de system.

Maybe is a very expensive option but this could be the most useful for achivist organizations who need to share structures and reducing costs operations. On a p2p network archive data are running over the organization.

In this picture, some organizations share a P2P netowrk in order to have backup files, and software to run this backup files.



SCENARIO 4 – Electronic Art

Founded in 1987, the Prix Ars Electronica is an interdisciplinary platform for digital art and media culture. The Prix Ars Electronica is one of the most important awards for creativity and pioneering spirit in the field of digital media. With the rapid change of software tools and frameworks for multimedia authoring their artworks are in danger of becoming inaccessible and unusable. You have been asked to preserve four of these historical digital artworks for future generations and to develop appropriate digital preservation strategies.

Task

Your task is to:

1. Display the multimedia art, provide screenshots and a description of the steps taken.
2. Decide which aspects of the artworks to preserve, and identify their significant properties.
3. Develop a set of different preservation strategies for the four pieces of multimedia art provided, that have the potential to address different aspects of the artwork.
4. Point out the differences in the strategies with respect to the characteristics of the preserved artworks and their suitability.
5. (Optional) Implement part of the preservation strategies you have developed and submit code for this.

TASK 4A- Display the multimedia art, provide screenshots and a description of the steps taken

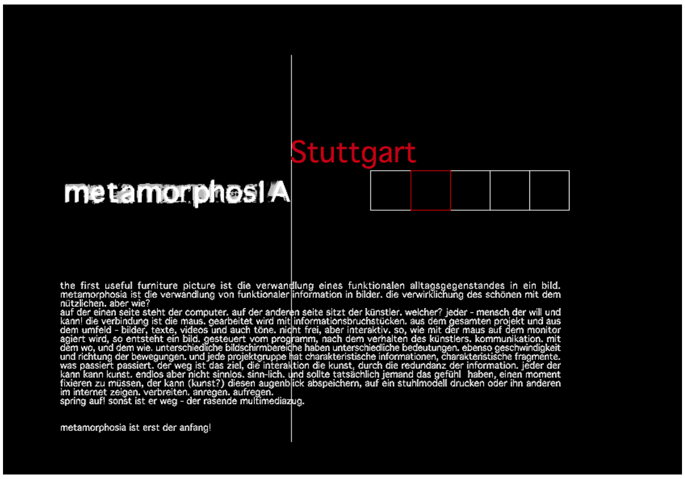
There were 4 folders. These folders were called

1998_PRIX_IA_90_MarcKleindienst_StefanBeuter_Metamorphosia
1998_PRIX_IA_104_SergioPerezMoretto_VaninaSteiner_CyberinstallationBotschaftAnEuropa2098
cyborg
interactive_paint

1998_PRIX_IA_90_MarcKleindienst_StefanBeuter_Metamorphosia

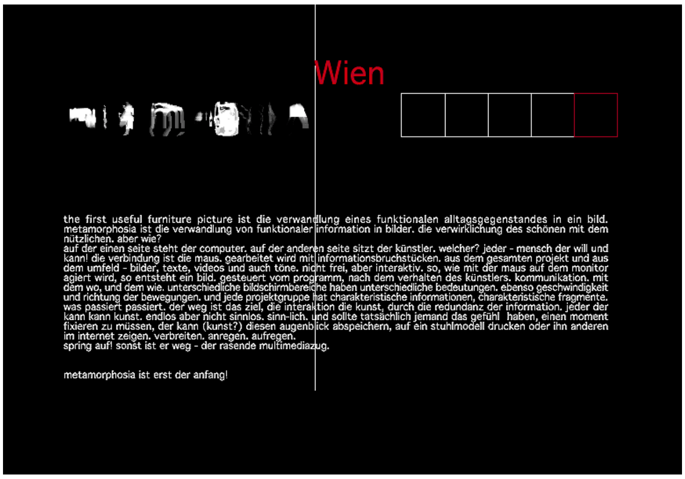
This is an application that works automatically double-click on the A logo. When is running is an interactive aplication with sounds and pictures.

Steps taken to this screenshots is a general wiew of the artwork. Main screenshots are shown. First part consist of a menu of artwork related to different cities. These cities München , Stuttgart, Hamburg, , Hannover and Wien



the first useful furniture picture ist die verwandlung eines funktionalen alltagsgegenstandes in ein bild. metamorphosia ist die verwandlung von funktionaler information in bilder. die verwirklichung des schönen mit dem nützlichen. aber wie?
auf der einen seite steht der computer. auf der anderen seite sitzt der künstler. welcher? jeder - mensch der will und kann! die verbindung ist die maus. gearbeitet wird mit informationsbruchstücken. aus dem gesamten projekt und aus dem umfeld - bilder, texte, videos und auch töne. nicht frei, aber interaktiv. so, wie mit der maus auf dem monitor agiert wird, so entsteht ein bild. gesteuert vom programm, nach dem verhalten des künstlers. kommunikation. mit dem wo, und dem wie. unterschiedliche bildschirmbereiche haben unterschiedliche bedeutungen. ebenso geschwindigkeit und richtung der bewegungen. und jede projektgruppe hat charakteristische informationen, charakteristische fragmente. was passiert passiert. der weg ist das ziel, die interaktion die kunst, durch die redundanz der information. jeder der kann kann kunst. endlos aber nicht sinnlos. sinn-lich. und sollte tatsächlich jemand das gefühl haben, einen moment fixieren zu müssen, der kann (kunst?) diesen augenblick abspeichern, auf ein stuhlmodell drucken oder ihn anderen im internet zeigen. verbreiten. anregen. aufregen.
spring auf! sonst ist er weg - der rasende multimediazug.

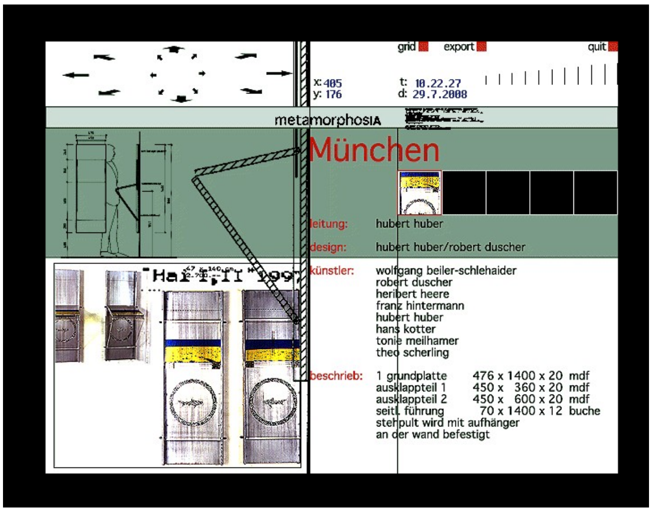
metamorphosia ist erst der anfang!



the first useful furniture picture ist die verwandlung eines funktionalen alltagsgegenstandes in ein bild. metamorphosia ist die verwandlung von funktionaler information in bilder. die verwirklichung des schönen mit dem nützlichen. aber wie?
auf der einen seite steht der computer. auf der anderen seite sitzt der künstler. welcher? jeder - mensch der will und kann! die verbindung ist die maus. gearbeitet wird mit informationsbruchstücken. aus dem gesamten projekt und aus dem umfeld - bilder, texte, videos und auch töne. nicht frei, aber interaktiv. so, wie mit der maus auf dem monitor agiert wird, so entsteht ein bild. gesteuert vom programm, nach dem verhalten des künstlers. kommunikation. mit dem wo, und dem wie. unterschiedliche bildschirmbereiche haben unterschiedliche bedeutungen. ebenso geschwindigkeit und richtung der bewegungen. und jede projektgruppe hat charakteristische informationen, charakteristische fragmente. was passiert passiert. der weg ist das ziel, die interaktion die kunst, durch die redundanz der information. jeder der kann kann kunst. endlos aber nicht sinnlos. sinn-lich. und sollte tatsächlich jemand das gefühl haben, einen moment fixieren zu müssen, der kann (kunst?) diesen augenblick abspeichern, auf ein stuhlmodell drucken oder ihn anderen im internet zeigen. verbreiten. anregen. aufregen.
spring auf! sonst ist er weg - der rasende multimediazug.

metamorphosia ist erst der anfang!

Once on the application, is one city is selected there is a sequence of images related to each cities on the left corner on the screen. When a click is done on this corner there is a possibility to interact with images and the mouse. Moving the mouse on the screen produces a set of sounds and pictures above this same image.



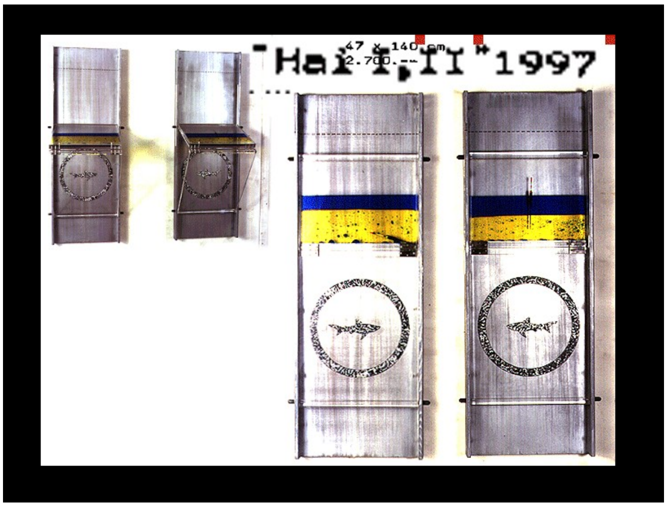
München

leitung: hubert huber

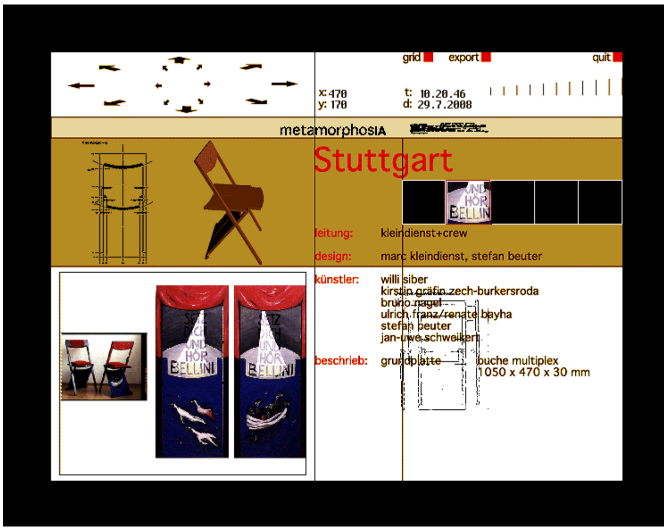
design: hubert huber/robert duscher

künstler: wolfgang beiler-schlehaider
robert duscher
heribert heere
franz hintermann
hubert huber
hans kotter
tonie meilhamer
theo scherling

beschrieb: 1 grundplatte 476 x 1400 x 20 mdf
ausklappteil 1 450 x 360 x 20 mdf
ausklappteil 2 450 x 600 x 20 mdf
seitl. führung 70 x 1400 x 12 buche
stehpult wird mit aufhänger
an der wand befestigt



As it may be seen on the next image is possible to see a difference on the image shown on the left and the image represented on the right. This is because when moving the mouse on the screen, image is filled with other images.



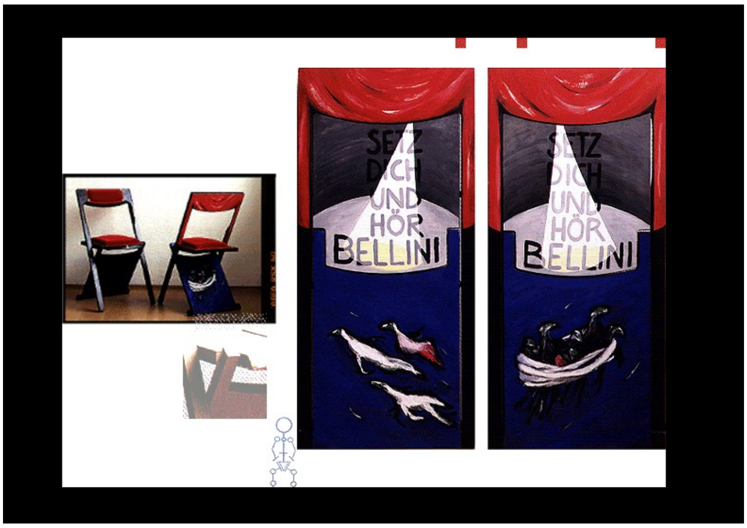
Stuttgart

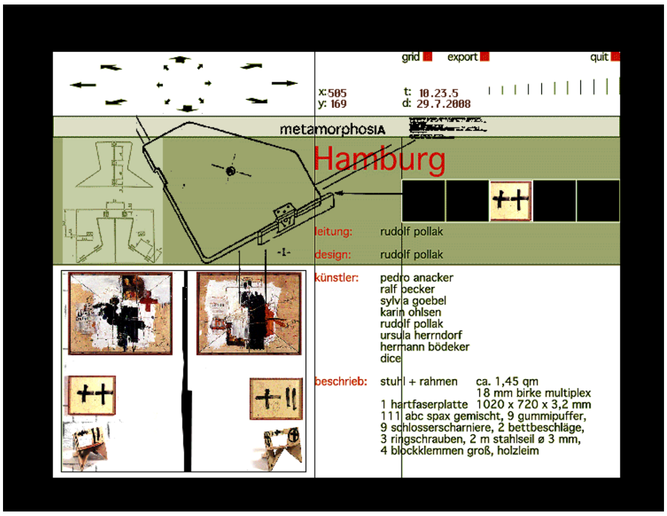
leitung: kleindienst+crew

design: marc kleindienst, stefan beuter

künstler: willi siber
kirstin gräfin zech-burkersroda
bruno nagel
ulrich franz/renate bayha
stefan beuter
jan-uwe schweikert

beschrieb: grundplatte buche multiplex
1050 x 470 x 30 mm





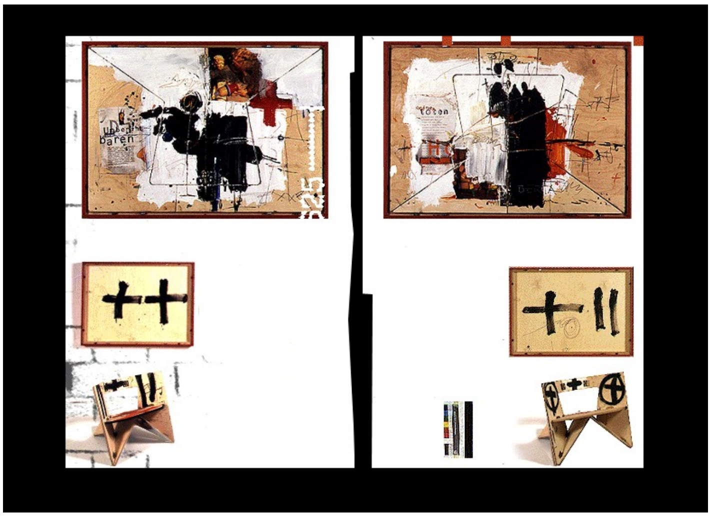
Hamburg

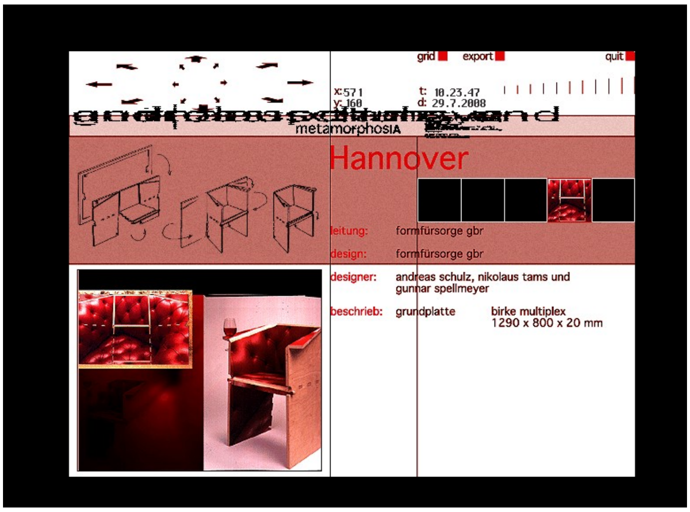
leitung: rudolf pollak

design: rudolf pollak

künstler: pedro anacker
ralf becker
sylvia goebel
karin ohlsen
rudolf pollak
ursula herrndorf
hermann bödeker
dice

beschrieb: stuhl + rahmen ca. 1,45 qm
18 mm birke multiplex
1 hartfaserplatte 1020 x 720 x 3,2 mm
111 abc spax gemischt, 9 gummipuffer,
9 schlosserscharniere, 2 bettbeschläge,
3 ringschrauben, 2 m stahlseil ø 3 mm,
4 blockklemmen groß, holzleim





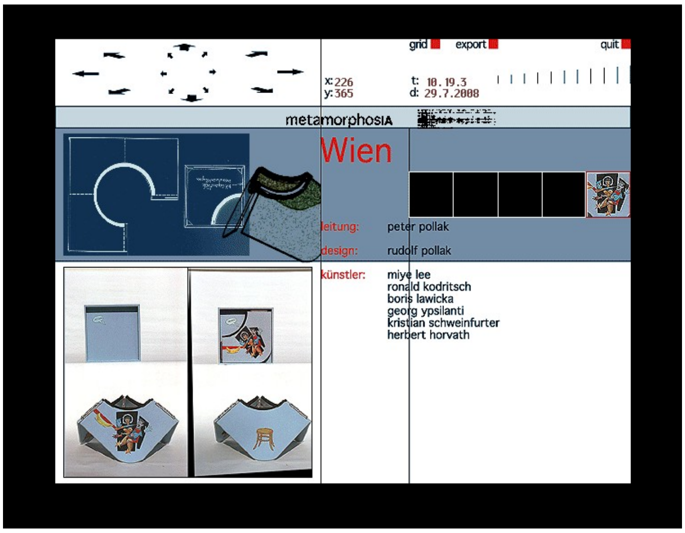
Hannover

leitung: formfürsorge gbr

design: formfürsorge gbr

designer: andreas schulz, nikolaus tams und gunnar spellmeyer

beschrieb: grundplatte birke multiplex 1290 x 800 x 20 mm

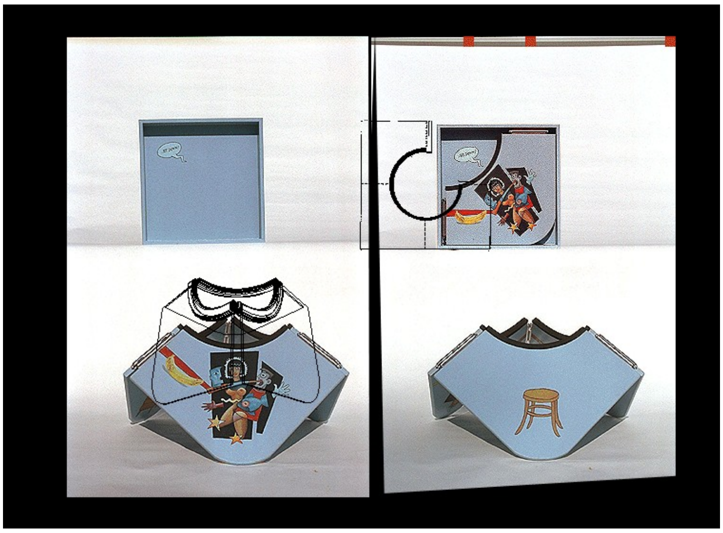


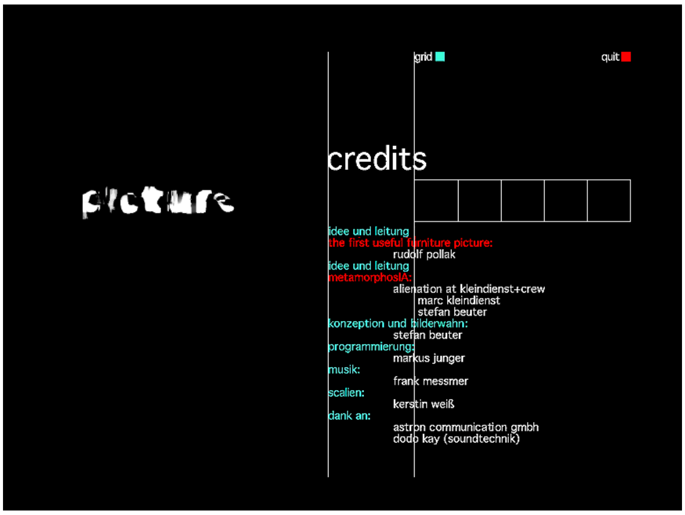
Wien

leitung: peter pollak

design: rudolf pollak

künstler: miye lee
ronald kodritsch
boris lawicka
georg ypsilanti
kristian schweinfurter
herbert horvath





credits

idee und leitung the first useful furniture picture:
rudolf pollak

idee und leitung metamorphosIA:
alienation at kleindienst+crew
marc kleindienst
stefan beuter

konzeption und bilderwahn:
stefan beuter

programmierung:
markus junger

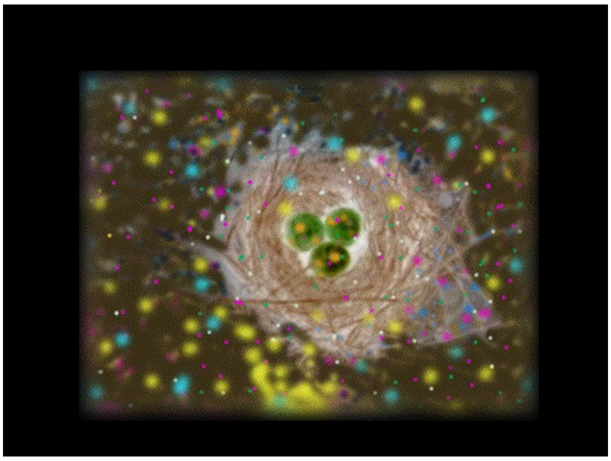
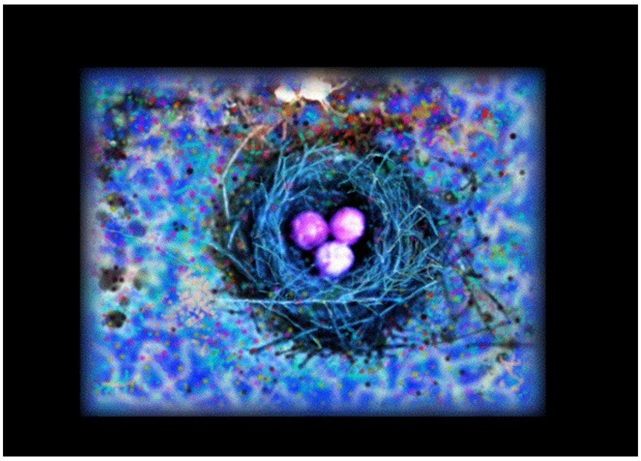
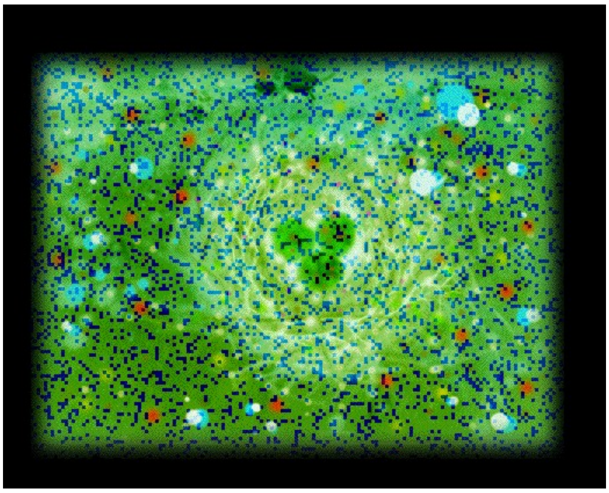
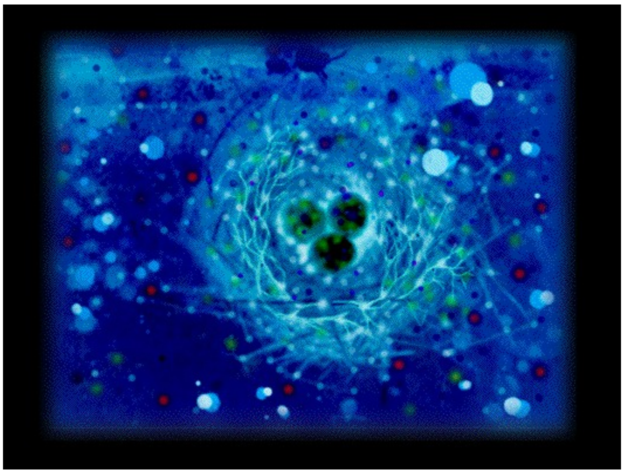
musik:
frank messmer

scalien:
kerstin weiß

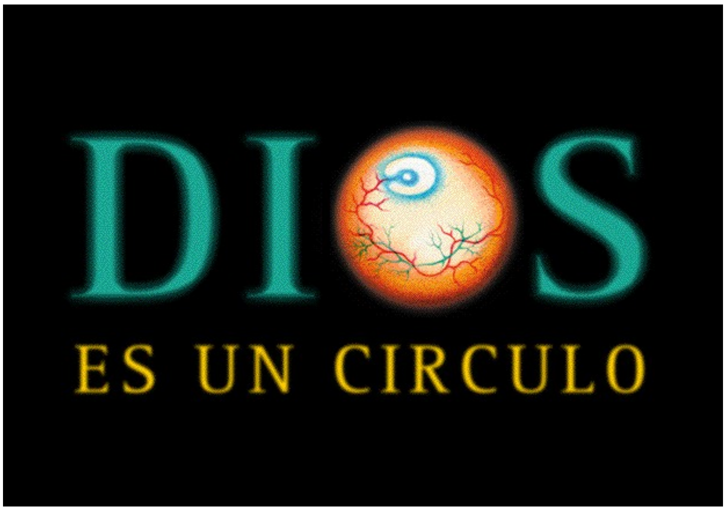
dank an:
astron communication gmbh
dodo kay (soundtechnik)

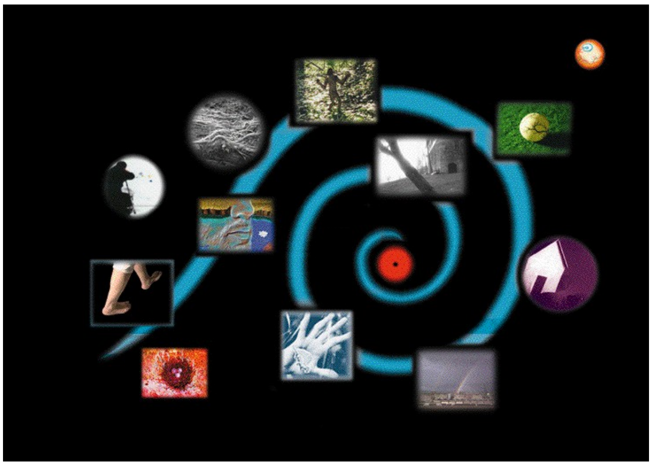
1998_PRIX_IA_104_SergioPerezMoretto_VaninaSteiner_CyberinstallationBotschaftAnEuropa2098

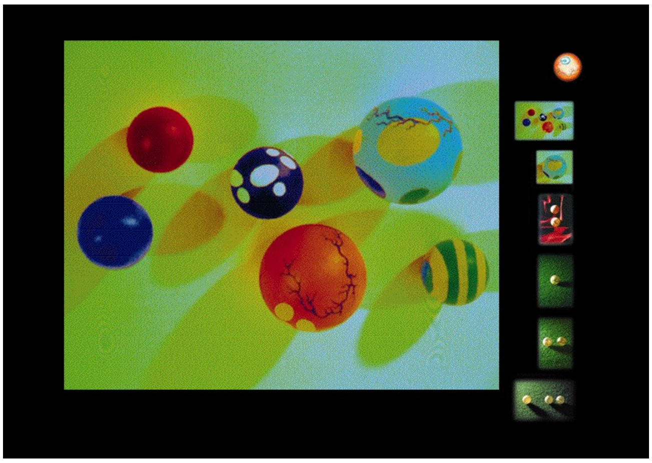
This is also a 16 bit application that runs on a 256 colour configuration. Screenshots taken have been a general view of the artwork. In this case are represented main screenshots from the three parts of the artwork, begging, intermedia part and general part

The images above is the beginnig of the artwork with a set of images that its colour is changing over time. Next image is the middle part and the message of the artwork. An screenshot is also provided



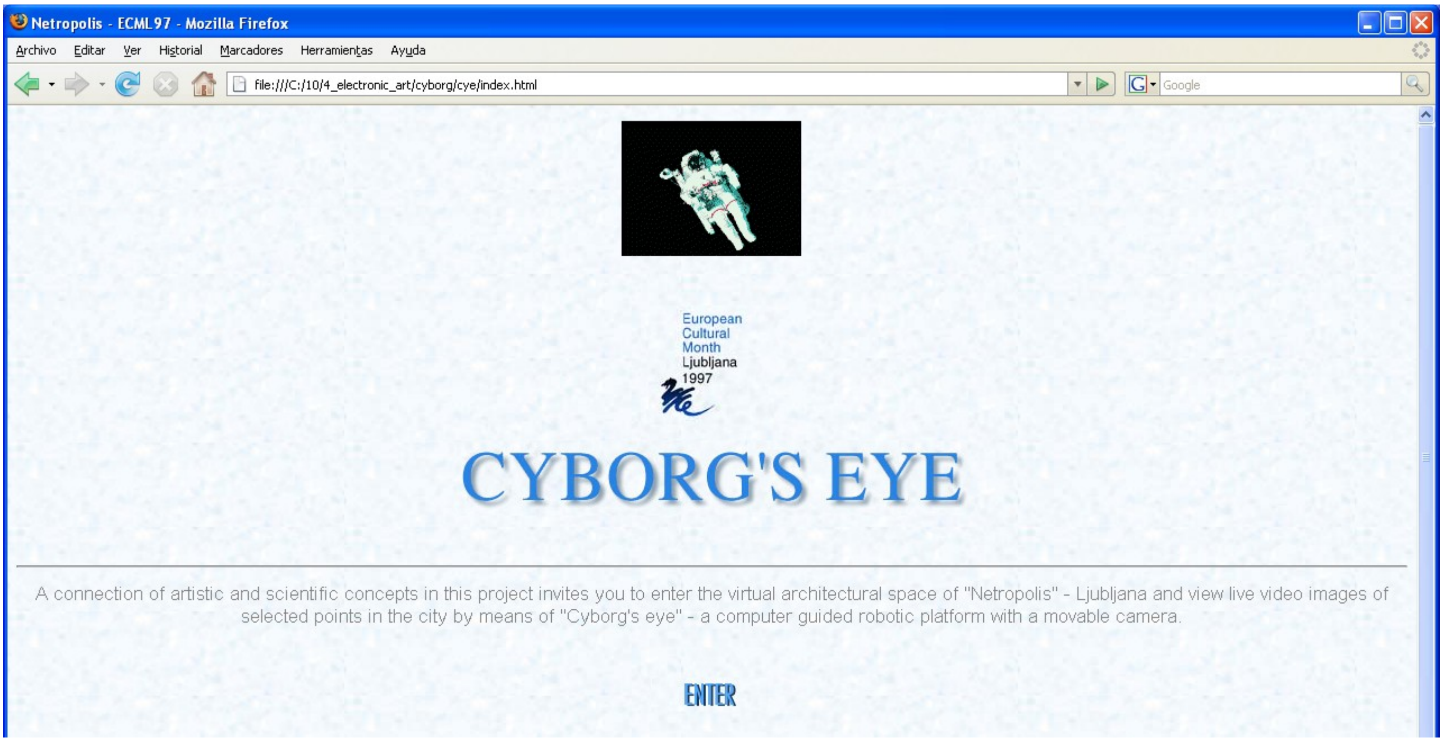




Third part of this artwork is this sequence of images when each of this images can be selected and it is possible to see mores images related to the selection as it can be seen on the images above. A representative part of this part can be seen on the right part.

cyborg

Cyborg's Eye was a project prepared for the European month of culture in Ljubljana (ECML), 15 May - 5 July 1997. In this artwork it is possible to see Virtual Ljubljana Walk Through.



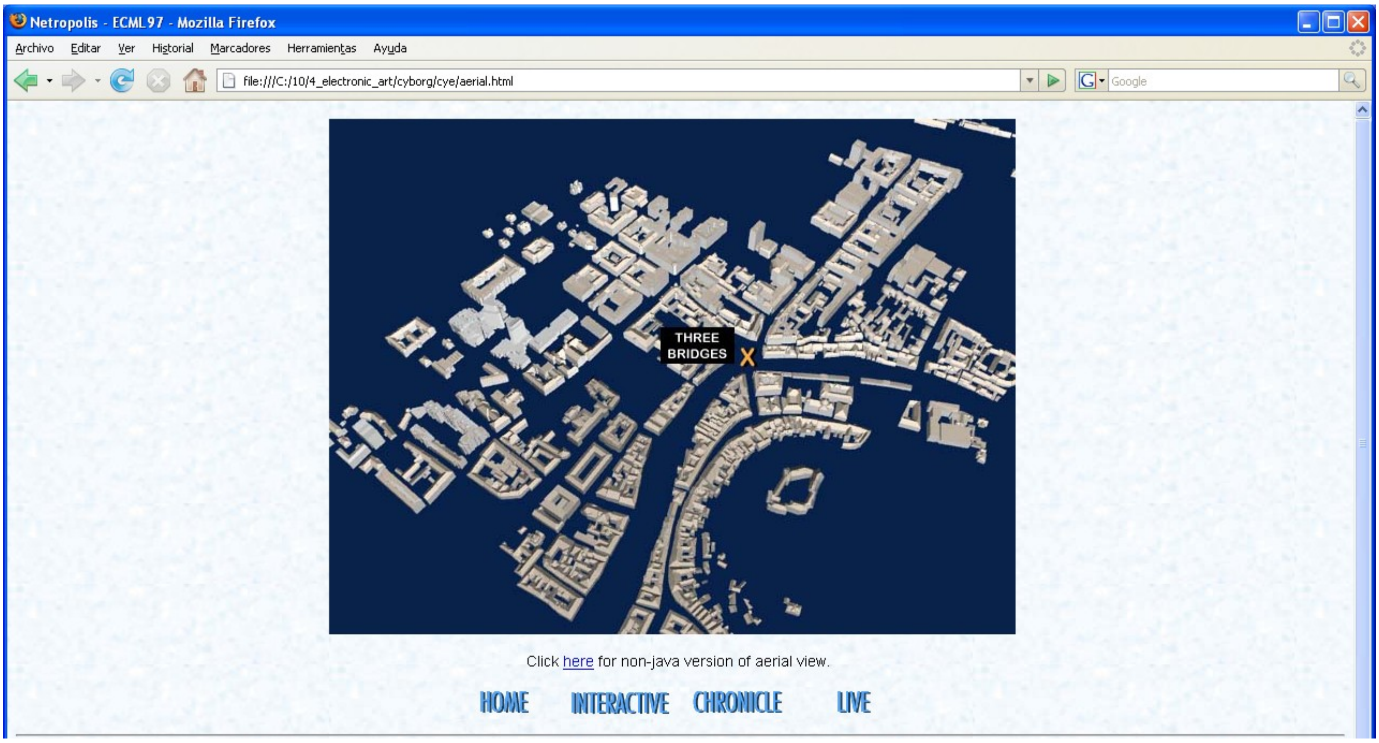
Netropolis - ECML97 - Mozilla Firefox
Archivo
Editar
Ver
Historial
Marcadores
Herramientas
Ayuda
file:///C:/10/4_electronic_art/cyborg/cye/index.html
Google

European
Cultural
Month
Ljubljana
1997

CYBORG'S EYE

A connection of artistic and scientific concepts in this project invites you to enter the virtual architectural space of "Netropolis" - Ljubljana and view live video images of selected points in the city by means of "Cyborg's eye" - a computer guided robotic platform with a movable camera.

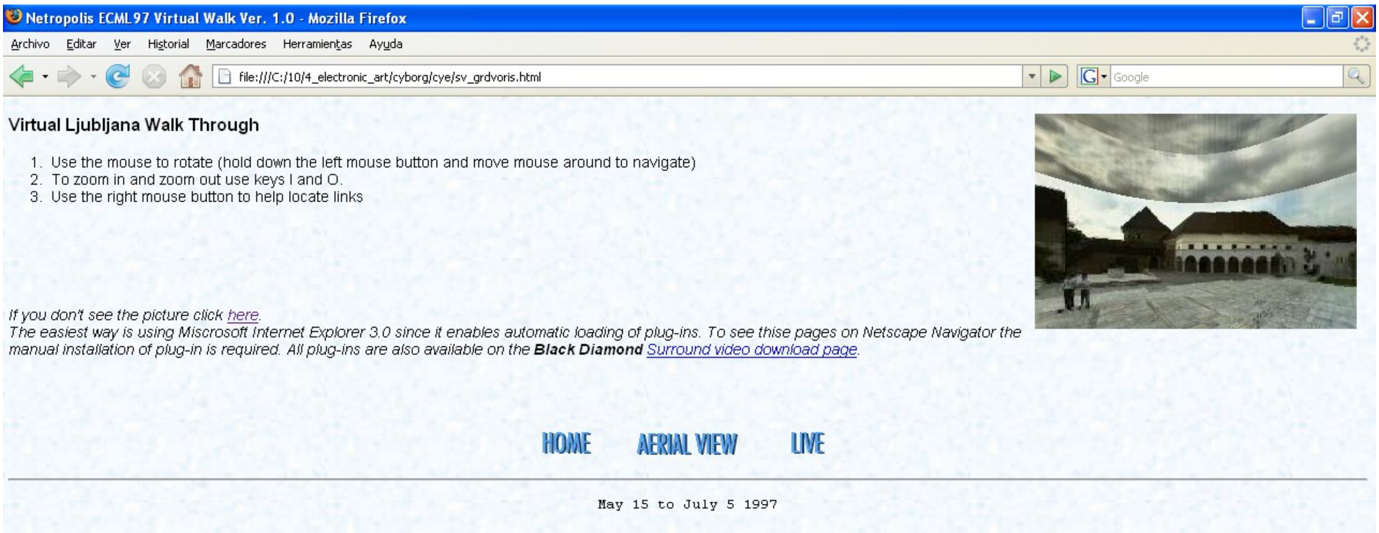
ENTER


Netropolis - ECML97 - Mozilla Firefox
Archivo Editar Ver Historial Marcadores Herramientas Ayuda
file:///C:/10/4_electronic_art/cyborg/cye/aerial.html
Google
THREE BRIDGES
HOME
INTERACTIVE
CHRONICLE
LIVE

Click here for non-java version of aerial view.

Virtual Ljubljana Walk Through is possible to be seen through the map above clicking on the part of the map that we want to see. Clicking on the interactive link it may be seen a panoramic view of Ljubljana and it is possible to interact with this panoramic view and going to the same places that can be selected on the map above.

On the image below there is a screenshot of the Virtual Ljubljana Walk Through with panoramic view. In the last it is possible to see that with this image it can be possible to interact through the browser.

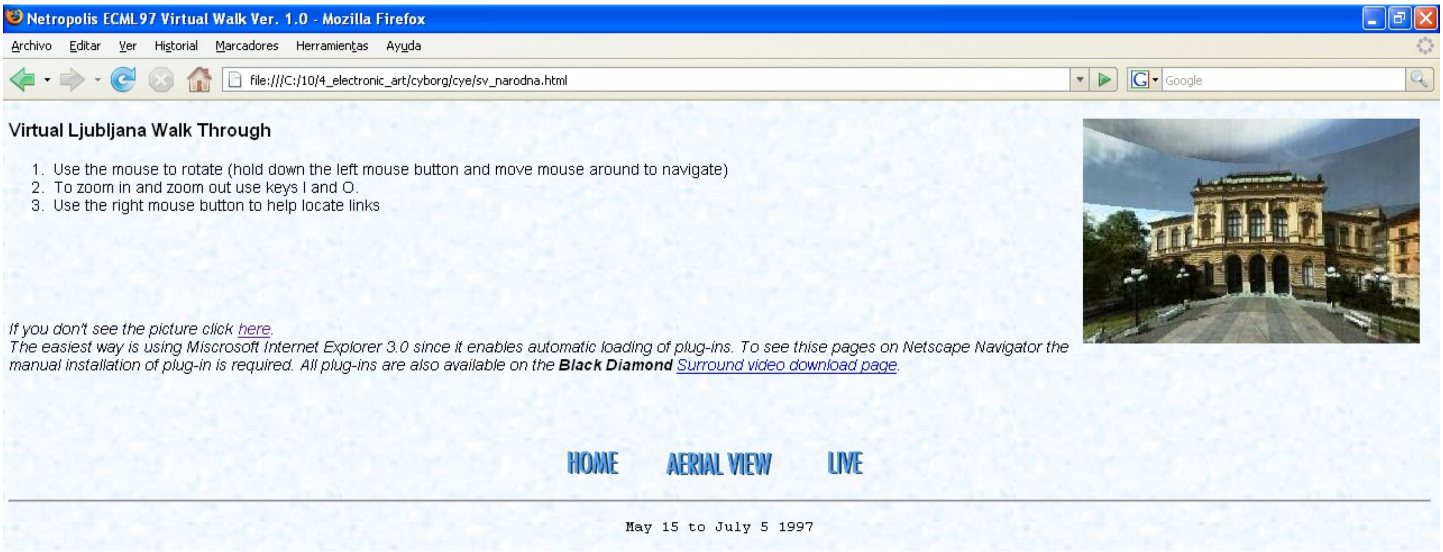

Netropolis ECML97 Virtual Walk Ver. 1.0 - Mozilla Firefox
Archivo Editar Ver Historial Marcadores Herramientas Ayuda
file:///C:/10/4_electronic_art/cyborg/cye/sv_grdvoris.html
Google
HOME
AERIAL VIEW
LIVE

Virtual Ljubljana Walk Through

1. Use the mouse to rotate (hold down the left mouse button and move mouse around to navigate)
2. To zoom in and zoom out use keys I and O.
3. Use the right mouse button to help locate links

If you don't see the picture click here.
The easiest way is using Miscrosoft Internet Explorer 3.0 since it enables automatic loading of plug-ins. To see thise pages on Netscape Navigator the manual installation of plug-in is required. All plug-ins are also available on the ***Black Diamond*** *Surround video download page.*

May 15 to July 5 1997



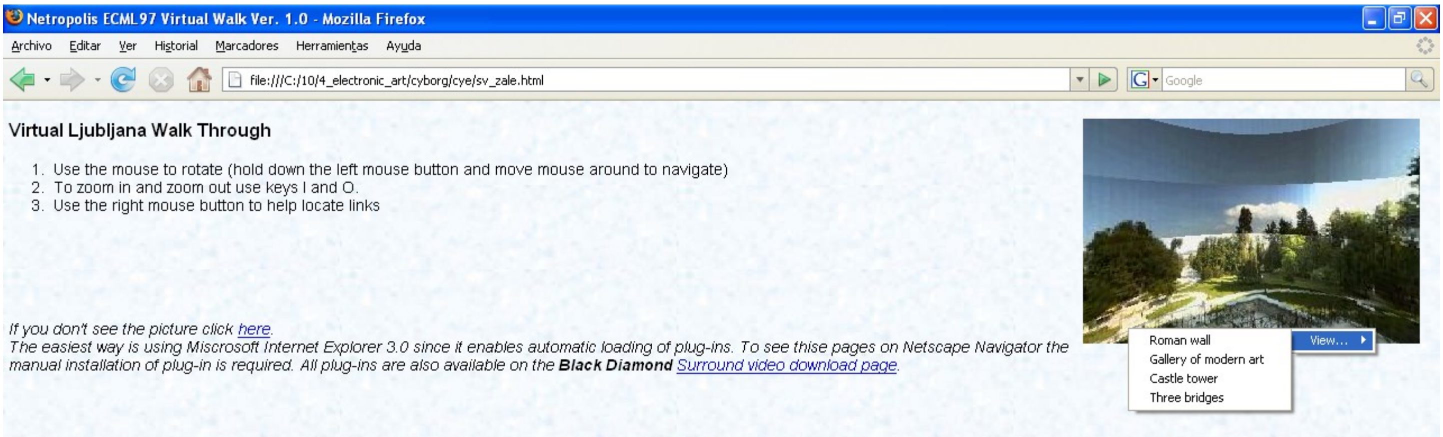
Netropolis ECML97 Virtual Walk Ver. 1.0 - Mozilla Firefox
Archivo Editar Ver Historial Marcadores Herramientas Ayuda
file:///C:/10/4_electronic_art/cyborg/cye/sv_narodna.html
Google

Virtual Ljubljana Walk Through

1. Use the mouse to rotate (hold down the left mouse button and move mouse around to navigate)
2. To zoom in and zoom out use keys I and O.
3. Use the right mouse button to help locate links

If you don't see the picture click here.
*The easiest way is using Miscrosoft Internet Explorer 3.0 since it enables automatic loading of plug-ins. To see thise pages on Netscape Navigator the manual installation of plug-in is required. All plug-ins are also available on the **Black Diamond** Surround video download page.*

May 15 to July 5 1997



Netropolis ECML97 Virtual Walk Ver. 1.0 - Mozilla Firefox
Archivo Editar Ver Historial Marcadores Herramientas Ayuda
file:///C:/10/4_electronic_art/cyborg/cye/sv_zale.html
Google
Roman wall
Gallery of modern art
Castle tower
Three bridges
View...

Virtual Ljubljana Walk Through

1. Use the mouse to rotate (hold down the left mouse button and move mouse around to navigate)
2. To zoom in and zoom out use keys I and O.
3. Use the right mouse button to help locate links

If you don't see the picture click here.
*The easiest way is using Miscrosoft Internet Explorer 3.0 since it enables automatic loading of plug-ins. To see thise pages on Netscape Navigator the manual installation of plug-in is required. All plug-ins are also available on the **Black Diamond** Surround video download page.*

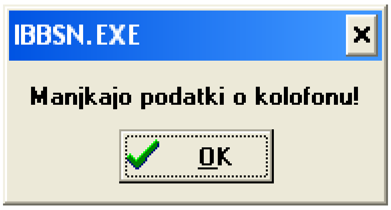
On of the main problems to see this artwork was precisely a plug-in that should be installed to the browser to se the panoramic view. The plug-in was NPSVIDEO.DLL that should be installed on

C:\Archivos de programa\Mozilla Firefox\plugins

This plug-in must installed manually not, automatically as the browser does.
According to the user help for this artwork, some links are provided to install the plug-in, but this links do not work anymore. This plug-in must be look for over the internet.

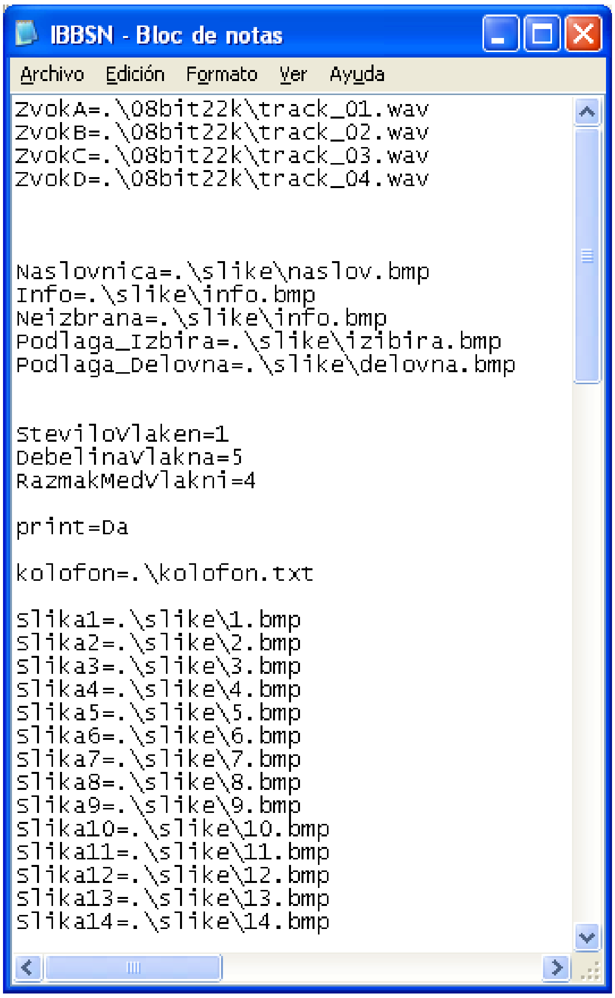
interactive_paint

This application has not been possible to run it. This is due to in the configuration file where “kolofon.txt” is indicated, was missed and there’s a message where it claims for data. It has been supposed that this program has not run to this reason.



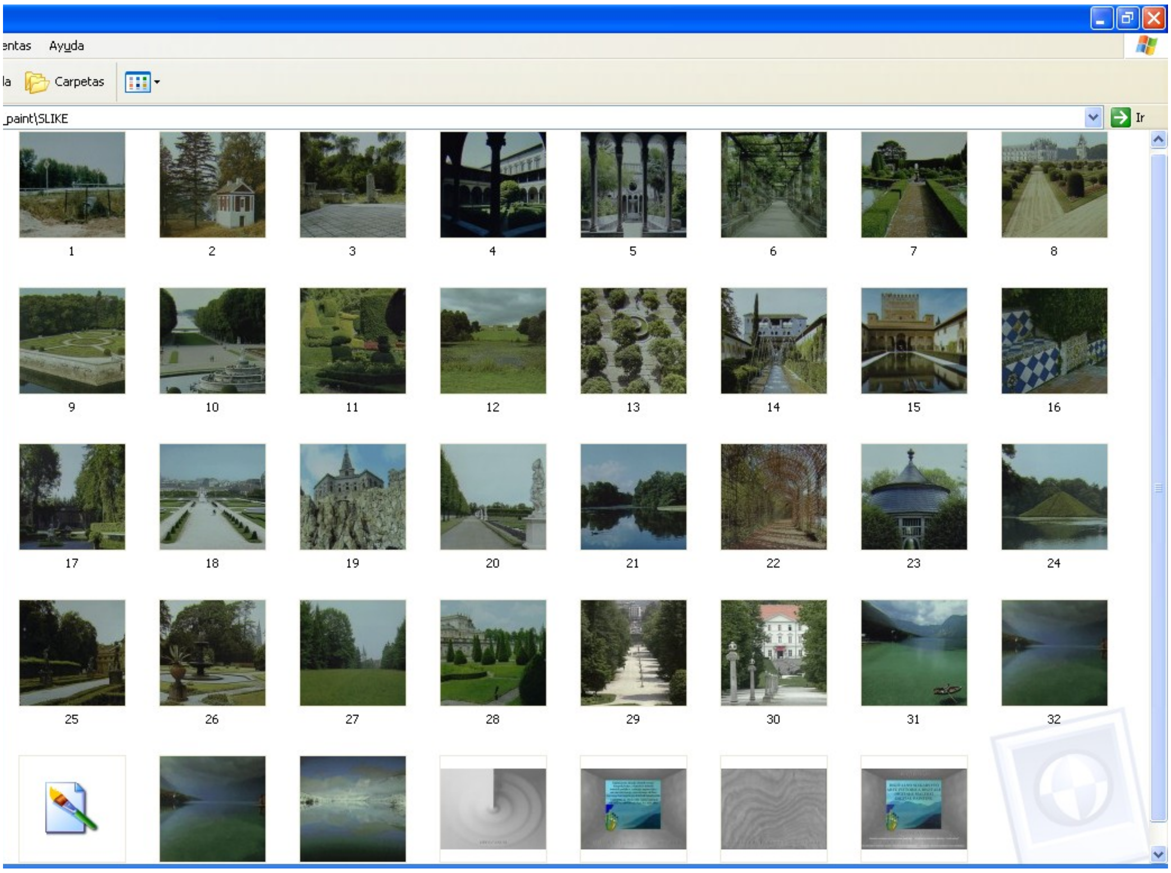
Manjkajo podatki o kolofonu!

This is the message that the programs displays. The message is in slovenian language and it says that data are required



```
ZvokA=.\08bit22k\track_01.wav
ZvokB=.\08bit22k\track_02.wav
ZvokC=.\08bit22k\track_03.wav
ZvokD=.\08bit22k\track_04.wav



Naslovnica=.\slike\naslov.bmp
Info=.\slike\info.bmp
Neizbrana=.\slike\info.bmp
Podlaga_Izbira=.\slike\izibira.bmp
Podlaga_Delovna=.\slike\delovna.bmp


SteviloVlaken=1
DebelinaVlakna=5
RazmakMedVlakni=4

print=Da

kolofon=.\kolofon.txt

Slika1=.\slike\1.bmp
Slika2=.\slike\2.bmp
Slika3=.\slike\3.bmp
Slika4=.\slike\4.bmp
Slika5=.\slike\5.bmp
Slika6=.\slike\6.bmp
Slika7=.\slike\7.bmp
Slika8=.\slike\8.bmp
Slika9=.\slike\9.bmp
Slika10=.\slike\10.bmp
Slika11=.\slike\11.bmp
Slika12=.\slike\12.bmp
Slika13=.\slike\13.bmp
Slika14=.\slike\14.bmp
```

On the other side, there a folder with images related to IBSSN application that may be seen as it is shown on the next image.



In this case, images images would be important to be preserved..

TASK 4B - Decide which aspects of the artworks to preserve, and identify their significant properties.

1998_PRIX_IA_90_MarcKleindienst_StefanBeuter_Metamorphosia

Associated to this application there is a folder with files that run with the application with .kc extension and folders with sound tracks.

Each time the mouse is sttoped and not moved, there are also screensavers that are relative to the authors of the artwork and the company who has done it.

Each presentation of expositions in München, Stuttgart, Hamburg, Hannover and Wien represents different interactions and different senses of methamorphosis.

There are also for different exposition cities, different authors and artist.

Main properties:

-Sounds
-Interaction with the screen
-Mixture of sounds
-The relationship between exposition and authors

In this case it would be desirable to preserve the whole application with its interactions, because interactions with this software allow to produces sounds and mixing of sounds. In this case it would be desirable the emulation in long term., because, sounds are important for the artwork.

This would represent to elaborate an emulator of the operating system for future generations.

1998_PRIX_IA_104_SergioPerezMoretto_VaninaSteiner_CyberinstallationBotschaftAnEuropa2098

In this job is important here the background sound and images, the transitions from one image to another.
Interaction between mouse and screen is also important to preserve because it allow us to view different artwork. Images and sound are also important and colours palette should be conserved.

Main properties:
-It's important transitions between the first images, different pixelation is used.

-Interaction between screen and images

-It's a nonfinished application. This means that there is no way to quit the application. This artwork is prepared to be constantly interactive

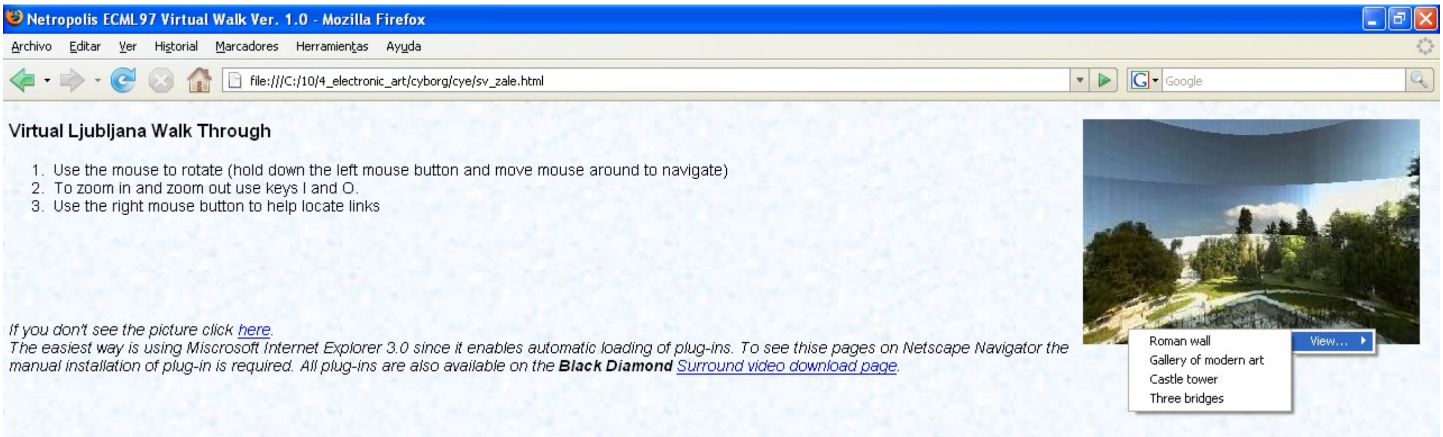
-There's a relation between the music and the draws on the screen. Transition in time and change of colours.

Cyborg

Cyborg is a virtual architectural space of "Netropolis" – Ljubljana. Due that is done on html, it could be interesting to translated the site to xml with a presentation layer.

Interactive images that run with javascript could be better transformed on a Flash® interactive image and transformed later to xml, due to Flash® allow xml transformations. Migration to a new plaftorm.

Main properties:
-Map of the city interactive with the user. Different view of the city.
-Walk around different part of the city. Sorround effects that are interactive with the mouse.
-Chronicle about cultural events and description with some images
-Interaction with the mouse to see other panoramic views as the image above.



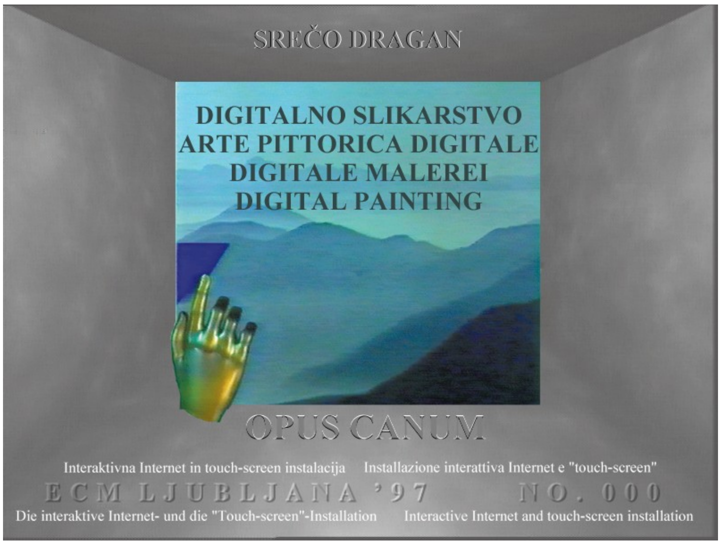
Netropolis ECML97 Virtual Walk Ver. 1.0 - Mozilla Firefox
Archivo Editar Ver Historial Marcadores Herramientas Ayuda
file:///C:/10/4_electronic_art/cyborg/cye/sv_zale.html
Google

Virtual Ljubljana Walk Through

1. Use the mouse to rotate (hold down the left mouse button and move mouse around to navigate)
2. To zoom in and zoom out use keys I and O.
3. Use the right mouse button to help locate links

If you don't see the picture click here.
The easiest way is using Miscrosoft Internet Explorer 3.0 since it enables automatic loading of plug-ins. To see thise pages on Netscape Navigator the manual installation of plug-in is required. All plug-ins are also available on the ***Black Diamond*** *Surround video download page.*

interactive_paint
Interactive paint is a program that allows paint over images and modify them, sign them , with an interactive screen as it is shown in one the images.


SREČO DRAGAN
DIGITALNO SLIKARSTVO
ARTE PITTORICA DIGITALE
DIGITALE MALEREI
DIGITAL PAINTING
OPUS CANUM
Interaktivna Internet in touch-screen instalacija
Installazione interattiva Internet e "touch-screen"
ECM LJUBLJANA '97
NO. 000
Die interaktive Internet- und die "Touch-screen"-Installation
Interactive Internet and touch-screen installation

Main properties:
-Images of the application may be preserved.
-As it was not possible to run the program, and all images are possible to see the whole images should be preserved as a part of the artwork.

TASK 4C - Develop a set of different preservation strategies for the four pieces of multimedia art provided, that have the potential to address different aspects of the artwork.

1998_PRIX_IA_90_MarcKleindienst_StefanBeuter_Metamorphosia
EMULATION.
This software shoudl be conserved as is and having emulation in the future. Due to its interaction with sound and music it would be necessary to write emulators in the future for windows 3.1 to run this kind of applications. It will have no much sense just to have screenshots without the music and the interaction either.

As a part of the strategy to have records explaining what kind of art is it and what kind of sound produces the interaction, this means to explain what happens in any screen related to a city and which is the connection to the music.

1998_PRIX_IA_104_SergioPerezMoretto_VaninaSteiner_CyberinstallationBotschaftAnEuropa2098
EMULATION
In this religious artwork it is also interesting to write emulators as in the artwork before and allow this applications to run. In case the work is interactive and its interesting to preserve also the interaction with the music, but to realize that the music is just a sound for the whole artwork. There is no other music when the user interacts with the screen.

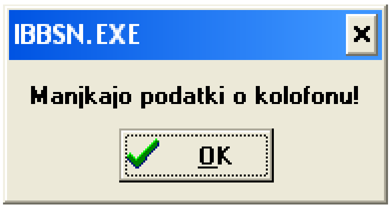
Other option to this artwork due to slow interaction that can be sequentially done may be a video screen with sound included.

SCREEN VIDEO
This is another option. In this case, as interactions are not so fast, in order to preserve, a video screening with sounds, following an order and doing this application step by step. Later the video conserved with MPEG2[30] format with sounds.
It's relevant in this case to produce a videoscreen as an alternative to the emulation because the artwork can be done in a sequential order, and the transitions between pictures does not depend on the mouse-click or the music. Everything seem to have an order in this artwork, so it's an alternative to emulation to preserve a video with interactions rather than preserve the software with an emulator. Cost is lower.
Cyborg

[30] http://www.esds.ac.uk/aandp/create/data.asp

XML translation. In order to preserve this job, XML translation could be an option. This would requiere an effort because an image who is interactive due to javascript, could be converted to flash interactive image and later with FLASH transformed to XML.
Panoramic views should be in a better flash format than .svh because a plug-in is very difficult to find it on the internet and they maybe converted to flash. The reason to convert this panoramic view in flash format is because that allo interactions in videos and later maybe converted to a file XML marked-up file.
In case it is also possible just convert svh files to MPEG-4 but inteaction with the screen as it is shown before wil be loose. It will on be possible to see panoramic views as a uninteractive film.

Interactive_paint
As the program does not run properly due to the lack of files, images can be preserved in a PNG format and having a record from this images.


IBBSN.EXE

Manjkajo podatki o kolofonu!

What it is interesting to preserve are the photos, who actually are in BMP format, so translate it to PNG all of them. The program could not be preserved because files are omitted and it cannot run accurately in the future.

TASK 4D - Point out the differences in the strategies with respect to the characteristics of the preserved artworks and their suitability.

Differences in strategies are quite different because all of them are different artworks with very different objectives.

Emulation will allow to have interactions with the whole application and having the possibility to run the whole artwork and seeing the interaction between mouse-clicking, art, design and sounds.

Screen Video will allow to see the whole artwork but in an order. The order which will be filmed.
Interaction with mouse in this case it may not be so important as in the first artwork, rather than to preserve the colours of the artwork and the music.

XML transforming seems to be logical to a website where the interactivity resides on the links and in the panoramic view. Losing interactivity with the panoramic view may be it’s important as the artwort itself, and videos may be recovered each time a place in Lublijana map is selected.

Converting the interactive panoramic view to a MPEG2 video it seems a solution just for not being forced to preserve the plug-in which is obsolete.

SCENARIO 5 -Web Archiving

Your company wants to preserve their website to document their growth and evolution over time. You are asked to analyse different preservation strategies for websites. The developed strategies will be applied to two internet domains and to analyse their advantages and disadvantages.

Task

Your task is to:

1. Devise at least two preservation strategies for websites, highlighting their respective advantages and disadvantages.
2. Harvest the two following internet domains documenting the date, time and method of harvesting:
   - o www.digitalpreservationeurope.eu with a depth of 3 (approximately 15MB of data)
   - o
   - o www.rai.it with a depth of 2 (approximately 40MB of data)
   - o
3. Apply the identified preservation strategies to the harvested websites, compare and document the results (for example storage size, processing time, presentation quality). Give an estimate of the resources (such as time, storage, effort, costs) required to deploy the strategies.

TASK 5 A- Devise at least two preservation strategies for websites, highlighting their respective advantages and disadvantages.

Preservation web strategies must assure in the future accessibility, authenticity and reliability of the data.

One of the mains problems concerned to web archiving strategies will be also author rights where legislation which is different in every country. Some general rules about authors rights are set in Europe through the Official Journal of European Union Directive 93/98/CEE, but it leaves to any country to make its own legislation over this topic.

One of the questions in a the future will be who will have the right to see web archiving and which information is relevant to be archived. Taking decisions about what is relevant or not will not have the same value now than in a future.

Other question to count on, is the depth of the harvested sites. Decide until which level a site must be harvested will depend also upon the value or the important of its information to the future.

According to these questions at least two strategies can be considered:

<u>MIGRATION:</u>

Once a site is harvested, it must be decided what to do with files that are harvested. Actually there is a very heterogeneity of files on the internet.

This means that in some years most of this files cannot be read with internet tools unless those files are updated or migrated. It will depend on the technology to support former files.

A sample of this question is also in the fourth task of this challenge with **cyborg eye's** when a plug-in can be downloaded form the original link and it has been looked for on the internet.

Migration of these files will include migrate, streams, pdf files, hypermedia text files and even files scripting languages such as php pages, perl pages and others.
Migration will include also images in a wide variety of formats, i.e: GIFF, TIFF, JEG,PNG, etc.

Migration will force to analyse each document of each site and migrate it according to new systems in order to be accessible, reliable and authentic.

Relationship between the cost and the effectiveness of migration should be considered before migration is going to be done.

The main advantage for this strategy would be that the whole information which has been harvested will be possible to been with new tools in a future.

Main disadvantage is that elaborate a migration plan for harvested websites will be very expensive due to the heterogeneity of the documents.

LET THE HARVESTED SITE AS ORIGINAL

If document are left as they're actually will mean that probably to see harvested documents is probably that an emulator servers will be needed.

Relationship between cost and effectiveness in this case will be better and costless because there should be just on software to run an emulator server, but this emulators should support all kind all files that actually may be seen on the internet.

The main advantage of this strategy is that information will be kept as is and information will have authenticity, reliability but accessibility is not guaranteed.

The main disadvantage is that to assure accessibility future server web emulators must be able to read all kind of files unless accessibility to the information will not be possible.

TASK 5 B
Harvest the two following internet domains documenting the date, time and method of harvesting:

Method of harvesting:

According to Marill et al[31], there are three kind of crawler according to some questions: open source, documented prior use, and an active community of developers and users. HTTRACK, HETRITRIX and NEDLIB.

HTTACK is very valuable for site analysis but not for wide-scale harvesting. NEDLIB depends on a database on MYSQL and it lacks a direct user interface. Heritrix is driven by an XML configuration language, which supports complex crawl definitions and filtering. In addition, it appears to support advanced customisation via Java plug-ins. Heritrix includes a Web hosted control panel for managing and monitoring crawls.

In our case HERITRIX have been used. It's a software that has been tested with Linux[32], so Heritrix has been used under LINUX Ubuntu.

Both domains have been harvested at the same time having count on that one of them had 2 levels to be hardvested and the other 3 levels to be harvested. According to the task three levels from www.digitalpreservationeurope.eu have been crawled and two levels form www.rai.it have been also crawled. No subdomains have been crawled.

As a reports crawl-report, mimetype-report, seeds-report are included. Other files relative to the harvester action are included on the files handed

In our case files have been left original and not migrated.

[31] Marill et al; (2004); Tools and Techniques for Harvesting the World Wide Web, http://csdl2.computer.org/comp/proceedings/jcdl/2004/2493/00/24930403.pdf

[32] Internet Archive (2008); Heritrix –Home page, http://crawler.archive.org/index.html

Crawl-report

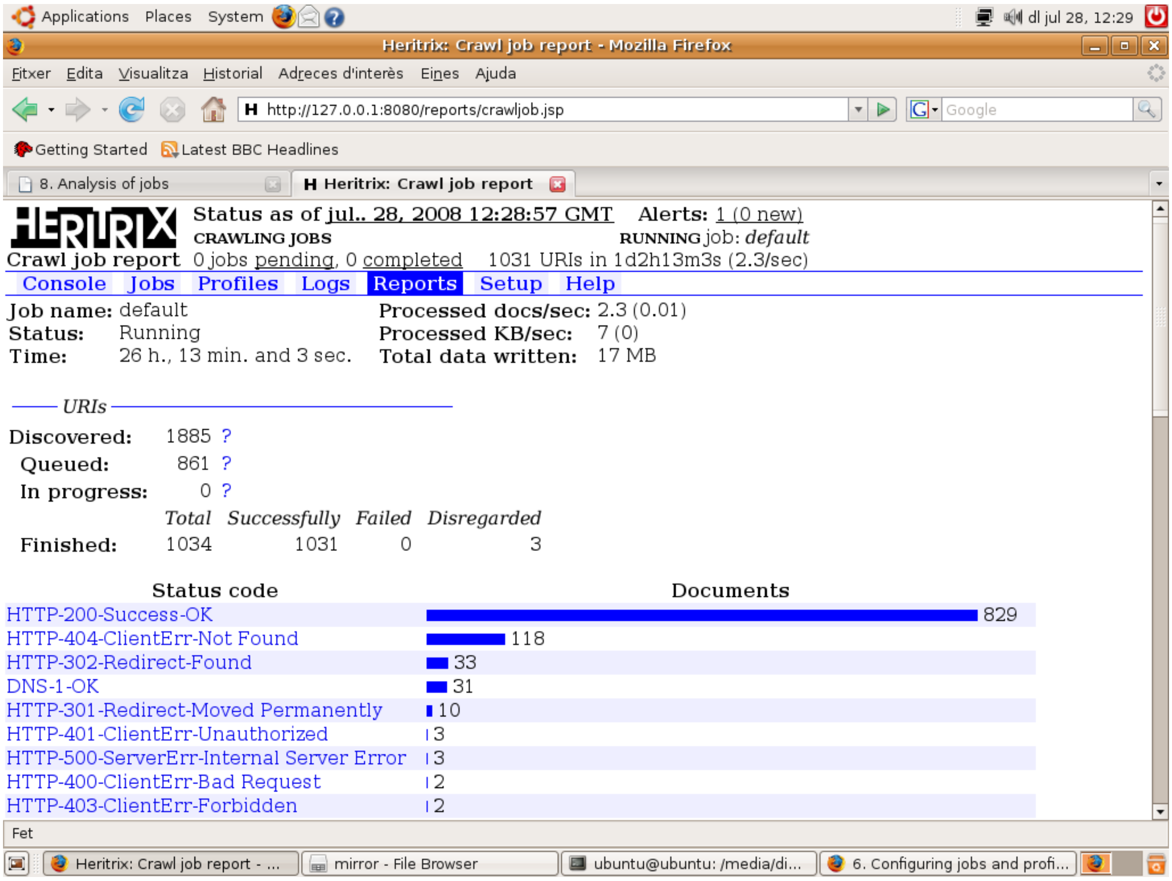
```
Crawl Name: TODO
Crawl Status: Finished - Ended by operator
Duration Time: 6h10m54s992ms
Total Seeds Crawled: 2
Total Seeds not Crawled: 0
Total Hosts Crawled: 150
Total Documents Crawled: 18797
Processed docs/sec: 0.84
Bandwidth in Kbytes/sec: 14
Total Raw Data Size in Bytes: 335122090 (320 MB)
Novel Bytes: 335122090 (320 MB)
```

Mimetype-report

```
[#urls] [#bytes] [mime-types]
14309 151604054 text/html
1735 6008744 image/gif
1275 23032537 image/jpeg
254 1376514 image/png
227 113823651 application/pdf
170 1910642 application/x-javascript
150 10437 text/dns
142 1469225 text/css
123 4522146 text/plain
100 32742 audio/x-pn-realaudio
36 12188378 application/x-shockwave-flash
33 5412 application/x-netgravity
10 18945 image/x-icon
10 156234 text/xml
8 6979157 application/vnd.ms-powerpoint
5 39849 application/javascript
5 4749 text/x-component
4 47246 application/xml
4 1713 no-type
4 11443716 video/x-ms-wmv
1 38287 application/atom+xml
1 219403 application/msword
1 31097 application/rdf+xml
1 31692 application/rss+xml
1 101081 image/bmp
1 1566 image/vnd.microsoft.icon
1 22873 text/javascript
```

seeds-report

```
[code] [status] [seed] [redirect]
200 CRAWLED http://www.digitalpreservationeurope.eu/
200 CRAWLED http://www.rai.it/
```



Applications Places System
dl jul 28, 12:29
Heritrix: Crawl job report - Mozilla Firefox
Fitxer Edita Visualitza Historial Adreces d'interès Eines Ajuda
http://127.0.0.1:8080/reports/crawljob.jsp
Google
Getting Started
Latest BBC Headlines
8. Analysis of jobs
Heritrix: Crawl job report
HERITRIX
Console Jobs Profiles Logs Reports Setup Help
Fet
Heritrix: Crawl job report - ...
mirror - File Browser
ubuntu@ubuntu: /media/di...
6. Configuring jobs and profi...

Crawl job report

Status as of jul.. 28, 2008 12:28:57 GMT Alerts: 1 (0 new)

CRAWLING JOBS RUNNING job: *default*

0 jobs pending, 0 completed 1031 URIs in 1d2h13m3s (2.3/sec)

Job name:	default	**Processed docs/sec:**	2.3 (0.01)
Status:	Running	**Processed KB/sec:**	7 (0)
Time:	26 h., 13 min. and 3 sec.	**Total data written:**	17 MB

URIs

Discovered:	1885 ?			
Queued:	861 ?			
In progress:	0 ?			
	Total	*Successfully*	*Failed*	*Disregarded*
Finished:	1034	1031	0	3

Status code	Documents
HTTP-200-Success-OK	829
HTTP-404-ClientErr-Not Found	118
HTTP-302-Redirect-Found	33
DNS-1-OK	31
HTTP-301-Redirect-Moved Permanently	10
HTTP-401-ClientErr-Unauthorized	3
HTTP-500-ServerErr-Internal Server Error	3
HTTP-400-ClientErr-Bad Request	2
HTTP-403-ClientErr-Forbidden	2

a screenshot of some essay done before with heritrix

TASK 5C - Apply the identified preservation strategies to the harvested websites, compare and document the results (for example storage size, processing time, presentation quality). Give an estimate of the resources (such as time, storage, effort, costs) required to deploy the strategies

Here there is a table with harvested files on both sites.

http://www.digitalpreservationeurope.eu		
Type of files	Number	Size
PDF	143	62.0 MB
HTML	2668	48,7 Mb
GIF	44	66.9 Kb
JPG	60	1.9 Mb
PNG	129	525.2 Kb
JAVASCRIPT	1	23.9 Kb
CSS	5	39.9 Kb
PPT	8	6.7 Mb
DOC	2	8.5 Kb
XML	1	22.8 Kb
SWF	1	630 Kb

http://www.rai.it		
Type of files	Number	Size
PDF	14	2, 5Mb
HTML	6045	66.2 Mb
GIF	649	3.3 Mb
JPG	907	14.4 Mb
PNG	37	382.4 Kb
JAVASCRIPT	1	23.9 Kb
CSS	27	357 Kb
PPT	0	0 Kb
DOC	2	8.5 Kb
XML	18	41.1 Kb
SWF	27	1.4 Mb

Comparing the volume of the downloaded information it can be seen the major weigth in www.digitalpreservationeurope.eu is to harvest the pfd files and ppt because are heaviest than html files.

On the other hand, www.rai.it has a big volume of html files and may slow the harvest of the site.
Most of these pages comes from a CMS application and probably those pages will be changed on the next harvested of the site.

JPG images are also a big quantity of files.

As an estimate effort for this strategy, it can be seen than harvesting www.digitalpreservationeurope.eu in short and mid term will cost more than www.rai.it because information weights more due to the PDF documents and ppt document and its download is slower than a html file which is really small in comparison to pdf file.

Storage needed will be greater in rai.it due to its CMS applications. Updates of this sites a are more frequent than in www.digitalpreservationeurope.eu so, www.rai.it should be harvested more frequently.

APPENDIX

Software employed

LINUX UBUNTU

A distribution of LINUX UBUNTU 7.04 – The Feisty Fawn – released in April 2007.
It has been used to harvest the websites. Ubuntu is an entirely open source operating system built around Linux Kernel. It's under GNU licence.

WINDOWS XP PROFESSIONAL

Windows XP professional is a proprietary Operating system from Microsoft®. Most of the task to be done where possible to do it in this proprietary operating systems.

LOTUS 1-2-3 SMARTSUITE

A version from this software has been used to do the first task. Lotus Smartsuite® is a software with spreadsheets, text processors, agenda, etc

HERITRIX

Heritrix is the Internet Archive's open-source, extensible, web-scale, archival-quality web crawler project.
Heritrix (sometimes spelled *heretrix*, or misspelled or missaid as *heratrix*/*heritix*/ *heretix*/*heratix*) is an archaic word for *heiress* (woman who inherits). Since our crawler seeks to collect and *preserve* the digital artifacts of our culture for the benefit of future researchers and generations, this name seemed apt.

ATARI EMULATORS

Different ATARI emulators have been used to do the second task. The mainly used emulator has been XLFORMER due to its usability in front of the other emulators.

XMFORMER2000 is the Atari 8-bit emulator for Windows Is compatible with all 32-bit versions of Windows, including Windows 2000, Windows XP, and Windows 2003.

GRAPH2FONT

The Graph2Font (G2F) program treats the image as fields with dimensions of 4x8, 2x8, 8x8 pixels. These fields are charset(s) (fonts) and if only some field is repeated then it's exchanged with adequate sign (font). This is some kind of loseless image compression. At the beginning this program had been used only to convert graphics to chars but now it also allows you to do multicolor graphics for the 8-bit Atari.

WINISIS

Winsis is a generalised Information Storage and Retrieval system. The Windows version may run on a single computer or in a local area network. In our case 1.3 version has been used.

ISIS2XML

Is a freeware software that allow translate ISIS file to XML files. The programmer is Pierre Chabert.

REFERENCES

APE . Atari Pheripheral Emulator for DOS and Windows – Home Page; *Atari Disk Image FAQ* [online] http://www.atarimax.com/ape/docs/DiskImageFAQ/ [available: july 2008]

AHDS History (2005) *Preservation Handbook spreadsheets.* Arts and Humanities Data Service [online] http://ahds.ac.uk/preservation/spreadsheets-preservation-handbook.pdf [available: july 2008]

ATARI; Wikimedia Foundation, Inc., [online] http://en.wikipedia.org/wiki/ATARI [available: july 2008]

Brickley, D. et al. *Resource Description Framework (RDF)* Semantic Web Interest Group Chair [online] http://www.w3.org/RDF/ [available: july 2008]

Dracon, M (2005); *:: Graph2Font :: (Atari XE/XL) program is freeware* [online] http://g2f.atari8.info/ [available: july 2008]

Definition of Marc ibiblio.org [online] http://www.ibiblio.org/msmckoy/marc2.html [available: july 2008]

Emulators Online – Run the Mac Os on Windows!; Emulators Inc [online] http://www.emulators.com/ [available: july 2008]

Electronic Resources for Special Collections; James Hardiman Library, National University of Ireland [online] http://www.library.nuigalway.ie/resources/special_collections/Special_Collections_-_Electronic_resources.html [available: july 2008]

Heritrix Open Source Harvester, Best Practices Exchange 2009 Arizona State Library, Archives and Public Records [online] http://www.bpexchange.org/2008/materials/IA%20Heritrix%20preso.ppt [available: july 2008]

Herman, Ivan.; *Web Ontology Language (OWL).* (W3C) Semantic Web Activity Lead [online] http://www.w3.org/2004/OWL/ [available: july 2008]

IBM Lotus Software. International Business Machines Corp.[online] http://www-306.ibm.com/software/lotus/ . [Available: july 2008]

Internet Archive (2007) *Heritrix 1.12.0 – Crawling Smarter* Internet Archive [online] http://wa.archive.org/blog/2007/03/17/heritrix-1120-crawling-smarter/

Internet Archive (2008) *Heritrix Home Page*; Internet Archive [online] http://crawler.archive.org [available: july 2008]

Irish Virtual Research Library and Archive [online] http://www.ucd.ie/ivrla/workpackages.html [available: july 2008]

Jelsoft Enterprises Ltd, [online] http://filext.com/index.php . [available july 2008]

Liebowtiz, Stan; *Stan Liebowtiz Home Page.*[online] University of Texas at Dallas http://www.utdallas.edu/~liebowit/rsle1.sam; [available: july 2008]

Lilley, Chris; PNG (Portable Graphics Network). World Wide Web Consortium [online] http://www.w3.org/Graphics/PNG/ [available: july 2008]

Lotus Software; Wikimedia Foundation, Inc., [online] http://en.wikipedia.org/wiki/Lotus_Software. [Available july 2008]

Marill, J et al. (2004) *Tools and Techniques for Harvesting the World Wide Web* ; The Library of Congress [online]; http://csdl2.computer.org/comp/proceedings/jcdl/2004/2493/00/24930403.pdf [available: july 2008]

Ministerio de Ciencia e Innovación; *Instituto de Estudios Documentales sobre Ciencia y Tecnología* [online] http://www.cindoc.csic.es/ [available: july 2008]

Plan de Acción 1998-2000.[on line] Sociedad Matemática Thales. Consejeria de Educación y Ciencia. Junta de Andalucia. 1998. http://thales.cica.es/rd/Recursos/rd99/ed99-0286-01/acti-ami.sam [accessible: july 2th, 2008]

Phillipe PVBest's Atari 8bits Web Site [online] http://pvb.free.fr/Atari/index.php [available: july 2008]

Rich Text Format (RTF) Specification, version 1.6 Microsoft Corporation 2008 [online] http://msdn.microsoft.com/en-us/library/aa140277.aspx, [available: july 2008]

Stehlik, Peter (2003) *Atari800 – Windows* [online] http://atari800.sourceforge.net/download.html [available: july 2008]

Technical Background – Web Capute (Library of Congress); The Library of Congress [online] http://www.loc.gov/webcapture/technical.html [available: july 2008]

The Atari ++ HomePage [online] http://www.math.tu-berlin.de/~thor/atari++/ [available:july 2008]

The Library Of Congress (2008) *MARC STANDARDS*; The Library of Congress[online] http://www.loc.gov/marc/ [available: july 2008]

UNESCO (2008); *CDS/ISIS database software: UNSECO-CI;* UNESCO [online] http://portal.unesco.org/ci/en/ev.php-URL_ID=2071&URL_DO=DO_TOPIC&URL_SECTION=201.html
WinIsis. BIREME [online] http://productos.bvsalud.org/product.php?id=winisis&lang=es [available: july 2008]

Universities of Essex and Manchester (2003-2008) *Data Formats and Software*. Economic and Social Data Service [online] http://www.esds.ac.uk/aandp/create/data.asp [available: july 2008]