



DIPL.-INF. JENS-MARTIN LOEBEL
INFORMATIK IN BILDUNG UND GESELLSCHAFT
INSTITUT FÜR INFORMATIK
HUMBOLDT-UNIVERSITÄT ZU BERLIN

AG-LANGZEITBEWAHRUNG & AG-MULTIMEDIA
DEUTSCHER MUSEUMSBUND
12.10.2009
ZUSE INSTITUT



Langzeitarchivierung digitaler multimedialer Objekte

Ein Überblick

Gliederung

- * Einführung
- * Begriffsdefinitionen
- * Probleme digitaler Medien
- * „Multimedia“
 - * Probleme und Perspektiven
- * Workshop Emulation

Langzeitarchivierung

Langzeitarchivierung (engl.: digital preservation, long-term preservation)

'Langzeit' ist die Umschreibung eines nicht näher fixierten Zeitraumes, währenddessen wesentliche nicht vorhersehbare technologische und soziokulturelle Veränderungen eintreten, die sowohl die Gestalt als auch die Nutzungssituation digitaler Ressourcen in rasanten Entwicklungszyklen vollständig umwälzen werden. (Zitat aus: Grundlagen der praktischen Information und Dokumentation. - München, 2004)

Langzeitarchivierung digitaler Objekte umfasst alle Maßnahmen, die dazu dienen, digitale Objekte für die Nachwelt dauerhaft zu erhalten.

NESTOR - Kompetenznetzwerk Langzeitarchivierung

(BILD AUS URHEBERRECHTLICHEN
GRÜNDEN ENTFERNT)

Analog vs. Digital

- unmittelbar lesbar
- an physikalisches Trägermedium gebunden
- keine verlustfreie Kopie
- repräsentiert durch Bitstrom
- an physikalisches Trägermedium gebunden
- nur mittelbar lesbar, müssen interpretiert werden
- relativ einfach verlustfrei kopierbar

(BILD AUS URHEBERRECHTLICHEN
GRÜNDEN ENTFERNT)

Haltbarkeit

(BILD AUS URHEBERRECHTLICHEN
GRÜNDEN ENTFERNT)

uralte Steintafel vs.
Magnetband 80er Jahre

(BILD AUS URHEBERRECHTLICHEN
GRÜNDEN ENTFERNT)

Medium	Lebensdauer
Steintafeln	mehrere tausend Jahre
Bücher (säurefreies Papier)	mehrere hundert Jahre
Mikrofilm	ca. 500 Jahre (eher 50 Jahre)
CD-ROM / DVD (optische Medien)	25-100 Jahre (eher 3-5 Jahre)
Disketten	5-10 Jahre
Festplatten	10 Jahre
Magnetbänder	bis 30 Jahre

Probleme digitaler Medien

- schneller Medien- und Systemwechsel
 - veraltete Datenformate, veraltete Dateisysteme
- proprietäre Formate
- Abspielsoftware nicht mehr ausführbar, Hardware nicht mehr vorhanden

Probleme digitaler Medien

- Fülle von Datentypen und -formaten
 - (formatierter) Text
 - Daten, Datenbanken
 - Programme, Spiele
 - Schriften
 - Grafiken, Bilder, Fotos
 - Video
 - Audio - Tonaufnahme, Noten
 - Multimedia



exe
bat
com
a.out
java
WIN
MAC
UNIX
AMIGA
dmg
img
iso
gb
atr
nes
...

txt
rtf
doc
otf
pdf
html
wri
wps
ppt
xls
dot
css
js
ini
cfg
ttf
xml
...

iff
bmp
pcx
psd
gif
jpg
png
ico
ai
wsz
brush
tiff
bw
fif
raw
kdc
...

aiff
wav
mp3
ram
rm
au
flac
ogg
riff
tfmx
mod
midi
...

mov
avi
swf
fli
anim
mpg
mp2
rm
smac
dv
3gp
asf
tivo
...

„Multimedia“

- * keine eindeutige Definition
- * Wort des Jahres 1995
- * Steinmetz 1999: *„Ein Multimediasystem ist durch die rechnergestützte, integrierte Erzeugung, Manipulation, Darstellung, Speicherung und Kommunikation von unabhängigen Informationen gekennzeichnet, die in mindestens einem kontinuierlichen und einem diskreten Medium kodiert sind.“*(1)
- * Klimsa: *„... bedeutet "Multimedia" zahlreiche Hardware- und Softwaretechnologien für Integration von digitalen Medien, wie beispielsweise Text, Pixelbilder, Grafik, Video oder Ton. Neben diesem Mediaspekt – Multimedialität – spielen aber auch Interaktivität, Multitasking (gleichzeitige Ausführung mehrerer Prozesse) und Parallelität*

(1) STEINMETZ, R.: *MULTIMEDIA-TECHNOLOGIE*. SPRINGER VERLAG, 1990.

(2) ISSING, L.J.; KLIMSA, P. (HRSG.): *INFORMATION UND LERNEN MIT MULTIMEDIA UND INTERNET*. BELTZ PVU, 2002.

Speicherplatzbedarf Beispiel

PIXEL	RGB 3 BYTE	PAL INTERLACED 25 BILDER/SEK	
768 * 576	*	3	* 25
			≈ 31,64 MB/s

- * PAL Video unkomprimiert
 - * 1 Minute ≈ 1,85 GB
 - * 1 Stunde ≈ 111,23 GB
- * HDTV 2 Mpx (25 fps) unkomprimiert
 - * 1 Minute ≈ 8,69 GB
 - * 1 Stunde ≈ 521,42 GB

CD
~0,65 GB

DVD-R DL
~8,6 GB

BlueRay
~25 GB

LTO (U4)
~800 GB

Festplatte
~1500 GB

* ABWEICHUNGEN DES PLATZBEDARFES BEI WAHL ANDERER (HD-)FORMATE UND FARBSYSTEME

Kompressionstechniken

- * Redundanzkompression (verlustlos)

11111111000 --> 8*1 3*0

- * Lauflängenkodierung

- * Huffman-Kode, LZW-Kompression

- * Relevanzkompression (verlustbehaftet)

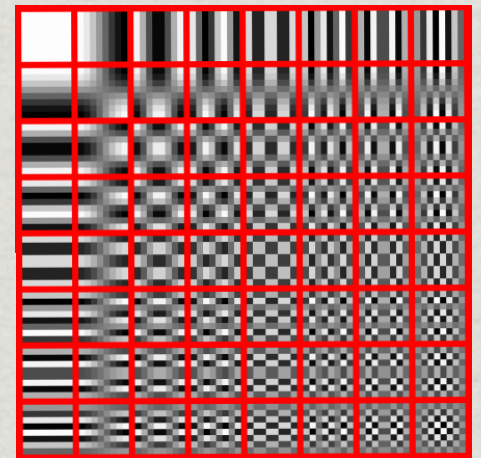
- * diskrete Kosinustransformation

- * YUV-Farbraum, Kontrastfilter

- * Bewegtbildkompression

- * Intraframe vs. Interframe (Differenzkodierung)

- * Erkennung von Bewegungsvektoren



Images

- * normalerweise exakte Kopie (Abbild) des Datenträgers
- * Rohdaten
- * enthält Informationen über Dateisystemsstruktur
- * direkter Bezug auf Hardwarestrukturen

Offene Formate, offene Standards

- * Bild: Pixel und Vektorbeschreibungen, RAW
- * Audio: Musiknotationen und Tonaufnahmen
- * verlustfrei
- * Quelloffen

Analoge Sicherung

- * Computerausdruck
- * Computer Output on Microfilm (COM)
- * aufwändig (einmalig), nur möglich für Text und Bilder
- * eingeschränkte Verfügbarkeit

(BILD AUS URHEBERRECHTLICHEN GRÜNDEN ENTFERNT)

Museumsansatz

- keine Migration, Bewahrung der Inhalte in Ihrer originalen Umgebung
- Versuch der Aufbewahrung
 - Originalhardware, -software
 - Lesegeräte
 - Ein- und Ausgabegeräte
- aufwändig, Zugang eingeschränkt
- keine Ersatzteile

Migration

- Umkodierung in aktuelle Speicherformate
 - Erhalt der Metadaten
 - verlustfreie vs. verlustbehaftete Konvertierung
 - sinnhafte Konvertierung
 - digitale Artefakte
 - Original wird verändert
 - manuelle Nachbearbeitung sehr aufwändig
 - bei multimedialen Daten nicht sinnvoll anwendbar
 - nicht möglich bei proprietären Formaten

Emulation

- aus dem lateinischen "aemulari", bedeutet nachahmen
- Emulation der Hardware und/oder Software (API)
- zuerst 1962 bei IBM
- originalgetreu, keine Konvertierung notwendig
- erfolgreich im Spielebereich
- Emulator ist selbst Software, muss angepasst werden

Workshop

- * Vielen Dank für die Aufmerksamkeit!
- * Jens-Martin Loebel
- * loebel@informatik.hu-berlin.de
- * <http://waste.informatik.hu-berlin.de/jml/>