Möglichkeiten und Gefahren
der Digitalisierung des visuellen Kulturguts

Jürgen Keiper
Deutsche Kinemathek – Museum für Film und Fernsehen

Herbsttagung der Fachgruppe Dokumentation im Deutschen Museumsbund 2010

Re: [AMIA-L] good practice for archival video formats?	08.09.2010	Ari Davidow	5.5 k
Re: [AMIA-L] good practice for archival video formats?	08.09.2010	C. Park Seward	6.9 k
Re: [AMIA-L] good practice for archival video formats?	08.09.2010	Isaiah Beard	6 k
Re: [AMIA-L] good practice for archival video formats?	08.09.2010	Ari Davidow	7 k
Re: [AMIA-L] good practice for archival video formats?	08.09.2010	andre roy	10 k
Re: [AMIA-L] good practice for archival video formats?	08.09.2010	David C. Crosthwait	5.4 k

[AMIA-L] good practice for archival vide... - Mozilla Firefox

195.3 h

[AMIA-L] good practice for archival video formats?

Von:	Association of Moving Image Archivists
im Auftrag von:	Ari Davidow
Datum:	08.09.2010 - 08:23:20
An:	AMIA-L@LSV.UKY.EDU

We have a lot of video, original created using a miniDV camera, that
we downloaded to .avi files and are archiving in that format.

We recently did a project where the videographer, using a couple of
the new HD cameras (sony Z7U, Z1U) and working on Mac hardware is
getting the downloads in .mov format. She is asking whether she should
save those raw .mov files for our archive, or process them into
"ProRes" files (still a .mov format?) or what? Conversion to .avi
sounds either unfamiliar or potentially just taking the time to
exchange one wrapper format for another.

This is a time when HD formats are pretty up in the air. My question
is how we want to store this video so that 10 years from now we can
re-edit it and/or regenerate our web files. What are current best
practices or thoughts?

Many thanks,

- Das Medium
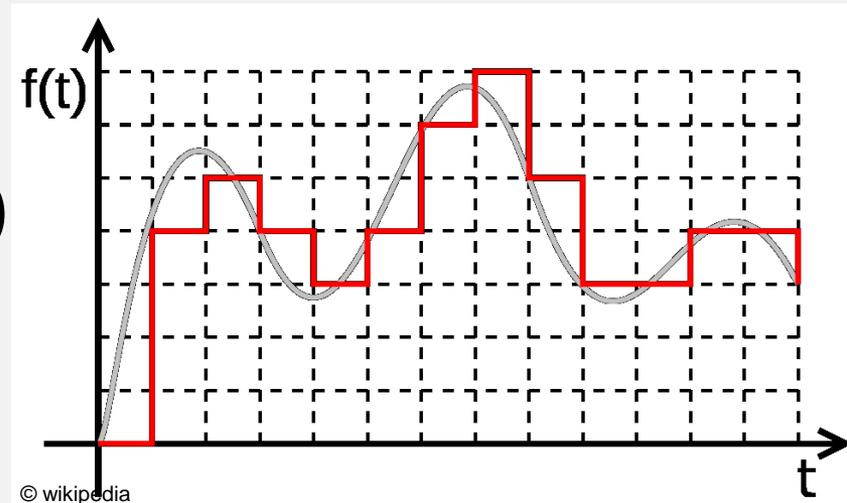- Die Produzenten
- Das Archiv

ANALOG > DIGITAL

Highres
- Methode: Telecine <> Abtastung (Scanner)
- Nachbearbeitung (Farbmanagement, Artefakte, Postproduktion, Restaurierung)
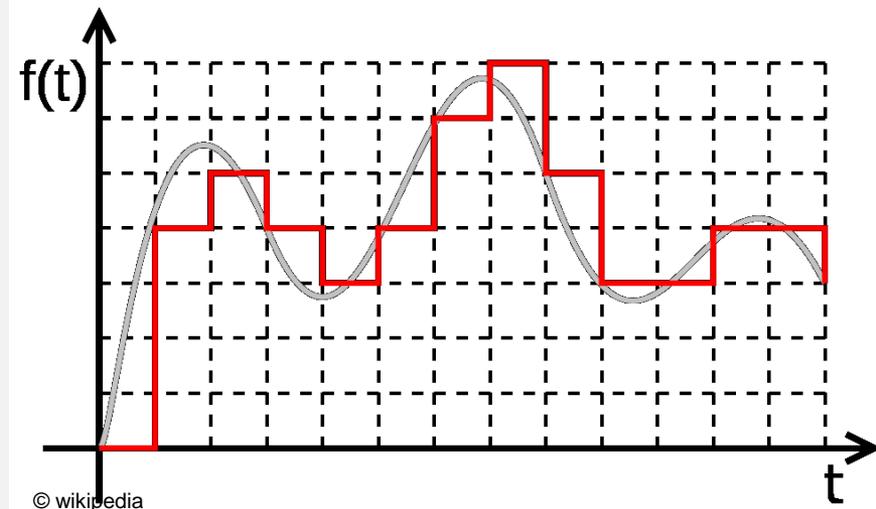- Zugriff
- Speicherung

Lowres
- Videokarten (VHS>File)

© wikipedia

Das digitale Medium

- Träger: Tape, optisch, HD, SSD, Flash …
- Ordnerorganisation (DVD, BD)
- Dateiorganisation: Containerformate wie AVI, Matroska, MPEG 4, MXF…
- Codecs: MPEG-4 (Part 2, Part 10 (AVC/H.264), VP8, DNxHD codec, Motion JPEG 2000…
- DRM

z. B. AVI-Datei

Kopfdaten

Video	Audio

© wikipedia

z. B. Datei mit Matroska-Container

Kopfdaten

franz. Untertitel

Video	Audio deutsch	Audio engl.

z. B. PDF-Datei

Kopfdaten
eingebettete Schriftarten
Text und Markup
Bild
Formular
Text und Markup

$f(t)$

t

© wikipedia
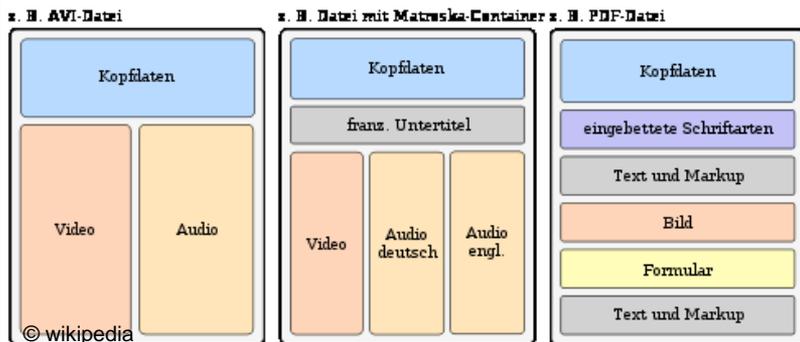
Das digitale Medium

- Formatmigration
- Redundanzkonzept und Konsistenzprüfung
- Interoperabilität
- Langzeitarchivierung

© wikipedia
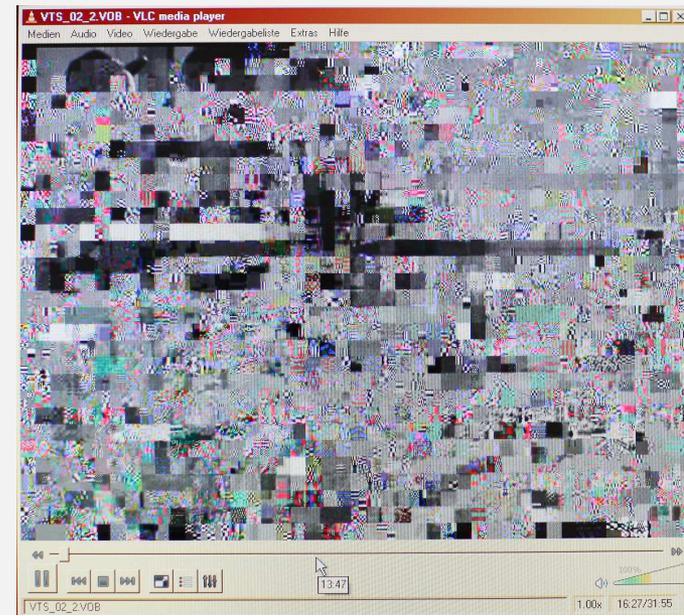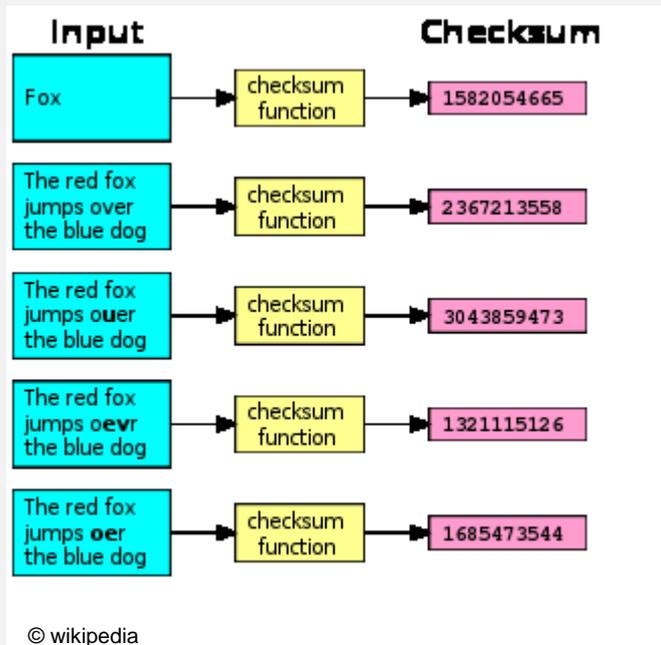
Das digitale Medium

Ingest Format	Migration format	Notes
Low quality media		
VHS tape	DVD	**Access** Perfectly adequate for VHS playback
VHS tape	MPEG-4 files	**Access** Adequate for quality. Minimum data rates (MPEG-4): 500k b/s. There are MANY potential access formats, and they come and go.
VHS tape	DV files	**Archive** (temporary) 25 M b/s, 12 GB/hr. Migrate to lossless for preservation.
'low end' digital files	Save as is, AND save as DV or lossless	**Archive** (temporary) Before format or DV format becomes obsolete, migrate to lossless for preservation.
DVD	DV files	**Archive** (temporary) 25 M b/s, 12 GB/hr. Migrate to lossless for preservation.
Medium quality		
U-Matic	DVD	**Access** Reduces quality; suitable only for viewing.
U-Matic	DV files	**Archive** (temporary) 25 M b/s, 12 GB/hr. Migrate to lossless for pre-servation.
DV, DVCAM	DV files (meaning .avi files with native coding)	**Archive** (temporary) transfers to computer at 25 M b/s, resulting in an .avi file 'clone' of the original DV tape. Migrate to lossless for pre-servation.
High Quality		
BetaSP, Digibeta, other pro formats	Uncompressed	**Archive** Uncompressed standard definition video: 200 Mb/s. About 100 GB for one hour, meaning 25 DVD-ROMs (or part of one data tape or hard drive).
BetaSP, Digibeta, other pro formats	Motion JPEG 2000 (lossless version)	**Archive** lossless compression, with a resultant data rate of around 90M b/s. About 40 GB for one hour.
DVCPRO50	.avi files, DV coding	**Archive** (temporary) As for DV, but at twice the data rate. Less suscep-tible to loss on future migrations. Migrate to lossless for preservation.
'High end' digital files	Save as is	**Archive** (temporary) Before format becomes obsolete, migrate to lossless for preservation.

Table 4: Video Migration Roadmap.

Das digitale Medium

Codecs und Komprimierung

interframe compression	H.264/MPEG-4 AVC, libtheora (Theora), VP8	
intraframe compression	DV (DCT), Motion JPEG 2000, (H.264/MPEG-4 AVC all-intra), DNxHD	Motion JPEG 2000 LL (MJ2), Lagarith
	lossy	**lossless**

Das digitale Medium

Codecs:

General video codec information — creator/company, license/price, etc.

Codec	Creator/Maintainer	First public release date	Latest stable version	License	Patented compression formats	Compression method
libtheora (Theora)	Xiph.org	2002-09-25	1.1.1 (2009)[3]	BSD-style[4]	none[*]	Lossy
dirac-research (Dirac)	BBC Research Department	2008-09-17	1.0.2 (2009)[5]	MPL 1.1, GNU GPL 2, GNU LGPL 2.1	none	Lossy/lossless
Schrödinger (Dirac)	David Schleef	2008-02-22	1.0.9 (2009)[5]	MPL 1.1, GNU GPL 2, GNU LGPL 2, MIT License	none	Lossy/lossless
x264	x264 team	2003	r1376 (2009)[6][7]	GNU GPL	MPEG-4 AVC/H.264	Lossy/lossless
Xvid	Xvid team	2001	1.2.2 (2009)[8]	GNU GPL	MPEG-4 ASP	Lossy
FFmpeg (libavcodec)	FFmpeg team	2000	0.5.1 (2010)[9]	GNU LGPL	MPEG-1, MPEG-2, MPEG-4 ASP, H.261, H.263, VC-3, WMV7, WMV8, VP5, VP6, MJPEG etc.	Lossy/lossless
FFavs (libavcodec)	FFavs team	2009	0.0.3[10]	GNU LGPL	MPEG-1, MPEG-2, MPEG-4 ASP etc.	Lossy/lossless
Blackbird	Forbidden Technologies plc	2006-01	2	Proprietary	Blackbird	Lossy
DivX	DivX, Inc.	2001	7.0 (2009)	Proprietary	MPEG-4 ASP, H.264	Lossy
DivX ;-)	a hack of Microsoft's MPEG-4v3 codec[11][12]	1998	3.20 alpha[13] (2000)	Proprietary	Microsoft's MPEG-4v3 (not MPEG-4 compliant)	Lossy
3ivx	3ivx Technologies Pty. Ltd.	2001	5.0.2 (2007)[14]	Proprietary	MPEG-4 ASP	Lossy
Nero Digital	Nero AG	2003	Unknown	Proprietary	MPEG-4 ASP, H.264[15]	Lossy
ProRes 422 / ProRes 4444	Apple Inc.	2007		Proprietary	Unknown	Lossy
Sorenson Video	Sorenson Media	1998		Proprietary	Sorenson Video	Lossy
Sorenson Spark	Sorenson Media	2002		Proprietary	Sorenson Spark	Lossy
VP3	On2 Technologies	2000		BSD-style[4]	none[*]	Lossy
VP4	On2 Technologies	2001		Proprietary	VP4	Lossy
VP5	On2 Technologies	2002		Proprietary	VP5	Lossy
VP6	On2 Technologies	2003		Proprietary	VP6	Lossy
VP7	On2 Technologies	2005		Proprietary	VP7	Lossy
VP8	On2 Technologies (Now bought by Google)	2008		BSD-style	VP8	Lossy
DNxHD	Avid Technology	2008		Proprietary	VC-3	Lossy
Cinema Craft Encoder SP2	Custom Technology Corporation	2000	1.00.01.09 (2009)[16]	Proprietary	MPEG-1, MPEG-2	Lossy
TMPGEnc Free Version	Pegasys Inc.	2001	2.525.64.184 (2008)[17]	Proprietary	MPEG-1, MPEG-2	Lossy
Windows Media Encoder	Microsoft	1999	9 (2003) (WMV3 in FourCC)	Proprietary	WMV, VC-1, (in early versions MPEG-4 Part 2 and not MPEG-4 compliant MPEG-4v3, MPEG-4v2)	Lossy
Cinepak	Created by SuperMac, Inc. Currently maintained by Compression Technologies, Inc.[18]	1991	1.10.0.26 (1999)	Proprietary	Unknown	Lossy
Indeo Video	Intel Corporation, currently offered by Ligos Corporation	1992	5.2	Proprietary	Indeo Video	Lossy
TrueMotion S	The Duck Corporation	1995		Proprietary	TrueMotion S	Lossy
RealVideo	RealNetworks	1997	RealVideo 10[19]	Proprietary	H.263, RealVideo	Lossy
ACT-L3	Streambox	Template:02-21-03	Template:3.4 (2009)	Proprietary	Unknown	Lossy
Huffyuv	Ben Rudiak-Gould	2000	2.1.1 (2003)[20]	GNU GPL 2	none	Lossless
Lagarith	Ben Greenwood	2004-10-04	1.3.20 (2009)[21]	GNU GPL 2	none	Lossless

Die Produzenten

- Ingest aus dem
  - Home-Bereich (Consumer-Formate)
  - Broadcast-Bereich
  - Filmproduktionsbereich

Beispiele:
Digitales Kino
(2048 × 1080 Bildpunkte, 24B/s

HD-Fernsehen
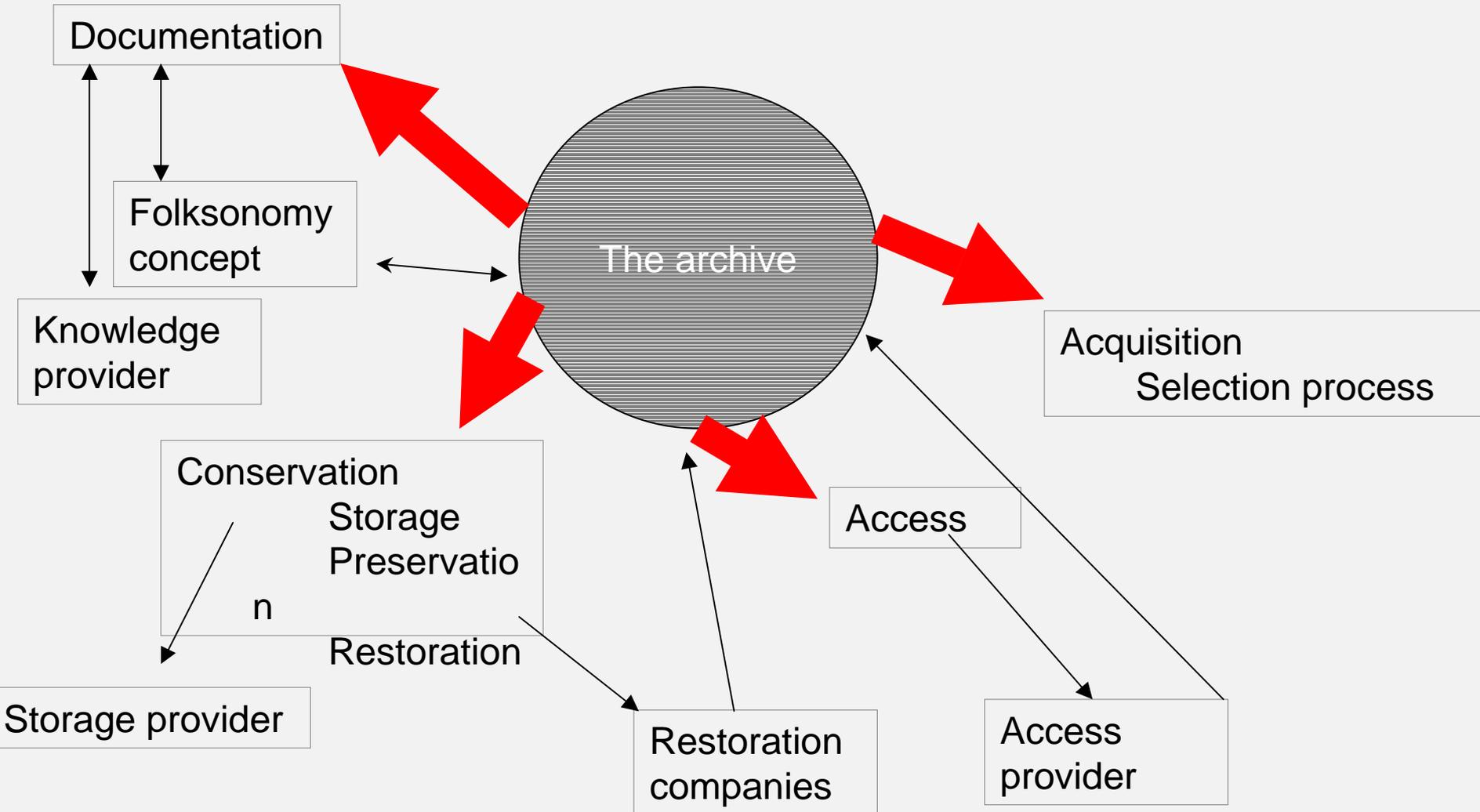1.920 × 1.080 Bildpunkte, 25B/s

Die Produzenten

© wikipedia

- Hochproprietäre Formate (Sony XDCAM>PFD50-Disk) und Lesegeräte; inzwischen schon „abgelöst" durch Sony XDCAM EX Speicherchips)
- Extrem hohe Innovationszyklen von Hard- und Software
- Mangelnde Interoperabilität
- Hohe technische Expertise notwendig

Die Archive

Strukturwandel der Archivlandschaft

Die Archive

Storage-Provider

Beispiel: CO2-Löschanlage im Rechen-zentrum des des Konrad Zuse-Instituts, Berlin

Die Archive

Storage Provider:

Konrad Zuse-Institut
LOC
Amazon:
Simple Storage Service (S3)
MozyPro

Taperoboter im Rechenzentrum
des des Konrad Zuse-Instituts, Berlin

© Jürgen Keiper

© Jürgen Keiper

Die Archive

- Frage des Verlustes rechtlich völlig ungeklärt, Versicherungen sind nicht möglich

USV-Anlage im Rechenzentrum des des Konrad Zuse-Instituts, Berlin

Die Archive

Wahlmöglichkeiten für die Archive:
- Adaption von heterogenen, temporären Industrielösungen
- Vorgaben gemäß eigenen Kompetenzen
- „Flaschenpost"-Strategie