

Eine **ARCHE** zur Rettung digitalen Kulturguts

Dr. Klaus Wendel · Wladimir Schwitin

Schlußbericht zum Verbundprojekt ARCHE „Entwicklung eines Farbmikrofilm-Laserbelichters zur Langzeitarchivierung digitaler bzw. digitalisierter Dokumente“, Teilvorhaben „Entwicklung eines durchgehenden Workflow für die Erstellung von Farbmikrofilmen“

Projektpartner des Teilvorhabens:

Universitätsbibliothek Stuttgart · Landesarchiv Baden-Württemberg

Laufzeit : März 2004 bis September 2006

Förderprogramm: InnoNet des BMWi

Förderkennzeichen: 16IN0271

Ausgabe: 13. März 2007

Das diesem Bericht zugrundeliegende Vorhaben wurde mit Mitteln des BMWi (16IN0271) gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autoren.

Kurzfassung des Schlußberichts

Das Verbundprojekt ARCHE war ein durch das Fraunhofer-Institut für Physikalische Messtechnik IPM in Freiburg geleitetes Gemeinschaftsunternehmen der Universitätsbibliothek Stuttgart, des Landesarchivs Baden-Württemberg, der Firma MicroArchive Systems GmbH und weiteren Industriepartnern. Es wurde vom Bundesministerium für Wirtschaft und Arbeit über das Programm „Förderung von innovativen Netzwerken“ - InnoNet gefördert.

Im Rahmen des ARCHE-Projektes wurde vom Fraunhofer Institut für Physikalische Messtechnik ein Laserbelichter entwickelt, welcher es erlaubt, digitale Bilddaten mit sehr hoher Genauigkeit und Farbtreue auf alterungsbeständigem Farbmikrofilm zu speichern.

Zielstellung des Teilprojektes der Universitätsbibliothek Stuttgart und des Landesarchivs Baden-Württemberg war die Entwicklung eines Workflow und der dazugehörigen technischen Lösung für die Ausbelichtung digitaler Dokumente auf Farbmikrofilm. Die Daten werden dabei in einer für den Mikrofilm optimierten Form aufbereitet, so daß zukünftige Redigitalisierungsprozesse möglichst automatisiert und fehlerfrei die analog gespeicherten Informationen zurücklesen können. Unterstützt werden dadurch Fachaufgaben wie die Bestandserhaltung, die verbesserte Nutzung durch retrospektive Digitalisierung und die Langzeitarchivierung digitaler Dokumente durch Konversion. Es wurde der Prototyp einer Software entwickelt, welche digitale Bild- und Metadaten in einem analogen Format kombiniert.

Zusätzlich leistete das Stuttgarter ARCHE-Team mit umfangreichen Digitalisierungsarbeiten einen wichtigen Beitrag zur Informationssicherung. In den Räumen des Ludwigsburger Instituts für Erhaltung wurden ca. 1,2 TByte Daten aus ca. 6000 Vorlagen gewonnen, welche von Einrichtungen des Landesarchivs (Hauptstaatsarchiv Stuttgart; Staatsarchiv Ludwigsburg; Generallandesarchiv Karlsruhe) und Einrichtungen der Universität Stuttgart (Universitätsbibliothek; Universitätsarchiv; Institut für Leichtbau, Entwerfen und Konstruieren) zur Verfügung gestellt wurden. Anschließend wurden die Daten erfolgreich auf Ilfachrome-Farbmikrofilm gesichert.

Abstract of the final report

The group project ARCHE („Arche“ [germ.] = „ark“) was a joint venture led by the Fraunhofer Institut fuer Physikalische Messtechnik (institute for physical measuring techniques) or physical measuring technique IPM at Freiburg, consisting of the Universitaetsbibliothek (university library) Stuttgart, the Landesarchiv (state archive) Baden-Wuerttemberg, the company MicroArchive Systems GmbH and further industrial partners. It was sponsored by the Federal Ministry for economics and labour by means of the program „promotion of innovative networks“ - InnoNet.

In the context of the ARCHE project the Fraunhofer Institut fuer Physikalische Messtechnik IPM at Freiburg has developed a laser microfilm recorder, enabling the storage of digital images with very high accuracy and colour fidelity on age-resisting coloured microfilm.

Primarily, the subproject of the Universitaetsbibliothek Stuttgart and the Landesarchiv Baden-Wuerttemberg aimed at the elaboration of a workflow and development of necessary software tools supporting computer output on colour microfilms. The software (prototype) combines digital images and its meta data in one common analogue format. For this purpose the data are prepared in a form optimized for microfilms, so that future redigitization processes can read backwards the stored information as automated as possible and error free similar to stored information. Thereby specialized tasks are supported like inventory maintenance, improved use by retrospective digitization and long-term archiving of digital documents by conversion.

Additionally by extensive digitization work the Stuttgart ARCHE team made an important contribution to the information safety device. The project got assistance by the Institut fuer Erhaltung (institute for preservation) at Ludwigsburg while scanning approx. 1.2 TByte data from approx. 6.000 sheets in different sizes, owned by institutes of the Landesarchiv Baden-Wuerttemberg (Hauptstaatsarchiv Stuttgart; Staatsarchiv Ludwigsburg; Generallandesarchiv Karlsruhe) and institutions of the Universitaet Stuttgart (Universitaetsbibliothek (library); Universitaetsarchiv (archive); Institut fuer Leichtbau, Entwerfen und Konstruieren ILEK (institute for design and construction)). Finally the data were successfully stored on Ilfachrome colour microfilm.

Inhaltsverzeichnis

1	Ansprechpartner	7
2	Zielsetzungen	8
2.1	Grundproblem bei der Langzeitarchivierung digitaler Daten	8
2.2	Der Arche-Ansatz: Konversion	11
2.2.1	Das Medium	11
2.2.2	Der ArchiveLaser [®]	11
2.2.3	Das Prinzip der Konversion – der ARCHE-Workflow	12
3	Die theoretischen Grundlagen zur ARCHE-Bildaufbereitung	17
3.1	Randbedingungen	17
3.2	Vorgehensweise für Archive und Bibliotheken	17
3.3	Anforderungen an eine Textdatei	18
3.4	Die Struktur der CSV-Datei	18
3.4.1	Regeln für Spaltennamen	19
3.4.2	Regeln für Spaltenarten	20
3.5	Vorbereitung der Bilddaten	25
3.5.1	Ideale Spezifikationen der Bildfiles	25
3.5.2	Variante 1: Auflösung der Signatur im Dateinamen	25
3.5.3	Variante 2: Auflösung der Signatur durch den Verzeichnisbaum	26
3.5.4	Sonderfall: Individuelle Metadatenergänzung zum Bild	28
3.6	Anordnung der Dateien auf dem Datenträger	28
4	Aufbereitung eines Bestandes am Beispiel eines Musterfilms	31
4.1	Ausgabe der Metadaten	31
4.1.1	meta.txt	31
4.1.2	meta.csv	32
4.2	Ausgabe der Bilddaten	34
4.2.1	Eigenschaften der „Frames“ zur Aufnahme von Bilddaten	34
4.2.2	Variante 1: Ausgabe ohne „Nesting“ (Bildhäufung)	36
4.2.3	Variante 2: Ausgabe mit „Nesting“ (Bildhäufung)	37
4.2.4	Schematische Anordnung eines ARCHE-Films	38
4.2.5	Ein „Musterfilm“ aus praktischen Beispielen	39
4.2.6	Weitere darstellbare Informationstypen	40

5	Digitalisierung von hochwertigen Testvorlagen	41
5.1	Unterschiedliche Verfahrensweisen	42
5.1.1	Der Zeilenscanner	42
5.1.2	Das Shot-System	43
5.1.3	Der Zeilenscanner in eine Kamera integriert	44
5.2	Rauschverhalten	44
5.3	Auflösung	44
5.4	Farbtiefe	45
5.5	Beleuchtung	46
5.5.1	Lichteinwirkung auf die Vorlage	46
5.5.2	Lichtverteilung	46
5.6	Scandauer	46
5.7	Ausstattung, Zubehör, Durchlichteinheit	47
5.8	Scanner-Software	47
5.9	Calibrierung, Farbmanagement, Wartung	48
5.10	Schlußfolgerungen	49
6	Der Warenkorb an Testvorlagen	50
6.1	Staatsarchiv Ludwigsburg	50
6.2	Generallandesarchiv Karlsruhe	51
6.3	Hauptstaatsarchiv Stuttgart	52
6.4	Universitätsbibliothek Stuttgart	52
6.5	Universitätsarchiv Stuttgart	53
6.6	Institut für Leichtbau, Entwerfen und Konstruieren der Universität Stuttgart (ILEK)	54
7	Entwicklungsschritte der Bildaufbereitungssoftware	55
7.1	ArcheShell	55
7.1.1	Funktionsweise der ArcheShell	56
7.1.2	Primäre Befehle der ArcheShell	56
7.1.3	Sekundäre Befehle	57
7.2	ArcheGui	58
7.3	Metainput	59
7.4	ArcheFilmcreator	59
7.4.1	Funktionsweise des ArcheFilmcreators	59
8	Zeitplan des Teilprojekts	61
	Literaturverzeichnis	67

1 Ansprechpartner

Teilprojektleitung

Dr. Gerhard Lehrmann
Universitätsbibliothek Stuttgart
Postfach 10 49 41
D-70043 Stuttgart
Tel. 0711/685-83520
gerhard.lehrmann@ub.uni-stuttgart.de

Dr. Gerald Maier
Landesarchiv Baden-Württemberg
Eugenstr. 7
D-70182 Stuttgart
Tel. 0711/212-4279
gerald.maier@la-bw.de

Mitarbeiter

Technische Realisierung:
Wladimir Schwitin
wladimir.schwitin@ub.uni-stuttgart.de

Inhaltliche Betreuung:
Dr. Klaus Wendel
klaus.wendel@ub.uni-stuttgart.de

Universitätsbibliothek Stuttgart
Postfach 10 49 41
D-70043 Stuttgart
Tel. 0711/685-82536

Gesamtprojektleitung

Andreas Hofmann
Fraunhofer Institut für Physikalische
Messtechnik (IPM)
Heidenhofstr. 8
D-79110 Freiburg
Tel. 0761/8857-136
Fax: 0761/8857-224
andreas.hofmann@ipm.fraunhofer.de

2 Zielsetzungen

2.1 Grundproblem bei der Langzeitarchivierung digitaler Daten

Zum besseren Verständnis muß man sich vergegenwärtigen, woraus sich eine Information zusammensetzt und welchen Übertragungsweg sie nimmt. Im Kern ist eine Information ein Gedanke. Das einfachste Medium bzw. den ersten Datenträger, den der Gedanke zur Verbreitung nehmen kann, ist die Sprache. Die Information wird also gedacht, gesprochen, gehört, interpretiert und weitergegeben. Dieses Medium ist jedoch sehr flüchtig und zur Verbreitung auf interpretierende Zwischenstationen angewiesen.

Als alternatives Medium bietet sich die Schrift an. Auf haltbaren Datenträgern kann die Information so fixiert werden, daß sie über einen längeren Zeitraum lesbar bleibt und zur Verbreitung keine (bzw. sehr viel seltenere) Interpretationzwischenritte durchlaufen muß. Im (allerdings unwahrscheinlichen) Idealfall würde die Information *einmal* fixiert und auch beim Kopieren nicht verändert.

Der digitale Datenträger bietet nun den ganz großen Vorteil, daß bei der Verbreitung nicht die gelesene und interpretierte Information, sondern das Medium selbst kopiert wird. Zudem geschieht der Kopiervorgang verifizierbar fehlerfrei. Bleibt also das digitale Medium, also die digitale Datenstruktur, erhalten und lesbar, steht auch die Information in ihrer Entstehungsform zur Verfügung. Doch genau hier liegt auch das Problem: Der Datenträger muß erhalten **und** die digitale Struktur lesbar bleiben, und beides ist derzeit nicht gewährleistet.

Derzeit werden zwei Hauptansätze zur Langzeitarchivierung digitaler Bestände verfolgt:

Migration Wandlung der Daten einer alten Technologie zur Nutzung mit einer neuen Technologie.

Vorteil: Die Arbeitsumgebung bleibt aktuell.

Nachteil: Teil-Datenverluste sind bei Migrationsschritten wahrscheinlich; können durch häufige kleine Migrationsschritte zwar minimiert, jedoch nicht völlig ausgeschlossen werden können.

Emulation Nutzung von Daten einer alten Technologie innerhalb einer neuen Technologie durch Schaffung einer künstlichen „alten“ Arbeitsumgebung.

Vorteil: Daten bleiben unverändert und authentisch.

Nachteil: Der Emulator muß seinerseits ständig auf die sich ständig verändernde Hardware migriert werden.

Es besteht also ein Gegensatz zwischen dem *Wert der digitalen Datenstruktur* und dem *Wert der in den digitalen Daten gebundenen Information*. Solange nämlich ein Medium in seiner Existenz verhältnismäßig sicher ist, ist seine Information jederzeit leicht aus dem Originalmedium zu rekonstruieren und entsprechend wenig kostbar. Mit steigendem Aufwand bzw. steigenden Kosten für die Migration des digitalen Masters nimmt die Bestandssicherheit hierfür ab, womit dem möglichst authentischen Erhalt der Information eine steigende Bedeutung zukommt.

Letztlich können drei Ereignisse für eine Unterbrechung der Migrationskette ursächlich sein:

1. Verlust der physikalischen Eigenschaften des Mediums.
2. Verlust der Lesbarkeit des Mediums.
3. Eine (unter Umständen sogar nur kurze) Unterbrechung des notwendigen technologisch hochangeregten Zustandes, der zur Speicherung und Auswertung digitaler Daten notwendig ist.

Führt die Unterbrechung der Migrationskette bei analogen Medien lediglich zum Verlust der Interpretationsfähigkeit (z.B. nicht entschlüsselte Schriftzeichen einer vergangenen Kultur), so hat eine Unterbrechung der Migrationskette bei digitalen Daten, den Verlust der Lesefähigkeit zur Folge, was einem Totalverlust der Information gleichkommt.

Und selbst unter rein wirtschaftlichen Gesichtspunkten haben Emulation und Migration zwei weitere Probleme:

1. Je älter ein Datenformat ist und je mehr Migrationszwischenschritte zu absolvieren sein sollten, desto höher ist der Aufwand, die wachsende Anzahl von Rücklesemechanismen zu überschauen. Analog dazu nimmt mit der zunehmenden und sich wandelnden Zahl von Systemumgebungen auch der Aufwand zu, stets funktionsfähige Emulatoren bereitzuhalten. Mit den wachsenden Anforderungen aber steigen die Kosten.
2. Je komplexer sich das (Rück-)lesen und der Erhalt digitaler Daten gestaltet, desto schwieriger ist es heute, für morgen eine gesicherte Prognose über die technischen Möglichkeiten zu geben, die uns ein Zurücklesen gestatten sollen. Wir können zwar vermuten und hoffen, daß der notwendige technologisch hochangeregte Zustand bei all seiner Komplexität auch für die Zukunft aufrecht erhalten werden kann. Eine Garantie können wir dafür nicht geben. Die Möglichkeiten der Datenmigration sind damit schon für die nahe Zukunft unüberschaubar und der Erhalt der Daten damit nicht kalkulierbar.

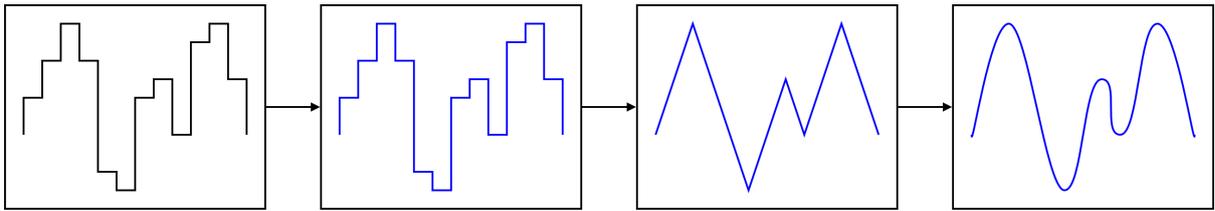


Abbildung 2.1: (1) Digitale Ausgangsdaten (2) werden maschinell gelesen, (3) analog visualisiert (bzw. „gesehen“) und (4) intellektuell interpretiert (bzw. „verstanden“).

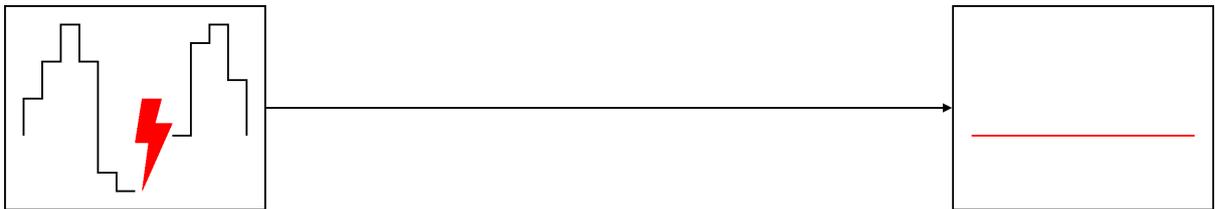


Abbildung 2.2: (1) Fehlerhafte digitale Ausgangsdaten (2) können maschinell nicht gelesen werden. Somit sind auch die Visualisierung und die Interpretation unterbunden.
→ Die Information ist vollständig verloren.

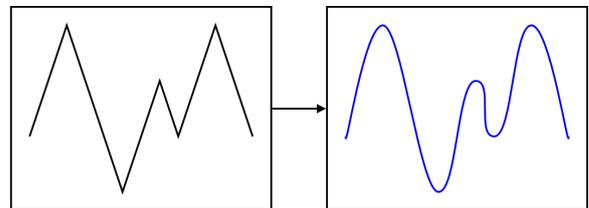


Abbildung 2.3: (1) Analoge Ausgangsdaten können mit weitaus geringerem oder gar keinem technischen Aufwand visualisiert werden und (2) werden anschließend intellektuell interpretiert.

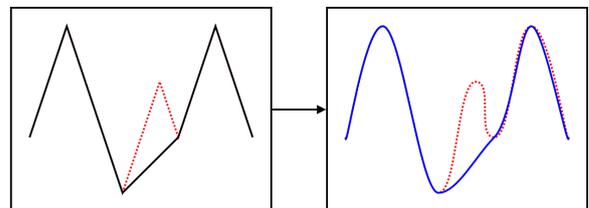


Abbildung 2.4: (1) Fehlerhafte analoge Ausgangsdaten können mit Ausnahme des fehlerhaften Bereichs weiterhin visualisiert werden (2) und werden anschließend als Restinformation intellektuell interpretiert.

2.2 Der Arche-Ansatz: Konversion

2.2.1 Das Medium

Jedes Medium hat eine begrenzte Haltbarkeit und jede darauf gespeicherte Information muß früher oder später auf ein neues Medium übertragen werden. Um die daraus resultierenden Konsequenzen (Datenverlust, Informationsverfälschung) so gering und so selten wie möglich tragen zu müssen, muß es das Ziel sein, die Migrationsintervalle so groß wie möglich zu machen. Entscheidend hierfür ist das Medium:

1. Es muß maschinell beschreib- und lesbar sein.
2. Der Aufwand, es zu beschreiben darf durchaus hoch sein, es muß aber auch mit möglichst einfachen Werkzeugen ausgelesen werden können.
3. Es muß bei vergleichbarem Raumbedarf eine so hohe Informationsdichte aufnehmen können, daß die Inhalte gängiger digitaler Medien gespeichert werden können.
4. Es sollte über eine Haltbarkeit verfügen, die mit mindestens der Lebensdauer säurefreier Buch und Papierbestände entspricht.
5. Es sollte kostengünstig herstellbar sein.

All diese Anforderungen erfüllt der Farbmikrofilm Ilfochrome[®] (Cibachrome) der Schweizer Firma Ilford. Es handelt sich dabei um einen langzeitstabilen Film mit sehr geringer Körnigkeit. Seine Ultra Stablen DIAZO-Farbstoffe sind bereits bei der Herstellung in der Emulsion vorhanden. Er ist außergewöhnlich temperaturbeständig und es wird eine Lagerzeit von über 500 Jahren erwartet. Allerdings ist der Film mit <1 ASA sehr langsam und benötigt sehr viel Licht bei der Aufnahme.

Die technischen Anforderung an das Zurücklesen eines (Mikro-)films sind denkbar gering. Es genügt eine einfache optische Aparatur. Selbst mit einer Lupe und einer Kerze wären bereits interpretierbare Informationen zu gewinnen.

Bei der Speicherung digitaler Daten auf analogem Film ist jedoch ein technisch aufwendigerer Zwischenritt notwendig.

2.2.2 Der ArchiveLaser[®]

Das Fraunhofer Institut für Physikalische Meßtechnik in Freiburg hat auf der technologischen Grundlage vorheriger Projekte einen Farbmikrofilmbelichter für sehr hohe Auflösungen entwickelt. Er wird unter dem Namen ArchiveLaser[®] von der Firma MicroArchive Systems GmbH¹ als Vertriebspartner angeboten.

¹ [8]

Anders als bei der herkömmlichen *analogen* Fotografie, die ein durch ein Objektiv verkleinertes Bild auf den Film belichtet, hat der in Freiburg entwickelte Laserbelichter den ganz erheblichen Vorzug, durch Nutzung von Lasern mit höchster Strahlqualität keine Abbildungsfehler akzeptieren zu müssen, die sich sonst selbst bei den besten Optiken einstellen. Die direkte Ansteuerung der Laser ermöglicht ein sehr konstantes und höchst genaues Farbmanagement mit 36 Bit Farbtiefe. Insgesamt drei den Farbkanälen rot, grün und blau entsprechend getaktete einfarbige Laserimpulse werden hier zu einem bunten Lichtstrahl gebündelt und als Bildpunkt direkt auf den Film geschrieben. Damit wurde eine Pixelgröße von 3 μm möglich, was einer Auflösung von 160 Linienpaaren / mm entspricht. Ein schneller Drehscanner bewegt dabei die Laserstrahlen senkrecht zum Film und belichtet jeweils eine 10666 Laserpunkte breite Zeile. Ein Bild besteht aus 15000 Zeilen, die in ca. 40 Sekunden geschrieben werden. Die Toleranz im Zeilenabstand beträgt max. 10 nm. Gemessen an den 32x45 mm großen „Frames“ eines perforationslosen 35 mm Films erreicht man so eine Auflösung für *Digitalbilder* von 10666x15000 Bildpunkten auf Ilfochrome[®] Mikrofilm.²

2.2.3 Das Prinzip der Konversion – der ARCHE-Workflow

ARCHE sieht also vor, die optische reproduzierbare Information auf ein haltbares Medium zu übertragen – zu „konvertieren“. Dabei ist zu berücksichtigen, daß sich eine Information zumeist aus zwei Teilen zusammensetzt: Der Bildinformation und den die Information beschreibenden *Metadaten*. Natürlich sollen beide Informationsteile bei der Konversion übernommen werden. Hier zeigt sich bereits der erste große Vorteil dieses Ansatzes. Müssen bei der Speicherung digitaler Daten die beiden Informationsteile häufig getrennt gespeichert (z.B. Bilddaten in Dateisystem, Metadaten in Datenbank) und auch getrennt migriert werden, so werden bei ARCHE beide Informationen auf einem einzigen Medium vereint, das zudem mit den gleichen Mitteln gelesen werden kann.

Schematisch sieht ARCHE zunächst folgenden Workflow vor (siehe [Abbildung 2.6](#)):

1. Gewinnung der Ausgangsdaten:
 - a) Bilddaten werden digitalisiert bzw. liegen bereits digital vor. Sie werden in ein einheitliches Bildformat gebracht.
 - b) Metadaten werden in ein einheitliches Datenbankformat gebracht. Metadaten sind nicht genau spezifiziert. Bei ihnen kann es sich um nahezu jedes Datenformat handeln. Vom Zettelkasten über das Findbuch bis hin zur Datenbank.
2. Bilddaten und Metadaten werden in einem gemeinsamen Format vereinigt. Dazu muß die Textinformation (Metadaten) visualisiert werden.
3. Speicherung der kombinierten Daten auf Archivierungsmedium:

² Siehe auch [9] und [7]

Funktionsweise des ArchivLasers

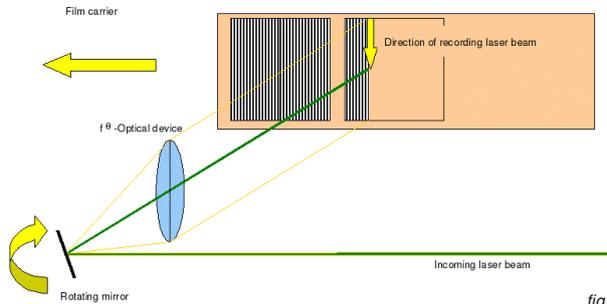
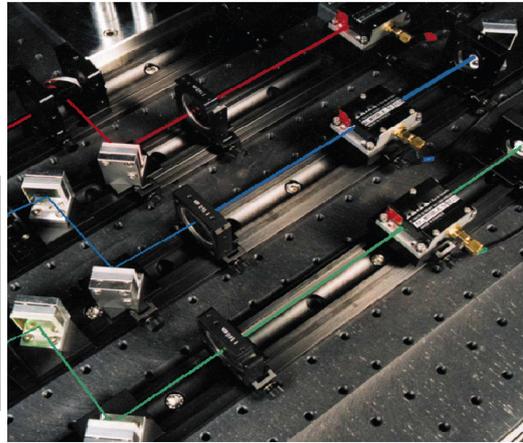
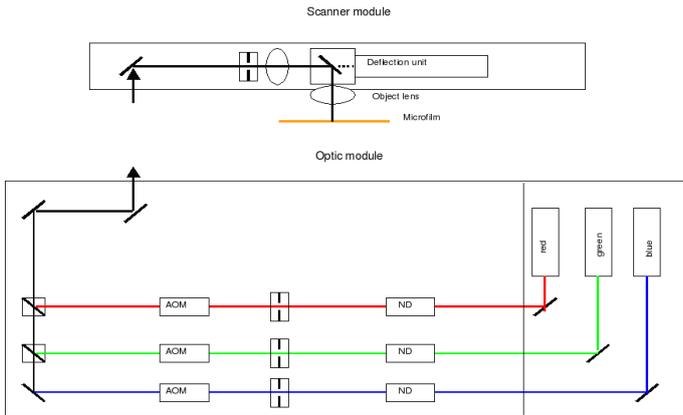
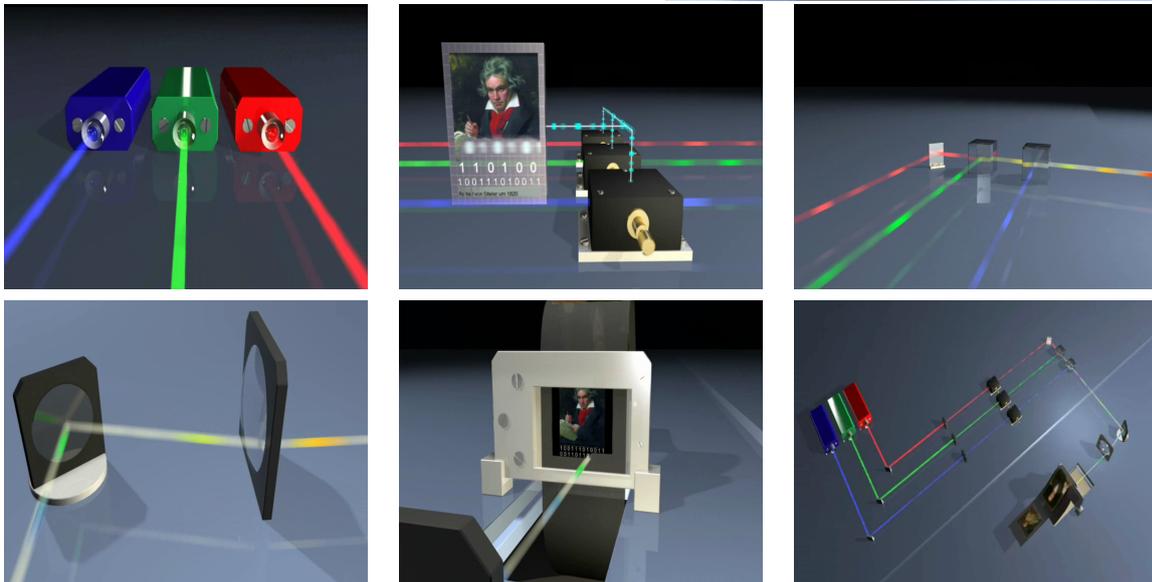


fig. 1

fig. 2



Technische Daten des ArchivLasers

Framegröße	32 x 45 mm (35 mm Film)	Belichtungszeit	40 sec / Frame
Pixelgröße	3 µm (160 Linienpaare / mm)	Filmlänge	≤ 600 m (13,000 Frames)
Pixel / Frame	10,666 x 15,000 Pixel	Zoomfaktor	34 (1 DIN A0 Blatt / Frame)
Farbtiefe	36 (3 x 12) Bit	Kapazität	16 Blätter DIN A4 / Frame bzw. 1 GB / 25 cm Film
Filmmaterial	S/W oder Color-Mikrofilm (Ilford, Kodak, Agfa)	Produktivität	≈ 1 km Film in vier Tagen

Abbildung 2.5: Die Funktionsweise des ArchivLasers und technische Daten

- a) Wird zur Archivierung ein digitales Speichermedium gewählt, so hat man nun noch immer den Vorteil, daß Bild- und Metadaten gemeinsam migriert werden können. (Denkbar wäre beispielsweise auch die kombinierte Speicherung von Bild- und Metadaten um bei einem Systemwechsel die Arbeitsdaten ohne verschiedene und eventuell verlustbehaftete Migrationszwischen Schritte zu umgehen.) Nachteil dabei ist, dass die Metadaten nicht recherchierbar sind.
 - b) Wird zur Archivierung ein analoges Speichermedium gewählt, so wird zur digitalen Weiternutzung eine ...
4. ... Redigitalisierung notwendig.
 5. Um Bild- und Metadaten erneut nutzen zu können, müssen sie wieder getrennt werden und ...
 6. ... auf die jeweilige Arbeitsumgebung übertragen werden.

Anmerkung: Zu Beachten ist, daß man bei einem digitalen Speichermedium davon ausgehen kann, am Ende des Bearbeitungsprozesses verglichen mit den Eingangsdaten die *selben* Ausgangsmetadaten und die *selben* Ausgangsbilddaten zu erhalten. Bei Benutzung eines analogen Mediums ist es allenfalls bei den Metadaten noch möglich, die *selben* Eingangs wie Ausgangsdaten zu erhalten. Bei den Bildern erhält man hier unter idealen Bedingungen jedoch „nur noch“ die *gleichen* Ausgangsdaten. Das ist damit zu erklären, daß bei fehlerfreier Interpretation Metadaten wieder in ihr Datenbankformat geschrieben werden können, die Bilddaten jedoch allenfalls die gleiche Bildinformation ausgeben, jedoch unmöglich die selbe digitale Datenstruktur annehmen werden!

Der ARCHE-Ansatz schließt jedoch die Benutzung eines digitalen Mediums aus, sondern sieht die Verwendung eines Mikrofilms in Verbindung mit dem ArchiveLaser[®] vor. Am in [Abbildung 2.7](#) gezeigten, auf diese Anwendung zugeschnittenen Workflow ist mit gestrichelten Linien und dem „unendlich“-Symbol ∞ nochmals der immerwiederkehrende Weg der Datenvorhaltung und Datenmigration angedeutet. Sein Aufwand ist kaum kalkulierbar.

In [Abbildung 2.8](#) ist der Ausschnitt des ARCHE-Workflows hervorgehoben, der vom Landesarchiv Baden-Württemberg und der Universitätsbibliothek Stuttgart bearbeitet wurde.

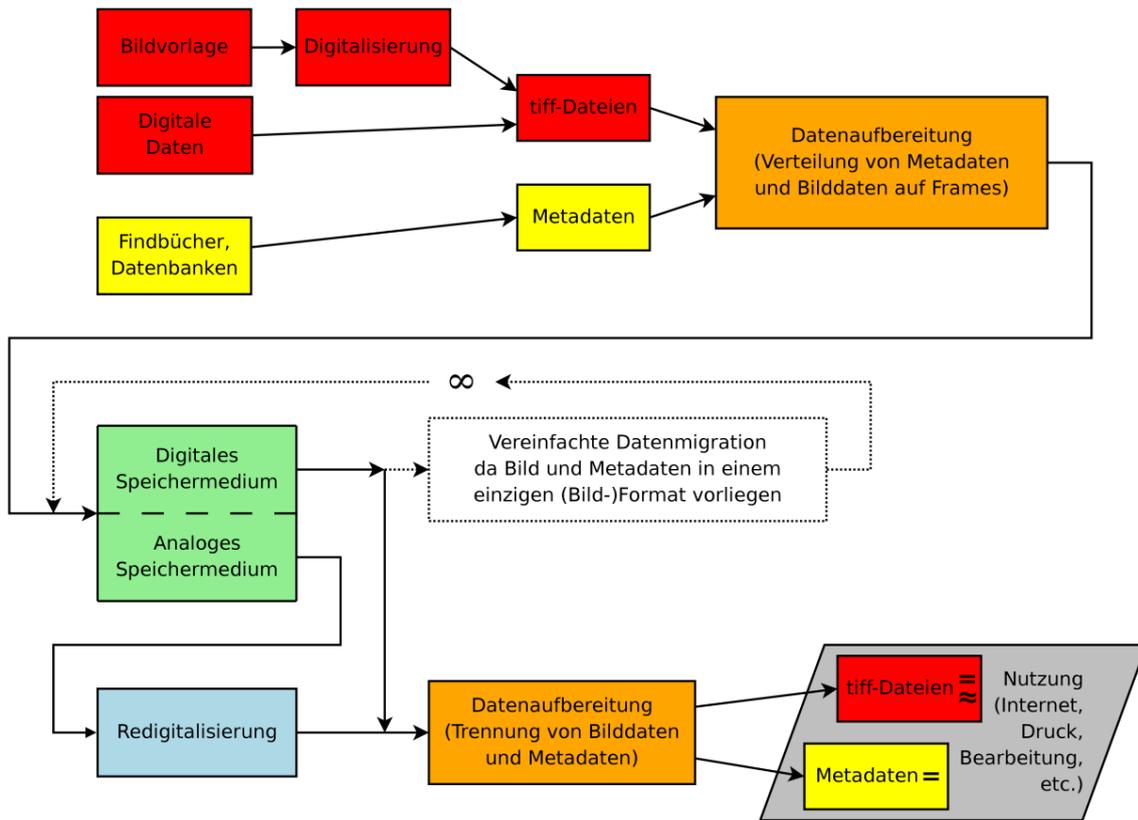


Abbildung 2.6: Der schematische ARCHE-Workflow

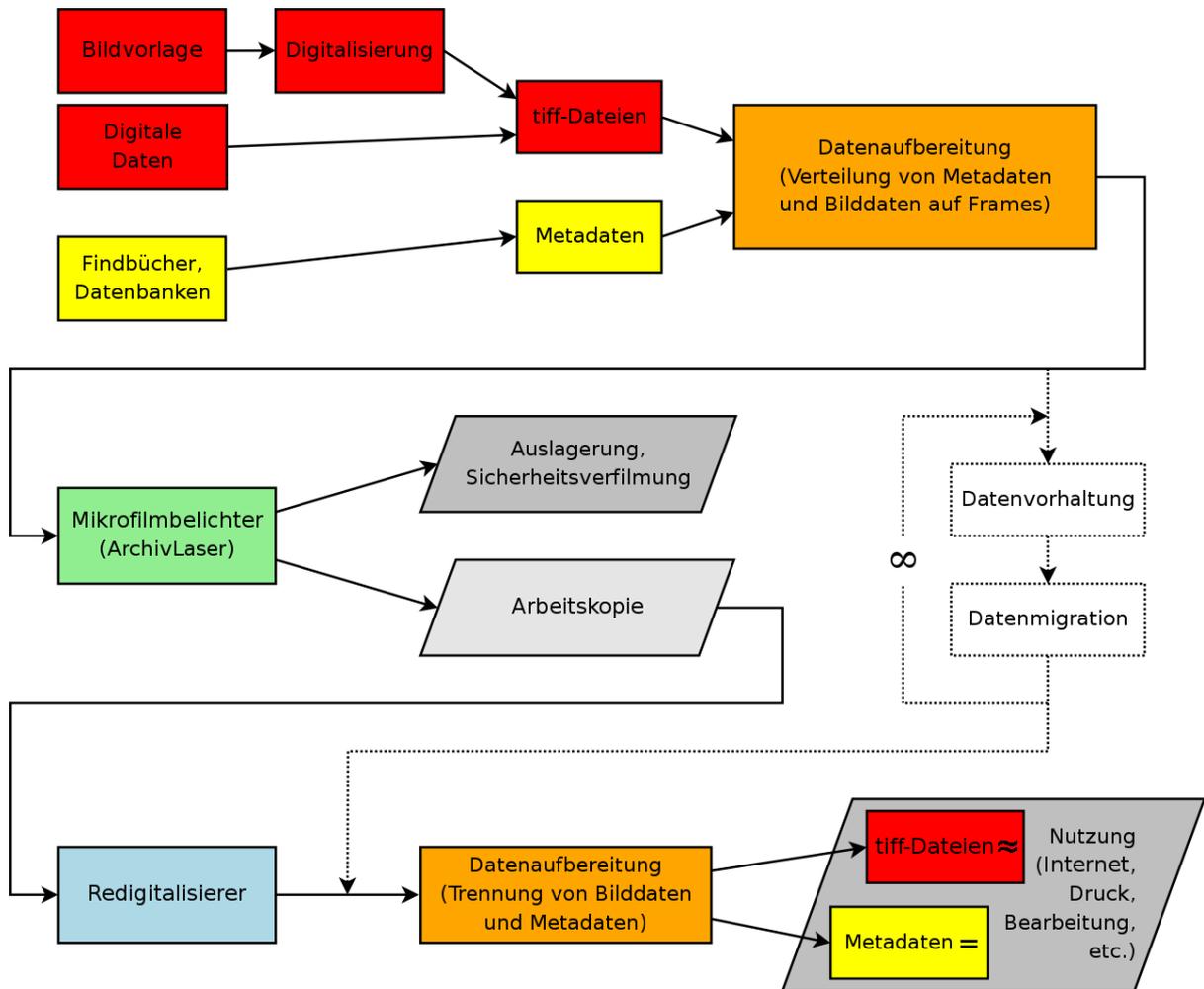


Abbildung 2.7: Der praktische ARCHE-Workflow mit dem Laserbelichter im Mittelpunkt

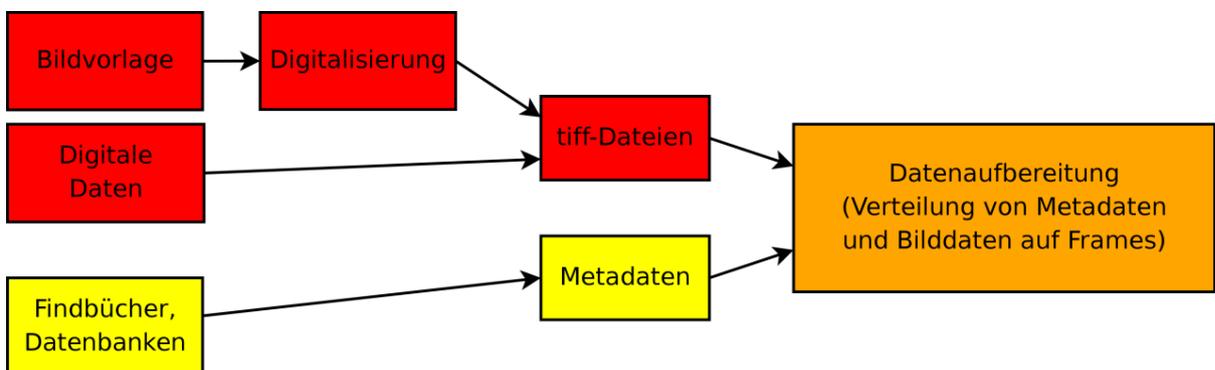


Abbildung 2.8: Der Stuttgarter Anteil am ARCHE-Workflow

3 Die theoretischen Grundlagen zur ARCHE-Bildaufbereitung

3.1 Randbedingungen

- Im Folgenden wird vorausgesetzt, daß ein Bestand bereits digitalisiert ist.
- Bei der Laserbelichtung sollen nicht nur die Bilder selbst auf dem Film konserviert werden, sondern auch die Metadaten (Findbuchinformationen).
- Ein Software-Prozess, das s.g. „Rendering“, erstellt eine Sequenz von Bildern im Tiff-Format plus eine s.g. JOB-Datei die die Reihenfolge der Ausbelichtung vorgibt.
- Die Hier gemachten Angaben dienen zur Vorbereitung der Datenbestände für das „Rendering“.
- Zur Vereinheitlichung der Terminologie wird nachfolgend zwischen Auftraggeber und Dienstleister unterschieden. Unter dem Auftraggeber sind die Archive, Bibliotheken etc. zu verstehen, die ihre digitalisierten Bestände zur Verfügung stellen. Das "Rendering" und die Filmbelichtung erfolgt bei einem Dienstleister.

3.2 Vorgehensweise für Archive und Bibliotheken

Die Auftraggeber stellen dem Dienstleister folgende Daten zur Verfügung:

1. Eine *Textdatei* mit Präliminarien zum Bestand selbst .
2. Eine *CVS-Datei* mit Metadaten der Bestände (Findbücher, Datenbankauszüge, etc.)
3. Die *Bilddateien*; optional können jedem einzelnen Bild individuelle *Metadaten* beigegeben werden.

3.3 Anforderungen an eine Textdatei

- **Die Textdatei enthält beliebige Metainformationen zum gesamten Bestand.** In der Textdatei können beispielsweise Informationen zur Herkunft des Bestandes, zur Überlieferungsgeschichte, ein Vorwort und weitere einleitende Texte enthalten sein. Sehr sinnvoll ist es, auch eine Beschreibung der Datenbank- bzw. Findbuchfelder beizulegen, sofern diese nicht eindeutig selbsterklärend sind!
- Als Format für die Textdatei ist einfach formatierter ASCII-Text (Plain Text) zu verwenden.
- Für die Textdatei gibt es keine Beschränkungen im Umfang.

Beispiel 1: Inhalt der Textdatei

Der Musterbestand

Die Bestandsgeschichte des Musterbestandes ist äußerst kurz. Er wurde künstlich generiert und dient dem ARCHE-Projekt zu Testzwecken.

3.4 Die Struktur der CSV-Datei

- Die CSV-Datei ist im Prinzip ebenfalls nur eine ASCII-Datei, welche allerdings im **CSV-Format** strukturiert ist.
- **Wahl eines geeigneten Zeichens als Feldtrenner.** Beim Datenexport in das CSV-Format fordern Datenbankprogramme üblicherweise zur Eingabe eines Zeichens für *Feldtrenner* und evtl. auch *Texttrenner* auf. Für den Texttrenner kann getrost das Anführungszeichen " beibehalten werden. Entscheidend ist der Feldtrenner. Für ihn muß unbedingt ein Zeichen gewählt werden, welches in keinem Datensatz vorkommt. Z.B. **p** (**AltGr+p**) eignet sich hierfür.
- **Auf die Relevanz der Informationen achten!** Vor dem Datenexport ist auf eine sinnvolle Feldauswahl zu achten. Gänzlich ungenutzte Datenbankfelder und Felder, die lediglich zum Betrieb des benutzten Datenbanksystems angelegt wurden, sollten nicht mit in das Digitalisierungsprojekt übernommen werden.
- **Es gibt drei verschiedene Spaltenarten.** Ein Findbuch besteht in der Regel aus *Signaturfeldern*, *Datenfeldern* und ist außerdem häufig in eine *Kapitelstruktur* eingebunden. Deshalb sind für die CSV-Datei drei Spaltenarten definiert:
 1. **A-Spalten** enthalten die Signaturbestandteile.
 2. **B-Spalten** enthalten die Bestandsdaten.

3. C-Spalten nehmen die Kapitelstruktur auf.

Anmerkung: Die Informationen der A-Spalten sind auch die einzigen Metainformationen der CSV-Datei, welche später in das jeweilige Bild eingebettet sind. B-Spalten enthalten Metainformationen, die nur am Anfang des Filmes im sog. TOC (*table of contents*) ausbelichtet werden. C-Spalten dienen zur Strukturierung der Metainformationen im sog TOC.

3.4.1 Regeln für Spaltennamen

- Die Spaltennamen werden folgendermaßen gebildet: **Spaltentyp@Feldname**

Beispiel 2:

```
A@Bestand, A@LfdNr, B@Titel, B@Enthält, B@Darin, B@Laufzeit, C@Kapitel, C@Unterkapitel
```

- Für den Spaltentyp sind Groß oder Kleinbuchstaben zulässig (ABC=abc).
- Feldnamen (nach dem @) dürfen enthalten:
 - Buchstaben (groß und klein, auch Umlaute)
 - Ziffern 0 bis 9
 - Interpunktionszeichen und _ (Unterstrich)

Anmerkung: Der Unterstrich _ wird als Leerzeichen interpretiert.

- Feldnamen dürfen keine Leer- und Sonderzeichen wie \ / : * () " enthalten.
- Das erste Zeichen eines Feldnamens darf keine Ziffer sein, es muß ein Buchstabe oder eines der erlaubten Interpunktionszeichen sein.
- In der Praxis sollte das erste Zeichen möglichst immer ein Buchstabe, allenfalls ein Unterstrich sein.
- Feldnamen dürfen nicht mit der Zeichenfolge xml (oder XML) beginnen.
- Feldnamen müssen aus mindestens einem Buchstaben bzw. Unterstrich bestehen.

Beispiel 3: unproblematische Feldnamen sind

```
Anschrift  
hausnummer  
plz_ort  
Enthält  
ZiNr  
AKZ  
a1
```

3.4.2 Regeln für Spaltenarten

- Jede Tabelle muß mindestens eine A-, mindestens eine B- und mindestens eine C-Spalte enthalten.
- Die Reihenfolge der Spalten bestimmt je nach Spaltentyp die Hierarchie der Inhalte stets von links nach rechts in absteigender Priorität. Die Einträge in einer rechtsstehenden Spalte eines Typs ist den Einträgen in den linken Spalten gleichen Typs untergeordnet.
- Es sollte eine eindeutige und übersichtliche Tabellenstruktur gewählt werden.

A-Spalten

- Die Zusammensetzung der Gesamtsignatur ergibt sich aus der Folge der A-Spalten von links nach rechts in absteigender Priorität.

Beispiel 4:

A@BestandA@LfdNr

A123	1	= z.B. A123/1 oder A123-1
A123	2	
...	...	
A123	101	
A123	102	

- Zusammensetzung der Einträge der A-Spalten ergibt eine Signatur bzw. eine ID, die innerhalb der gesamten Tabelle eindeutig ist. Doppelte Einträge sind unzulässig.

Beispiel 5:

A@BestandA@LfdNr

A123	101	
A123	102	
...	...	
A123	101	<-- unzulässig

- Zellen in A-Spalten dürfen nicht leer sein, sie müssen immer einen Inhalt haben.

B-Spalten

- B-Spalten enthalten die eigentlichen Bestandsdaten.
- Die Reihenfolge der B-Spalten hat keine besondere Bedeutung, sollte jedoch sinnvoll gewählt werden.
- Die *erste* B-Spalte muß einen Inhalt haben (z.B. Titel), alle übrigen B-Spalten dürfen leer sein.

C-Spalten

- C-Spalten werden verwendet, um die baumförmige Struktur der Metadaten abzubilden. Die Spaltenfolge bestimmt die Hierarchie der Kapitelstruktur von links nach rechts in absteigender Priorität.

Beispiel 6:

C@Kapitel_Ebene_1	C@Kapitel_Ebene_2	C@Kapitel_Ebene_3
Bauprojekte	Bauaufnahmen in Neresheim	Bezirk 1
Bauprojekte	Bauaufnahmen in Neresheim	Bezirk 2
Bauprojekte	Bauaufnahmen in Heidelberg	Nord
Bauprojekte	Bauaufnahmen in Heidelberg	Süd

Die Einträge aus der C@Kapitel_Ebene_2 sind den Einträgen aus C@Kapitel_Ebene_1 untergeordnet wobei die Einträge aus der C@Kapitel_Ebene_3 wiederum den Einträgen aus C@Kapitel_Ebene_2 untergeordnet sind. Gleichbedeutend könnte man in Anlehnung an eine Buchstruktur unterscheiden: Kapitel, Unterkapitel, Abschnitt, usw.

- Es dürfen keine Kapitelnummern angegeben werden. Die Kapitelnumerierung wird von der Software übernommen.

Beispiel 7: Das selbe Beispiel nach der automatischen Numerierung als Baumstruktur

```
1 Bauprojekte
  1.1 Bauaufnahmen in Neresheim
    1.1.1 Bezirk 1
    1.1.2 Bezirk 2
  1.2 Bauaufnahmen in Heidelberg
    1.2.1 Nord
    1.2.2 Süd
```

- Hier fälschlich eingetragene Kapitelnummern würden dagegen zu verfälschten Einträgen führen:

Beispiel 8: Baumstruktur mit fehlerhafter Numerierung des Bearbeiters

```
1 1. Bauprojekte
    1.1 1.1. Bauaufnahmen in Neresheim
        1.1.1 1.1.1. Bezirk 1
        1.1.2 1.1.2. Bezirk 2
    1.2 1.2. Bauaufnahmen in Heidelberg
        1.2.1 1.2.1. Nord
        1.2.2 1.2.2. Süd
```

- Die *erste* C-Spalte, also das C-Feld mit der höchsten Priorität, und übergeordnete C-Felder müssen einen Inhalt haben. Jeder Datensatz muß eindeutig in die Kapitelstruktur einzuordnen sein.

Beispiel 9: Beispiel für eine gültige Tabelle

C@Kapitel_Ebene_1	C@Kapitel_Ebene_2	C@Kapitel_Ebene_3
Bauprojekte	Bauaufnahmen in Neresheim	Bezirk 1
Bauprojekte	Bauaufnahmen in Neresheim	Bezirk 2
Bauprojekte	Bauaufnahmen in Heidelberg	!!!
Bauprojekte	!!!	!!!

Beispiel 10: Beispiel für eine ungültige Tabelle

C@Kapitel_Ebene_1	C@Kapitel_Ebene_2	C@Kapitel_Ebene_3
Bauprojekte	???	Bezirk 1
???	Bauaufnahmen in Neresheim	Bezirk 2
???	Bauaufnahmen in Heidelberg	!!!
Bauprojekte	!!!	!!!

Befüllung der Zellen

- Zellen können mit beliebigem Text angefüllt sein. Alle Zeichen sind zulässig mit Ausnahme des beim Datenexport selbst zu wählenden Feldtrenner-Zeichens!
- Die Zeichenkombinationen `
`, `
` oder `
` erzeugen einen Zeilenumbruch.

Findbuch als Ausgangspunkt
(Kapiteleinteilungen bzw. Strukturdaten = blau, Signaturdaten = rot)

Findbuch für den Bestand A123

Kapitel 1 - Bauprojekte

1.1 Bauaufnahmen in Neresheim

1.1.1 Bezirk 1

1	Gebäudeerfassung in Neresheim 1 Enthält: Lichtpausen	1968-1975
---	---	-----------

1.1.2 Bezirk 2

2	Gebäudeerfassung in Neresheim 2 Enthält: Lichtpausen	1968-1978
---	---	-----------

1.2 Bauaufnahmen in Heidelberg

1.2.1 Nord

3	Gebäudeerfassung in Heidelberg Nord	1980-1989
---	-------------------------------------	-----------

1.2.2 Süd

4	Gebäudeerfassung in Heidelberg Süd	1980-1989
---	------------------------------------	-----------

Kapitel 2 - Verwaltung

2.1 Sachgut

5	Instandhaltung der Betriebsmittel Enthält: Rechnungen	1954-2004
---	--	-----------

2.2 Personal

6	Beilaunehaltung der Betreiber Enthält: Korrespondenz	1980-2004
---	---	-----------

Abbildung 3.1: Beispiel 11 – Die gedruckte Form eines Findbuchs

	A	B	C	D	E	F	G	H
1	A@Bestand	A@LfdNr	B@Titel	B@Enthält	B@Laufzeit	C@Kapitel	C@Unterkapitel	C@Abschnitt
2	A123	1	Gebäudeerfassung in Neresheim 1	Lichtpausen	1968-1975	Bauprojekte	Bauaufnahmen in Neresheim	Bezirk 1
3	A123	2	Gebäudeerfassung in Neresheim 2	Lichtpausen	1968-1978	Bauprojekte	Bauaufnahmen in Neresheim	Bezirk 2
4	A123	3	Gebäudeerfassung in Heidelberg Nord		1980-1989	Bauprojekte	Bauaufnahmen in Heidelberg	Nord
5	A123	4	Gebäudeerfassung in Heidelberg Süd		1980-1989	Bauprojekte	Bauaufnahmen in Heidelberg	Süd
6	A123	5	Instandhaltung der Betriebsmittel	Rechnungen	1954-2004	Verwaltung	Sachgut	
7	A123	6	Beilauehaltung der Betreiber	Korrespondenz	1980-2004	Verwaltung	Personal	
8								
9								
10								
11								
12								
13								
14								
15								
16								
17								
18								
19								
20								
21								
22								
23								
24								
25								
26								
27								
28								
29								
30								

Abbildung 3.2: Beispiel 12 – Das Findbuch in Tabellenform

```
"A@Bestand";"A@LfdNr";"B@Titel";"B@Enthält";"B@Laufzeit";"C@Kapitel";"C@Unterkapitel";"C@Abschnitt"
"A123";1;"Gebäudeerfassung in Neresheim 1";"Lichtpausen";"1968-1975";"Bauprojekte";"Bauaufnahmen in Neresheim";"Bezirk 1"
"A123";2;"Gebäudeerfassung in Neresheim 2";"Lichtpausen";"1968-1978";"Bauprojekte";"Bauaufnahmen in Neresheim";"Bezirk 2"
"A123";3;"Gebäudeerfassung in Heidelberg Nord";"1980-1989";"Bauprojekte";"Bauaufnahmen in Heidelberg";"Nord"
"A123";4;"Gebäudeerfassung in Heidelberg Süd";"1980-1989";"Bauprojekte";"Bauaufnahmen in Heidelberg";"Süd"
"A123";5;"Instandhaltung der Betriebsmittel";"Rechnungen";"1954-2004";"Verwaltung";"Sachgut";
"A123";6;"Beilauehaltung der Betreiber";"Korrespondenz";"1980-2004";"Verwaltung";"Personal";
```

Abbildung 3.3: Beispiel 13 – Gültige CSV-Datei mit Semikolon ; als Feldtrenner

3.5 Vorbereitung der Bilddaten

3.5.1 Ideale Spezifikationen der Bildfiles

Format: TIFF

Farbtiefe: 48 Bit

Farbraum: ECI-RGB

Bildgröße: Bis maximal ca. 15000x10666 Pixel

(Nach Absprache sind auch unter Inkaufnahme eventueller Qualitätseinbußen und Einschränkungen andere Bildformate und Farbtiefen zulässig.)

Um einen eindeutigen Bezug zwischen den Bildfiles und der CSV-Datei herzustellen, muß die Signatur entweder im *Dateinamen* oder im *Verzeichnispfad* abgebildet werden.

3.5.2 Variante 1: Auflösung der Signatur im Dateinamen

- Jedes Bildfile wird mit einer eindeutigen und einmaligen Kombination aus Signatur und Bildnummer benannt.
- Die Signatursegmente werden durch einen einzelnen Unterstrich `_` getrennt und entsprechen den jeweiligen **A@**-Einträgen der CSV-Datei.
- Zwei weitere Unterstriche `__` ergänzen die Signatur um eine beliebige Anzahl von numerisch fortlaufenden Bildnummern.
 - Diese Bildnummern werden innerhalb der CSV-Datei nicht einzeln aufgelöst und ...
 - ... bestimmen die Reihenfolge der Aufnahme.
- Zu beachten ist, daß innerhalb der Bildnummern zwar **Lücken zulässig** sind, alle Bildnummern aber aus einer **identischen Anzahl von Ziffern** bestehen müssen!

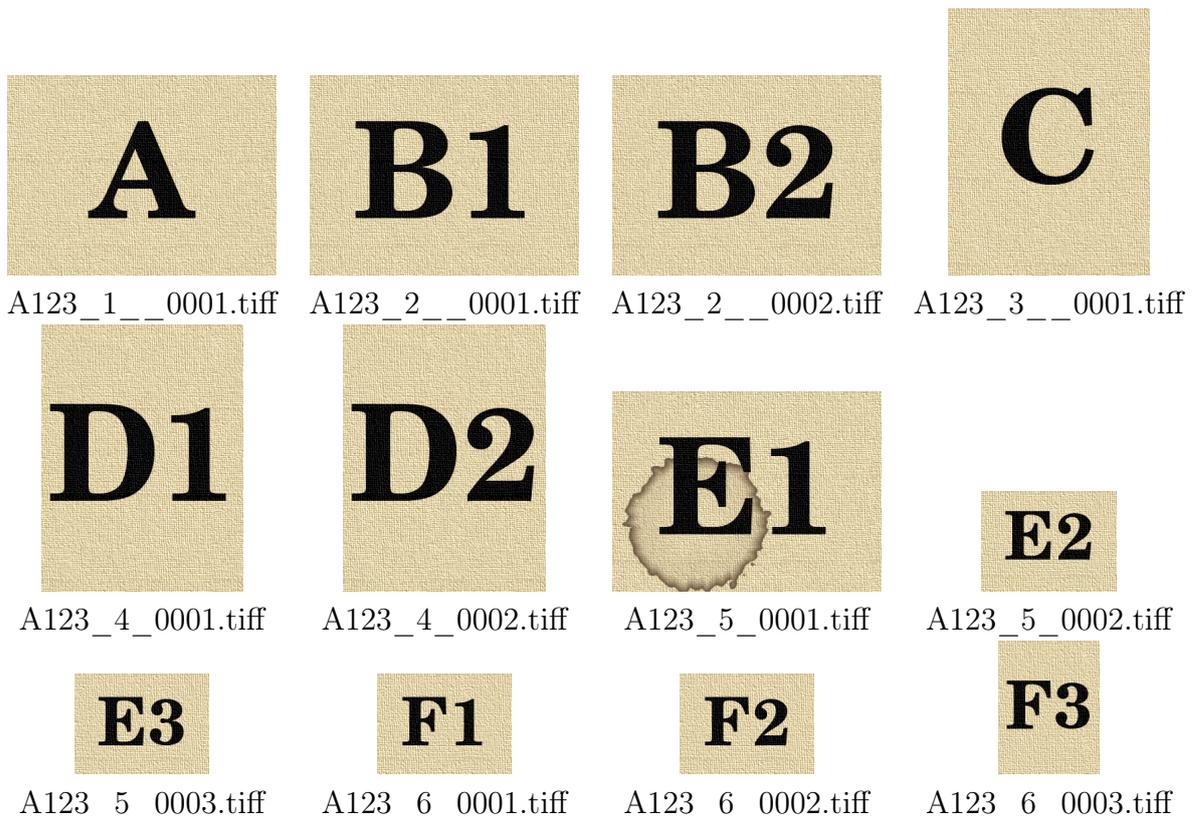
Beispiel 14: Beispiel für **gültige** Dateinamen

```
A123_1__0001.tif  
A123_1__0002.tif  
A123_1__0005.tif !!!  
A123_1__0006.tif
```

Beispiel 15: Beispiel für ungültige Dateinamen

```
A123_1__001.tif  
A123_1__002.tif  
A123_1__03.tif  
A123_1__0004.tif
```

Beispiel 16 – Korrekte Benennung eines Beispielbestandes



3.5.3 Variante 2: Auflösung der Signatur durch den Verzeichnisbaum

- Die Signatursegmente werden durch ein gleichnamiges Verzeichnis benannt. Untergliederung erfolgt durch Unterverzeichnisse.
- Die der Signatur zuzuordnenden Bilder liegen im jeweils untersten Verzeichnis.
- Jedem Bild ist eine eindeutige und alphanumerisch fortlaufende Bildnummer zugeordnet. Diese Bildnummern werden innerhalb der CSV-Datei nicht einzeln aufgelöst und bestimmen die Reihenfolge der Aufnahme.

Beispiel 17: Beispiel für eine gültige Verzeichnisstruktur

```
./A123/  
  1/  
    Datei0001.tif  
./A123/  
  2/  
    A0001.tif  
    0002.tif  
./A123/  
  3/  
    Bild.tif  
./A123/  
  4/  
    B0001.tif  
    A0002.tif  
./A123/  
  5/  
    0001.tif  
    0002.tif  
    0003.tif  
./A123/  
  /6  
    A123_6__0001.tif  
    A123_6__0002.tif  
    A123_6__0003.tif
```

Anmerkung: Diese Schreibweisen sind zwar prinzipiell möglich, doch ist zu beachten, daß durch die alphanumerische Sortierung folgende Reihenfolgen entstehen:

```
...  
./A123/  
  2/  
    A0001.tif  
    0002.tif  
...  
./A123/  
  4/  
    B0001.tif  
    A0002.tif  
...
```



```
...  
./A123/  
  2/  
    0002.tif  
    A0001.tif  
...  
./A123/  
  4/  
    A0002.tif  
    B0001.tif  
...
```

3.5.4 Sonderfall: Individuelle Metadatenergänzung zum Bild

Es besteht die Möglichkeit jedem Bild individuelle Metadaten beizufügen. Hierzu genügt es im Bildverzeichnis dem jeweiligen Bild eine gleichnamige xml-Datei beizufügen, die um das Attribut .xml ergänzt wird:

Beispiel 18:

```
...
A123_5__0001.tif
A123_5__0001.tif.xml
...
```

Die Eingabe dieser Daten wird durch das Tool metainput wesentlich erleichtert. Exemplarisch soll hier nur folgende xml-Datei gezeigt werden:

Beispiel 19: Struktur einer xml-Datei für die individuelle Metadaten-Ergänzung eines Bildes

```
<?xml version="1.0" encoding="UTF-8"?>
<metadata>
  <field>
    <fieldName>Zustand</fieldName>
    <fieldValue>schlecht</fieldValue>
  </field>
  <field>
    <fieldName>Bemerkung</fieldName>
    <fieldValue>Großer Fleck</fieldValue>
  </field>
</metadata>
```

Was zu folgender Ausgabe führt:

Beispiel 20: Ausgabe der individuellen Metadaten-Ergänzung eines Bildes

```
Zustand=schlecht; Bemerkung=großer Fleck;
```

3.6 Anordnung der Dateien auf dem Datenträger

Das Filesystem und die Verteilung der Daten auf dem Datenträger ist beliebig. Entscheidend ist die Anordnung sämtlicher für die Bearbeitung des Bestandes relevanten Daten innerhalb eines Wurzelverzeichnisses:

Beispiel 21a: Unix

```
./Beliebiger Pfad/  
./meta.txt  
./meta.csv  
./A123/  
1/  
A123_1__0001.tif  
./A123/  
2/  
A123_2__0001.tif  
A123_2__0002.tif  
./A123/  
3/  
A123_3__0001.tif  
./A123/  
4/  
A123_4__0001.tif  
A123_4__0002.tif  
./A123/  
5/  
A123_5__0001.tif  
A123_5__0001.tif.xml  
A123_5__0002.tif  
A123_5__0003.tif  
./A123/  
6/  
A123_6__0001.tif  
A123_6__0002.tif  
A123_6__0003.tif
```

oder:

Beispiel 21b: Dos/Windows

```
C:\Beliebiger Pfad\  
meta.txt  
meta.csv  
A123\  
1\  
A123_1__0001.tif  
A123\  
2\  
A123_2__0001.tif  
A123_2__0002.tif  
A123\  
3\  
A123_3__0001.tif  
A123\  
4\  
A123_4__0001.tif  
A123_4__0002.tif  
A123\  
5\  
A123_5__0001.tif  
A123_5__0001.tif.xml  
A123_5__0002.tif  
A123_5__0003.tif  
A123\  
6\  
A123_6__0001.tif  
A123_6__0002.tif  
A123_6__0003.tif
```

oder als praktisches Beispiel in einem Dateimanager:

Name	Größe	Dateityp	Geändert	Berechtigungen
a123	16,0 KB	Ordner	31.10.2005 22:39	drwxr-xr-x
1	16,0 KB	Ordner	31.10.2005 22:39	drwxr-xr-x
A123_1_0001.tif	768,4 KB	TIFF-Bild	31.10.2005 22:39	-rwxr-xr-x
2	16,0 KB	Ordner	31.10.2005 22:39	drwxr-xr-x
A123_2_0001.tif	768,4 KB	TIFF-Bild	31.10.2005 22:39	-rwxr-xr-x
A123_2_0002.tif	768,4 KB	TIFF-Bild	31.10.2005 22:39	-rwxr-xr-x
3	16,0 KB	Ordner	31.10.2005 22:39	drwxr-xr-x
A123_3_0001.tif	768,4 KB	TIFF-Bild	31.10.2005 22:39	-rwxr-xr-x
4	16,0 KB	Ordner	31.10.2005 22:39	drwxr-xr-x
A123_4_0001.tif	768,4 KB	TIFF-Bild	31.10.2005 22:39	-rwxr-xr-x
A123_4_0002.tif	768,4 KB	TIFF-Bild	31.10.2005 22:39	-rwxr-xr-x
5	16,0 KB	Ordner	13.02.2006 14:46	drwxr-xr-x
A123_5_0001.tif	768,4 KB	TIFF-Bild	31.10.2005 22:39	-rwxr-xr-x
A123_5_0001.tif.xml	483 B	XML-Dokument	22.03.2006 13:28	-rwxr-xr-x
A123_5_0002.tif	192,3 KB	TIFF-Bild	31.10.2005 22:39	-rwxr-xr-x
A123_5_0003.tif	192,3 KB	TIFF-Bild	31.10.2005 22:39	-rwxr-xr-x
6	16,0 KB	Ordner	22.03.2006 13:44	drwxr-xr-x
A123_6_0001.tif	192,3 KB	TIFF-Bild	31.10.2005 22:39	-rwxr-xr-x
A123_6_0002.tif	192,3 KB	TIFF-Bild	31.10.2005 22:39	-rwxr-xr-x
A123_6_0003.tif	192,3 KB	TIFF-Bild	22.03.2006 13:41	-rwxr-xr-x
A123_6_0003.tif.xml	383 B	XML-Dokument	22.03.2006 13:44	-rwxr-xr-x
A123_6_0004.tif	192,3 KB	TIFF-Bild	31.10.2005 22:39	-rwxr-xr-x
A123_6_0005.tif	192,3 KB	TIFF-Bild	22.03.2006 13:41	-rwxr-xr-x
A123_6_0005.tif.xml	383 B	XML-Dokument	22.03.2006 13:44	-rwxr-xr-x
meta.csv	761 B	Textdatei mit komma-getrennten Werten	12.11.2005 09:49	-rwxr-xr-x
meta.txt	152 B	Einfacher Text	31.10.2005 22:39	-rwxr-xr-x
meta.xml	3,7 KB	XML-Dokument	22.03.2006 22:56	-rwxr-xr-x

27 Elemente - 20 Dateien (6,6 MB insgesamt) - 7 Ordner

Abbildung 3.4: Verteilung der Daten auf dem Dateisystem

4 Aufbereitung eines Bestandes am Beispiel eines Musterfilms

Wie in [Kapitel 3](#) beschrieben, werden die Metadaten in den Dateien *meta.txt* und *meta.csv* gespeichert. Die Bilddaten können in einer Verzeichnisstruktur untergebracht werden oder in einem einzigen Verzeichnis vorliegen. Diese durch die Software aufbereiteten Daten werden beim Dienstleister in dieser Reihenfolge ausbelichtet:

1. Inhalt der Datei **meta.txt**.
2. Inhalt der Dateien **meta.csv** in lesbarer Form und in XML-Format.
3. Die **Bilddaten**. Sortierung nach Signatur.

4.1 Ausgabe der Metadaten

4.1.1 meta.txt



00001_ t.tiff

4.1.2 meta.csv

In lesbarer Form

<p style="text-align: center;">Seite 1 von 5</p> <p>1 Bauprojekte 1.1 Bauaufnahmen in Neresheim 1.1.1 Bezirk 1 Bestand A123 LfdNr 1 Titel Gebäudeerfassung in Neresheim 1 Enthält Lichtpausen Laufzeit 1968-1975</p>	<p style="text-align: center;">Seite 2 von 5</p> <p>1.1.2 Bezirk 2 Bestand A123 LfdNr 2 Titel Gebäudeerfassung in Neresheim 2 Enthält Lichtpausen Laufzeit 1968-1978 1.2 Bauaufnahmen in Heidelberg 1.2.1 Nord Bestand A123</p>
<p style="text-align: center;">Seite 3 von 5</p> <p>LfdNr 3 Titel Gebäudeerfassung in Heidelberg Nord Enthält Laufzeit 1980-1989 1.2.2 Süd Bestand A123 LfdNr 4 Titel Gebäudeerfassung in Heidelberg Süd</p>	<p style="text-align: center;">Seite 4 von 5</p> <p>Enthält Laufzeit 1980-1989 2 Verwaltung 2.1 Sachgut Bestand A123 LfdNr 5 Titel Instandhaltung der Betriebsmittel Enthält Rechnungen Laufzeit 1954-2004</p>

00002_.tiff

<p style="text-align: center;">Seite 5 von 5</p> <p>2.2 Personal Bestand A123 LfdNr 6 Titel Beilauehaltung der Betreiber Enthält Korrespondenz Laufzeit 1980-2004</p>

00003_.tiff

In XML-Format

<p style="text-align: center;">Seite 1 von 8</p> <pre><?xml version="1.0" encoding="iso-8859-1"? <metadata> <chapter title="Bauprojekte"> <chapter title="Baufnahmen in Neresheim"> <chapter title="Bezirk 1"> <row> <Bestand type="A">A123</Bestand> <LfdNr type="A">1</LfdNr> <Titel type="B">Gebäudeerfassung in Neresheim 1</Titel></pre>	<p style="text-align: center;">Seite 2 von 8</p> <pre> <Enthält type="B" >Lichtpausen</Enthält> <Laufzeit type="B">1968- 1975</Laufzeit> <files>1</files> </row> </chapter> <chapter title="Bezirk 2"> <row> <Bestand type="A">A123</Bestand> <LfdNr type="A">2</LfdNr> <Titel type="B">Gebäudeerfassung</pre>
<p style="text-align: center;">Seite 3 von 8</p> <pre>in Neresheim 2</Titel> <Enthält type="B" >Lichtpausen</Enthält> <Laufzeit type="B">1968- 1978</Laufzeit> <files>2</files> </row> </chapter> <chapter title="Baufnahmen in Heidelberg"> <chapter title="Nord"></pre>	<p style="text-align: center;">Seite 4 von 8</p> <pre> <row> <Bestand type="A">A123</Bestand> <LfdNr type="A">3</LfdNr> <Titel type="B">Gebäudeerfassung in Heidelberg Nord</Titel> <Enthält type="B" /> <Laufzeit type="B">1980- 1989</Laufzeit> <files>1</files> </row> </chapter> <chapter title="Süd"></pre>

00004_ t.tiff

<p style="text-align: center;">Seite 5 von 8</p> <pre> <row> <Bestand type="A">A123</Bestand> <LfdNr type="A">4</LfdNr> <Titel type="B">Gebäudeerfassung in Heidelberg Süd</Titel> <Enthält type="B" /> <Laufzeit type="B">1980- 1989</Laufzeit> <files>2</files> </row> </chapter> </chapter></pre>	<p style="text-align: center;">Seite 6 von 8</p> <pre> </chapter> <chapter title="Verwaltung"> <chapter title="Sachgut"> <row> <Bestand type="A">A123</Bestand> <LfdNr type="A">5</LfdNr> <Titel type="B">Instandhaltung der Betriebsmittel</Titel> <Enthält type="B" >Rechnungen</Enthält> <Laufzeit type="B">1954- 2004</Laufzeit></pre>
<p style="text-align: center;">Seite 7 von 8</p> <pre> <files>3</files> </row> </chapter> <chapter title="Personal"> <row> <Bestand type="A">A123</Bestand> <LfdNr type="A">6</LfdNr> <Titel type="B">Beilaunehaltung der Betreiber</Titel> <Enthält type="B" >Korrespondenz</Enthält> <Laufzeit type="B">1980-</pre>	<p style="text-align: center;">Seite 8 von 8</p> <pre>2004</Laufzeit> <files>3</files> </row> </chapter> </chapter> </metadata></pre>

00005_ t.tiff

4.2 Ausgabe der Bilddaten

4.2.1 Eigenschaften der „Frames“ zur Aufnahme von Bilddaten

- Ein Frame zur Aufnahme von Bilddaten enthält mindestens einen (oder mehrere) durch ihre Signalfarbgebung gekennzeichnete „Container“.
- Standardmäßig wird ein Container pro Frame dargestellt. Zur besseren Ausnutzung der Filmfläche kann die kombinierte Speicherung mehrerer Container in einem Frame optional hinzugewählt werden.
- Die Anzahl der in einem Frame darstellbaren Container wird durch den zur Verfügung stehenden Raum begrenzt. Container, die in einem Frame keinen Platz mehr finden, werden im Folgeframe dargestellt.

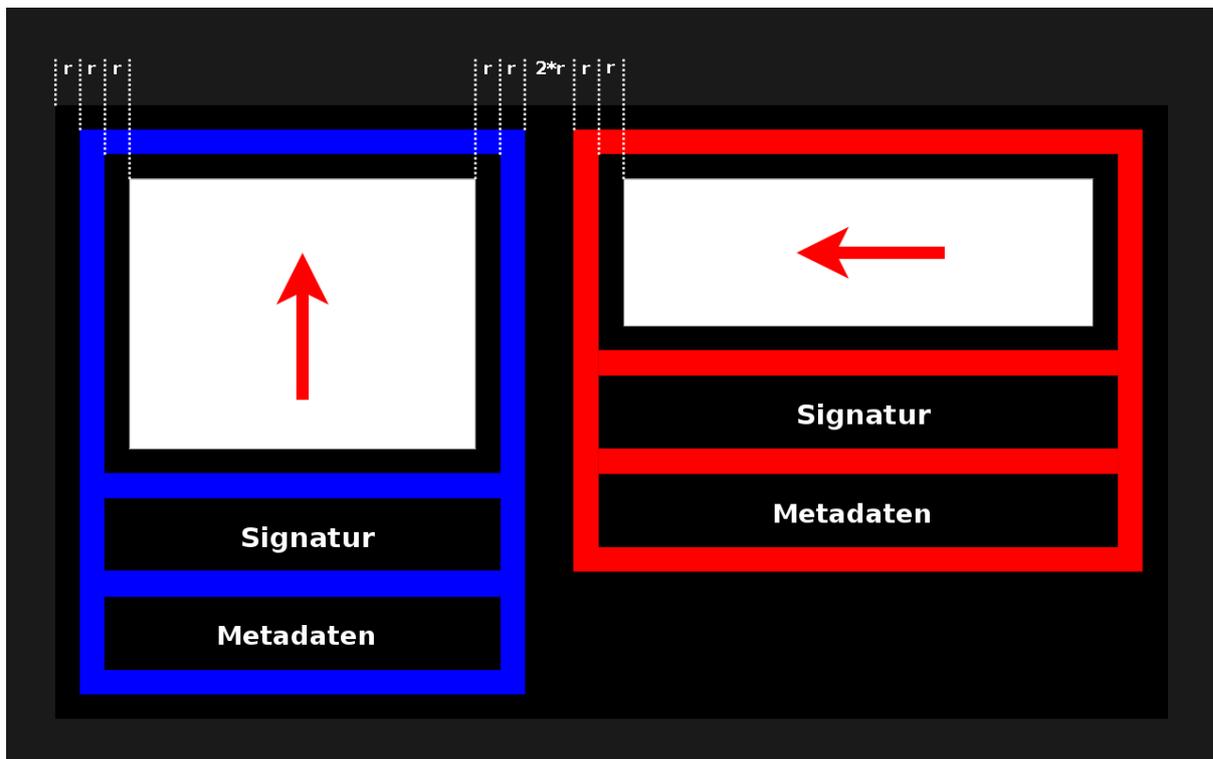


Abbildung 4.1: Der Rahmen eines Containers als Informationsträger für die Redigitalisierung. (r = Breite des sichtbaren Rahmenteils = $1/3$ der Breite des Gesamtrahmens)

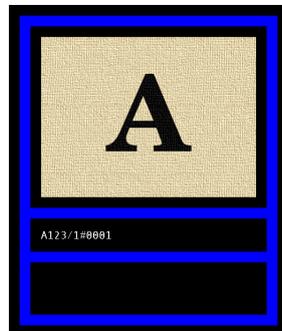
- Ein Container ist dreigeteilt. Der obere Abschnitt enthält das Bild, der mittlere die Signatur, der untere die für jedes Bild individuell speicherbaren Metadaten.

- Der Rahmen eines Containers besteht aus jeweils drei Linien mit gleicher Dicke, wovon jedoch die äußeren und die inneren Linien mit dem Filmhintergrund verschmelzen. Auf diese Weise wird es einem automatisierten Redigitalisierungsprozess erleichtert, zu erkennen welcher Bestandteil des Gesamtframes Information enthält. Auf diese Weise kann relativ zum dargestellten Bild eindeutig zwischen „innen“ und „außen“ unterschieden werden.
- Die Rahmenfarbe eines Containers signalisiert die Ausrichtung des Bildes:

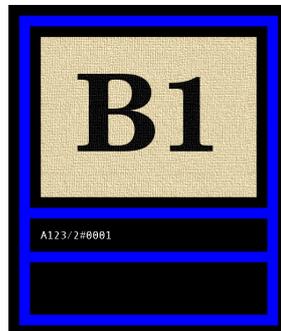
Blau Das Bild wurde unverändert mit der ursprünglichen Ausrichtung wiedergegeben.

Rot Bild wurde von der Rendering-Software um 270° im Uhrzeigersinn gedreht. Dieser kann notwendig werden, wenn ein Bild hochkant nicht auf die verfügbare Fläche paßt.

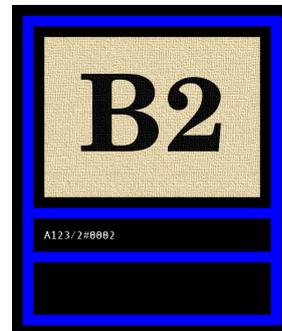
4.2.2 Variante 1: Ausgabe ohne „Nesting“ (Bildhäufung)



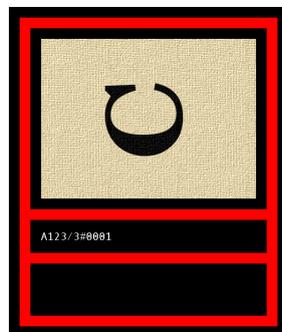
00006.tiff



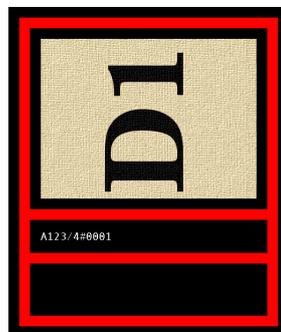
00007.tiff



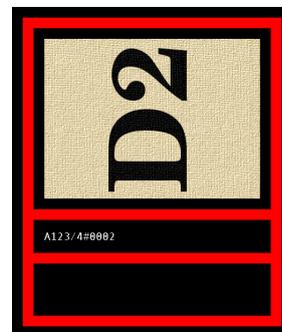
00008.tiff



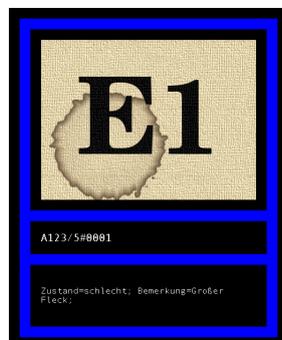
00009.tiff



00010.tiff



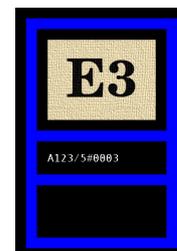
00011.tiff



00012.tiff



00013.tif



00014.tif



00015.tiff

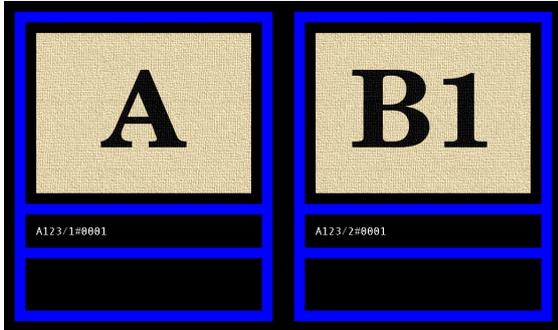


00016.tiff



00017.tiff

4.2.3 Variante 2: Ausgabe mit „Nesting“ (Bildhäufung)



00006.tiff



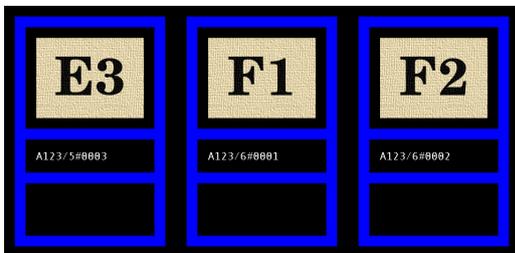
00007.tiff



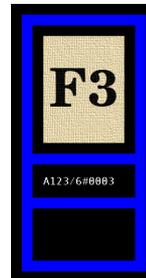
00008.tiff



00009.tiff



00010.tiff



00011.tiff

4.2.4 Schematische Anordnung eines ARCHE-Films

Nachfolgendes Beispiel soll noch einmal die Anordnung der einzelnen Frames auf einem Filmband verdeutlichen.



Abbildung 4.2: Schematische Darstellung eines ARCHE-Mikrofilms

Frame 1: Beschreibung des Bestandes

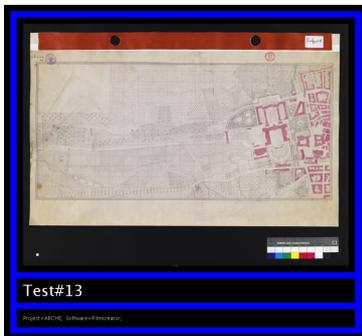
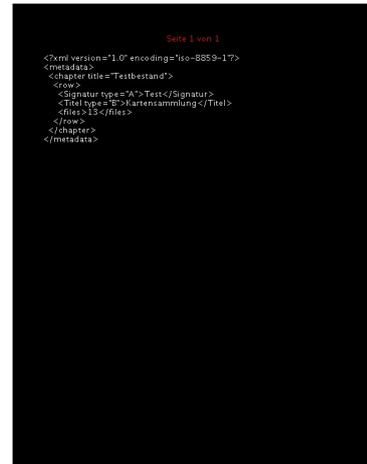
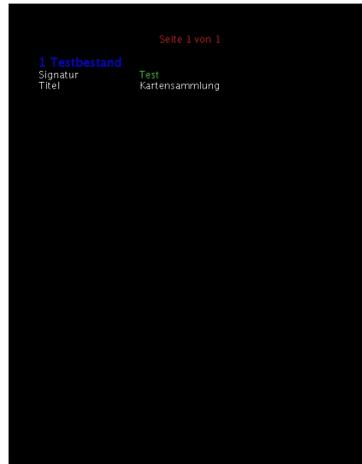
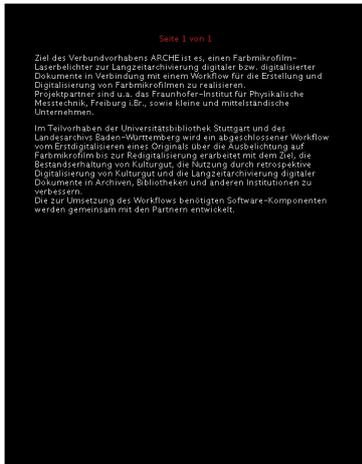
Frame 2: Datenbankauszug in lesbarer Form

Frame 3: Datenbankauszug in XML-Format

Frame 4-5: Speicherung einzelner Bilder pro Frame in horizontaler und vertikaler Ausrichtung

Frame 6-8: Beispiele für kombinierte Speicherung von Bildern durch Nesting

4.2.5 Ein „Musterfilm“ aus praktischen Beispielen



4.2.6 Weitere darstellbare Informationstypen

Die ARCHE-Bildaufbereitung kann grundsätzlich jedes zweidimensional visualisierbare Objekt darstellen, das den Spezifikationen aus [Abschnitt 3.5.1](#) entspricht. Denkbar wären also auch (mehrdimensionale) Strich- und Matrixcodes, deren digitale Daten auf dem analogen Film speicherbar sind!

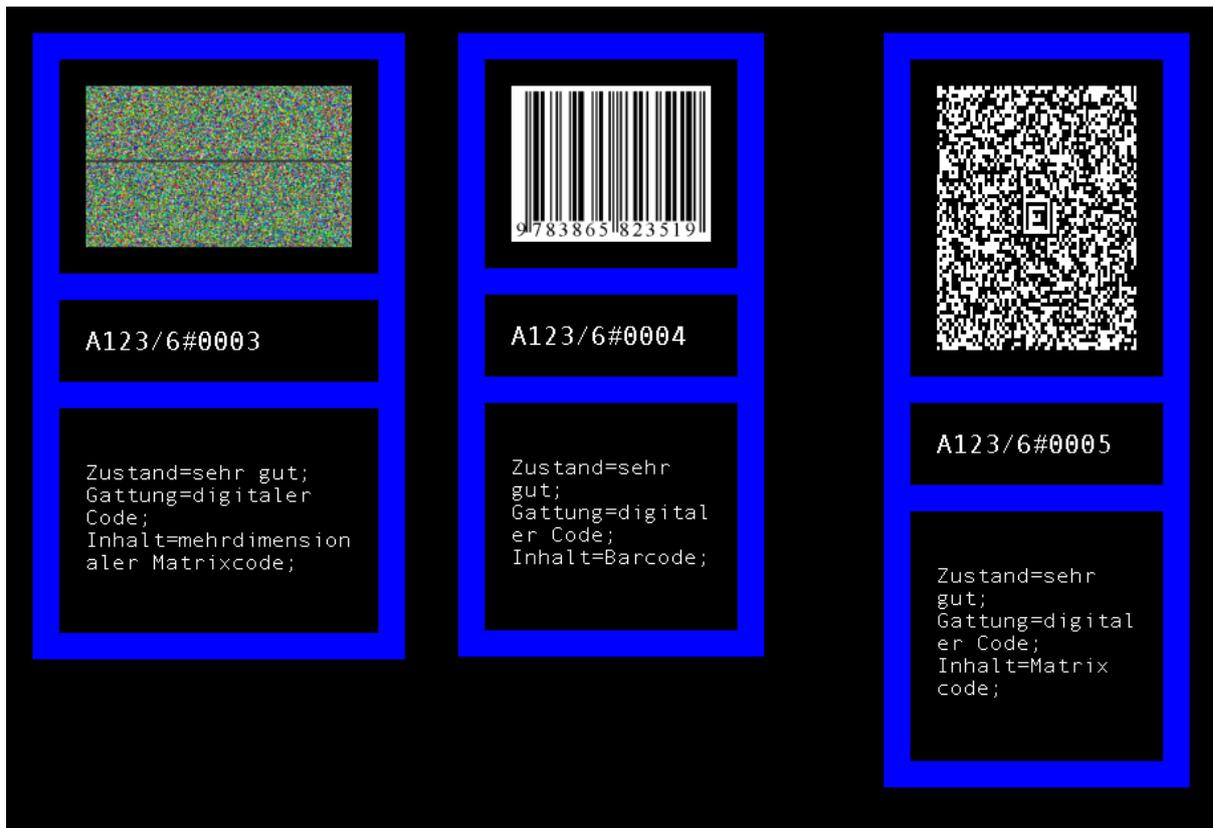


Abbildung 4.3: Auch visualisierte Digitalcodes können in den Containern gespeichert werden

5 Digitalisierung von hochwertigen Testvorlagen

In der Vorbereitungsphase wurden drei Reprodigitalisierungssysteme getestet, weitere in näheren Betracht gezogen.

Ausführlich getestete Geräte

- Von Februar bis April 2005 wurde auf Kosten des Landesarchivs Baden-Württemberg von der Stuttgarter Firma MM-Wandel ein **Zeutschel OS 10000 TT**-Reproscanner gemietet.
- Von April bis Juni 2005 stand dem Projekt ebenfalls auf Mietbasis durch das Landesarchiv Baden-Württemberg eine **Pentacon scan5000S** Reprokamera zur Verfügung.
- Von Juni 2005 bis Projektende erhielt das Projekt leihweise den Reproscanner **HIT Vario XL** der Firma Homrich.

Für sämtliche Tests hatten das Ludwigsburger Institut für Erhaltung von Archiv- und Bibliotheksgut und das Staatsarchiv Ludwigsburg in sehr großzügiger Weise alle benötigten Räumlichkeiten und Materialien ohne zeitliche Beschränkung zur Verfügung gestellt.

In näheren Betracht gezogene oder nur unter Aufsicht des Anbieters getestete Systeme

- **Zeutschel OS 10000 A0**, getestet bei Werkbesuch im April 2005
- **ProServ**, vorgestellt durch Vertreterbesuch im Frühjahr 2005, vorgeführt bei Deutschem Archivtag 2005 und Cebit 2006
- **Cruse**, vorgeführt bei Deutschem Archivtag 2005 und Cebit 2006
- **I2S**, vorgeführt bei Cebit 2006

Quasi „außer Konkurrenz“ ist noch der **Nikon Super Coolscan LS-4000 ED** Diascanner zu nennen, welcher dem Projekt im Juni 2005 vom Höchstleistungsrechenzentrum der Universität Stuttgart (HLRS) unentgeltlich zur Verfügung gestellt wurde.

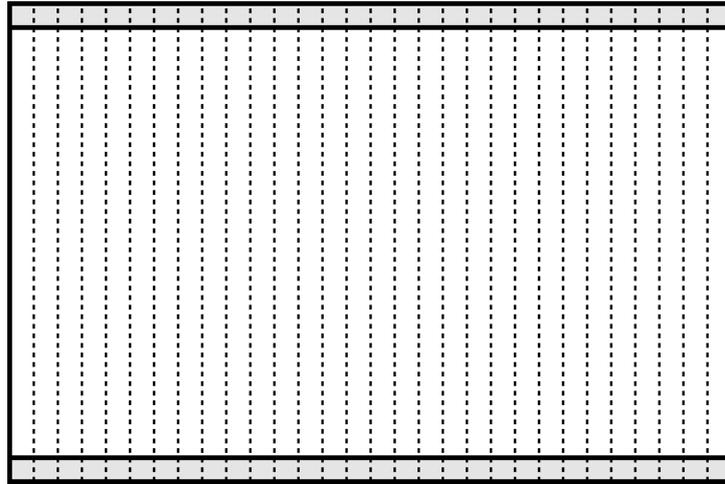


Abbildung 5.1: Prinzip des Zeilenscanners (die gestrichelten Linien stehen für die Ansatzpunkte der Zeile, die grauen Bereiche stehen für die optische Verzerrung und Vignettierung an den Rändern)

5.1 Unterschiedliche Verfahrensweisen

Es gibt zwei unterschiedliche Ansätze zur digitalen Reprographie: den Aufsicht-Zeilenscanner und das (Multi-)Shot-System.

5.1.1 Der Zeilenscanner

Der Zeilenscanner funktioniert nach dem selben Prinzip wie der gewöhnliche Flachbettscanner, außer daß die CCD-Zeile zusammen mit einer Optik über die Vorlage bewegt wird. Dabei nimmt der Zeilenscanner parallel zum Vorschub einzelne Zeilen auf und setzt diese zu einem Gesamtbild aneinander.

Der mechanische Vorschub bereitet dem Zeilenscanner den grundsätzlichen Nachteil, daß Vibrationen das System in Schwingungen versetzen, die auf dem Bild in Form von Unschärfen deutlich sichtbar werden.

Die Scanner der Firmen *Zeutschel*, *ProServ*, *I2S* und *Cruse* nutzen diese Technologie.

- *Cruse* gelingt es bei einigen Modellen, die Vibrationen dadurch zu minimieren, daß nicht der Scankopf bewegt wird, sondern bei sehr viel Raumbedarf der Auflagetisch.
- Die *Zeutschel*-Scanner der OS 10000-Serie haben die Besonderheit, daß der Scankopf hinter dem Vorlagentisch angebracht ist und beim Scan – um stets den gleichen Abstand zur Vorlage zu erhalten – in einer Art Parabelbewegung über den Tisch bewegt wird. Dabei ändert der Scankopf, dem Vorschub entsprechend, seinen Winkel zunehmend. Diese zwei Bewegungsrichtungen erzeugen nicht nur erheblich mehr

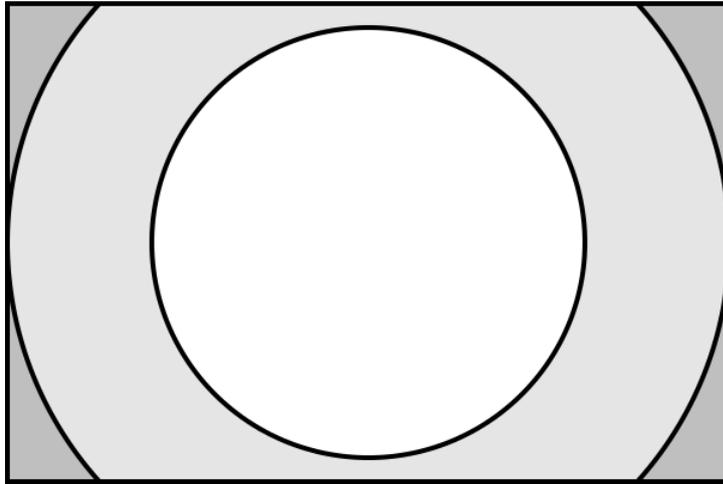


Abbildung 5.2: Prinzip des Shot-Systems (die grauen Bereiche stehen für die optische Verzerrung und Vignettierung an den Rändern)

Vibrationen, sondern führen bei dreidimensional in den Raum ragenden Vorlagen (z. B. Münzen, Orden, Siegel) zu sichtbaren Verzerrungen.

- Der Scankopf des *ProServ*-Scanners wird auf einem Schlitten über die Vorlage gezogen. Durch den geringen Abstand von Zeile zu Vorlage benötigt *ProServ* mehrere einzelne Optik-Zeile-Einheiten, die jeweils eigenständige Bilder erzeugen. Diese Bilder werden nach dem Scan softwaremäßig zusammengesetzt. Bei diesem s.g. „Stitching“ sind erhebliche Unschärfen (zusätzlich zu den Unschärfen des Zeilenscans) beim Zusammensetzen des Bildes unvermeidbar.

5.1.2 Das Shot-System

Das Shot-System ist im Prinzip eine Digitalkamera auf Spezialstativ und speziellen Aufnahmeoptionen zur Kombination mehrerer Bildsegmente zu einem großen Bild. Während der Zeilenscanner nur einen schmalen Streifen der Optik benötigt und dadurch eine höhere Parallelität erzielt, hat das Shot-System den grundsätzlichen Nachteil, einen weiteren Bereich der Optik nutzen zu müssen, wodurch man jene aus der Photographie bekannten Kissenverzerrungen und Belichtungsunterschiede (Vignettierung) an den Rändern berücksichtigen muß.

Diese Abweichung ist bei entsprechend guter Optik allerdings nahezu vernachlässigbar und hier bildlich sehr übertrieben dargestellt. Ein großer Vorteil des Shot-Systems ist die prinzipiell deutlich höhere Schärfe, da es mit Ausnahme des Kameraverschlusses gänzlich ohne mechanische Bewegung bei der Aufnahme auskommt und die Belichtungszeiten sehr kurz sind

Der Scanner der Firma *Homrich* nutzt diese Technologie.

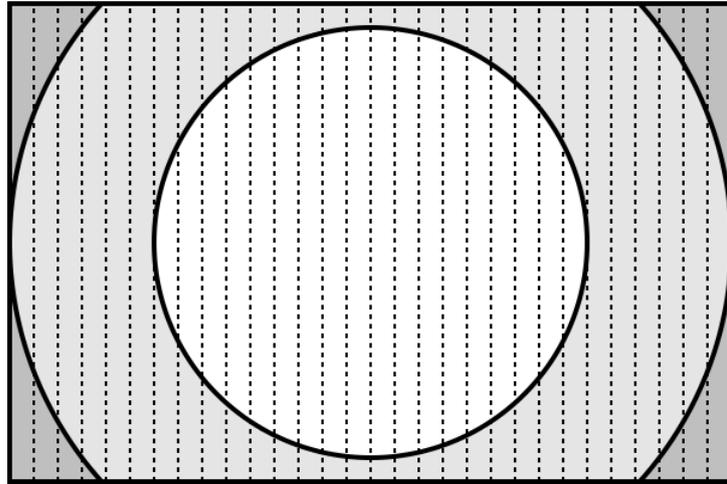


Abbildung 5.3: Prinzip des Zeilenscanners innerhalb einer Kamera (die gestrichelten Linien stehen für die Ansatzpunkte der Zeile, die grauen Bereiche stehen für die optische Verzerrung und Vignettierung an den Rändern)

5.1.3 Der Zeilenscanner in eine Kamera integriert

Einen Sonderweg geht der *Pentacon*-Scanner. Hier ist in einer Kamera, also hinter dem Objektiv, ein Zeilenscanner integriert. Dadurch werden auch die beiden Nachteile der Mechanik und der Optik miteinander kombiniert.

5.2 Rauschverhalten

CCDs haben produktionsbedingte Fehler, die mit zunehmender Temperatur ausgeprägter werden. Diese Fehler werden dadurch sichtbar, daß zwei benachbarte Bildpunkte, die auf der Vorlage die gleiche Farbe beschreiben, im Digitalbild unterschiedliche Farbtöne erhalten. Dieses s.g. „Rauschen“ ist für jeden CCD-Sensor charakteristisch. Das Rauschverhalten fällt beim Shot-System etwas weniger auf, als beim Zeilenscanner, da beim Zeilenscanner abweichende Bildpunkte parallel zueinander über das gesamte Bild gezogen werden. Lediglich beim *Pentacon*-Scanner war dieser Fehler sehr stark ausgeprägt.

5.3 Auflösung

Es muß berücksichtigt werden, wie man die Auflösung mißt. Einige Hersteller geben sie in Bildpunkten pro Inch (dpi) an, andere nennen die Anzahl der Bildpunkte auf der x- und Y-Achse. Den „dots per inch“ geben Hersteller insbesondere dann den Vorzug, wenn ihr System in der Auflösung nicht variabel ist. Völlig unbrauchbar sind Größenangaben in cm oder Blattformaten.

- *ProServ* und *Zeutschel* variieren von 300, 400 bis 600 dpi, wobei diese Auflösungen für die gesamte Fläche nutzbar sind. Allerdings ist der Scankopf in der Höhe nicht verstellbar, so daß keine höheren Auflösungen bei kleinerer Scanfläche möglich sind. *Zeutschel* benutzt einen großen Scankopf (OS 10000 A0: 21000x14800 Bildpunkte), *ProServ* mehrere kleine Scanköpfe (jeweils 5000 Pixel¹).
- *I2S* liefert einen Scanner mit 800 dpi-Scankopf (DigiBook 10000RGB) aus, welcher zudem den ganz großen Vorzug hat, höhenverstellbar zu sein und somit Auflösungen von 200 bis 800 dpi ermöglicht. Ein nennenswerter Vorzug des *I2S* ist die einzigartige Flexibilität beim Austausch des Scankopfes, was eine spätere Erweiterung sehr einfach macht!
- Ebenfalls höhenverstellbar ist der *Cruse* CS 155/450 P. Er erzielt zudem eine Auflösung von 15000x10000 Bildpunkten.
- Der *Pentacon*-Scanner, der an ein beliebiges Stativ montiert werden kann, erreicht eine Auflösung von 12000x8192 Bildpunkten. Damit sind Auflösungen von 0 bis ca. 6000 dpi möglich.
- Beim *Homrich* HIT Vario XL ist sowohl der Scankopf als auch der Tisch in seiner Höhe verstellbar. Das 22 Megapixel (5440x4080 Bildpunkte) große CCD kann im 16-Shot-Verfahren in 16 Bildsegmenten eine Gesamtauflösung von 87 Megapixeln (10880x8080) erzeugen. Eine Besonderheit stellen die zwei Objektive da: Ein Rodenstock-Zoomobjektiv für Aufnahmen aus der Distanz und ein Schneider-Kreuznach-Objektiv für Nahaufnahmen. Damit sind Auflösungen von 113 bis 5662 dpi möglich.

5.4 Farbtiefe

Die Farbtiefe wird gemessen an der Anzahl der Bits pro RGB-Farbkanal. Die begutachteten Scanner der Firmen *Zeutschel*, *ProServ* und *I2S* hatten eine Farbtiefe von 24 (3x8) Bit. Die aktuellen PDF-Prospekte auf der *Zeutschel*-Homepage sprechen neuerdings von 36 bit, was aber offenbar auf Kosten der Auflösung geht, da diese beim OS 10000 A0 stattdessen standardmäßig nur noch mit 400 dpi (statt 600 dpi im getesteten Modell im Sommer 2005) angegeben wird. *Cruse* erzeugt Bilder mit 36 Bit Farbtiefe, *Homrich* und *Pentacon* generieren sogar 48 Bit.

¹ Siehe „Stitching“ in [Abschnitt 5.1.1](#)

5.5 Beleuchtung

5.5.1 Lichteinwirkung auf die Vorlage

Alle Hersteller nutzen 5500 K Kaltlicht. Besonders schonsam ist *Zeutschel*, welches sich insbesondere durch die UV-Freiheit seiner Beleuchtung auszeichnet. Alle Zeilenscanner beleuchten beim Scan nur den aktuell gescannten Bereich. *Homrich* und *Pentacon* erfordern eine Ausleuchtung des Objektes für die gesamte Dauer der Aufnahme. Nach Aussage eines Homrich-Ingenieurs wird derzeit eine Jalousie entwickelt, die die Vorlage sofort nach dem Scan vor dem Lichteinfluß schützen soll.

5.5.2 Lichtverteilung

Zeutschel und *ProServ* haben ein festes Beleuchtungssystem, das nicht separat zu- oder aberschaltbar ist. Der variierende Einfallswinkel des Scankopfes bei *Zeutschel* läßt zudem ein Nachlassen der Lichtintensität zum Bildrand vermuten, welches vermutlich softwareseitig korrigiert wird.

I2S und *Cruse* verfügen über die Möglichkeit die Beleuchtung beidseitig getrennt zu schalten, so daß auch erwünschte Schatteneffekte (z.B. bei Münzen) möglich sind.

Völligen Gestaltungsfreiraum hat man bei *Pentacon* und *Homrich*. Die Beleuchtungskörper des HIT-Vario XL könnten einzeln geschaltet, beliebig gedreht und ausgerichtet werden. Zudem ist die Zuhilfenahme weiterer/anderer Lichtquellen jederzeit machbar. Ermöglicht wird dies durch eine Sonderfunktion des *Homrich*-Scanners, das s.g. „Shading“. Bei diesem Verfahren werden anhand einer Referenzfläche Beleuchtungsschwächen erkannt und später auf Wunsch softwareseitig ausgeglichen.

5.6 Scandauer

Diese Angaben sind stark abhängig von der Vorlagengröße und der gewählten Auflösung, entsprechend optimistisch sind die Angaben der Hersteller. Man muß zudem unterscheiden zwischen der reinen Scandauer und der Zeit, die das Bild zur Übertragung an den PC und zur Nachbereitung benötigt. Beides zusammen ist die Zeit, die der Bearbeiter benötigt, bis er das nächste Bild aufnehmen kann.

Genauere Erfahrungs-Angaben sind hier nur zu machen für *Zeutschel*, *Pentacon* und *Homrich*.

- *Zeutschel* wirbt mit 40 sec bei 400 dpi (OS 10000 A0), 5 sec bei 300 dpi und 3 sec bei 150 dpi (OS 10000 TT). Die Praxis zeigte, daß die Zeit zwischen zwei Bildern beim OS 10000 TT bei 300 dpi etwa 15-30 Sekunden betrug, bei einem Test mit 600 dpi beim OS 10000 A0 deutlich mehr als eine Minute.

- *Pentacon* läuft hier außer Konkurrenz. Bei höchster Auflösung und längster Belichtungszeit, welche auch bei Studio-Lichtverhältnissen noch benötigt wurde, benötigte ein Scan bis zu 19 Minuten.
- Bei *Homrich* ist ein einzelnes Bild (5440x4080 Bildpunkte) je nach Belichtungszeit in weniger als einer Sekunde gemacht. Der Transport zum PC dauert jedoch ca. 20 Sekunden, so daß ein Bearbeiter in diesem Modus ca. 25 Sekunden für eine Aufnahme benötigt. Im 4-Shot-Modus benötigt er ca. 40 Sekunden, im 16-Shot-Modus ca. 70 Sekunden für eine Aufnahme. Zu berücksichtigen ist aber noch, daß die Bilder hier nun erst im Roh-Format vorliegen und später in ein entsprechendes Grafikformat (z. B. Tiff) exportiert werden müssen. Dieser Export ist bei großen Datenmengen sehr zeitaufwendig, kann aber unbeaufsichtigt über Nacht erfolgen.

5.7 Ausstattung, Zubehör, Durchlichteinheit

- Ausstattung und Zubehör ist durch das von Haus aus große Portfolio die große Stärke von *Zeutschel*. Diverse Halterungssysteme und Buchwippen, mit und ohne Glasplatte, sind mit *Zeutschel*-Scannern zu kombinieren.
- Aber auch *I2S* und *Cruse* verfügen über Saugtische und Buchwippen. *I2S* hat sogar einen automatischen Umblättermechanismus konstruiert. *ProServ* hat eine Buchwippe und eine automatisch wegfahrbare Glasplatte integriert.
- Der *Pentacon*-Scanner verfügt über einen Aluminiumkoffer, in dem er halbwegs sicher transportiert werden kann.
- *Homrich* verfügt ebenfalls über eine Vakuumpalte, eine Buchwippe und eine Buchschwinge, die Aufnahmen ohne störende Glasplatte erlauben. Besonders hervorzuheben ist der mit einem Handgriff höhenverstellbare Arbeitstisch und eine Durchlichteinheit mit Dia- und Glasplattenhalterung, welche dieses System als einziges der geprüften Geräte mit geringem technischen Aufwand für die Digitalisierung von Film tauglich macht.

5.8 Scanner-Software

- Die Software OS 11 von *Zeutschel* ist sehr intuitiv zu bedienen. Sie ist so gestaltet, daß auch ein Ungeübter schnell mit dem System vertraut werden kann. Bei großen Dateien neigte sie im Test zu gelegentlichen Abstürzen, was aber auch im Windows-Betriebssystem seine Ursache haben kann.
- *Pentacon* nutzt die schon seit vielen Jahren auf dem Markt befindliche Software Silverfast. Von allen Systemen bot sie dem Bearbeiter die beste Mischung aus Komfort

und Bearbeitungsmöglichkeiten. Besonders angenehm war das einfache Farbmanagement mittels IT8-Profilen.

- *Homrich* setzt auf den in der Grafikerwelt heimischen Apple-Macintosh und die Scansoftware Sinar-Capture-Shop. Der große Vorteil des Mac liegt in seiner Stabilität.² Weniger ausgereift ist dagegen das Capture-Shop selbst. Capture-Shop hat einen ähnlichen Funktionsumfang wie Silverfast, doch wäre eine ähnlich intuitive Steuerung wünschenswert gewesen. Eine gelegentliche „Out of memory“-Fehlermeldung zwang zum Neustart der Applikation.
- Die Software von *Cruse*, *ProServ* und *I2S* wurde nicht eigenhändig bedient und kann deshalb nicht repräsentativ bewertet werden. Alle drei Systeme wirken ähnlich intuitiv wie *Zeutschels* OS 11, *ProServs* Scannersoftware ist außerdem komplett über einen Internetbrowser steuerbar. Das macht dieses System sehr interessant für den Netzbetrieb oder den Betrieb an Linux-Rechnern.

5.9 Calibrierung, Farbmanagement, Wartung

- *Zeuschel* bietet lediglich die Möglichkeit über einen kleinen vormontierten Streifen einen Weißabgleich durchzuführen. Ein echtes Farbmanagement (beispielsweise über IT8-Profil) ist dem Benutzer nicht selbständig möglich. Das Farbmanagement wies im Test mit dem OS 10000 TT große Schwächen auf. Das Farbmanagement des im Werk getesteten OS 10000 A0 gab in dieser Hinsicht keinen Anlaß zur Kritik. Leider speichert *Zeuschel* die Bilder nur im sRGB-Farbraum.
- Der *Pentacon*-Scanner hatte Dank der Silverfast-Software das einfachste Farbmanagement mittels IT8-Profile. Die Speicherung der Bilder erfolgte in jedem beliebigen Farbraum, für den ein Referenzprofil vorlag. Fast unmöglich war hingegen die genaue geometrische Kalibrierung des Systems, die zudem nach jedem Einstellungswechsel neu durchgeführt werden mußte. Das Kameragehäuse ließ sogar eine Libelle oder ein Skalensystem zur Vermessung der optischen Achse vermissen.
- *Homrich* erlaubt dem Benutzer nicht nur ein eigenständiges Farbmanagement über GretagMacbeth-Colour-Charts, Grau- oder Weißabgleiche, sondern liefert auch ein Sinar Piezo-Calibration-Chart zur Justierung der Kamera mit aus. Die Speicherung der Bilder erfolgt im ECI-RGB-Farbraum.

² In der beinahe einjährigen intensiven Nutzung ist dieses System nicht ein einziges Mal abgestürzt oder hängengeblieben! Die Macintosh-Arbeitsumgebung baut zudem auf einem Unix-Derivat (FreeBSD) auf, was den Vorzug hat, daß ein solcher Arbeitsplatz sehr viel leichter in eine Unix/Linux-Umgebung zu integrieren ist, als ein Windows-System.

5.10 Schlußfolgerungen

Für das ARCHE-Projekt bestand nicht die wirtschaftliche Notwendigkeit, mit gering qualifizierten Arbeitskräften schnell Masse an Material zu digitalisieren. In diesem Fall wären die Zeilenscanner von *Zeutschel*, *ProServ* oder *I2S* sicherlich geeignet gewesen. Insbesondere der höhenverstellbare und austauschbare Scankopf des *I2S* hätten dieses System von den beiden anderen abgehoben.

Dem Anspruch der ARCHE-Anwendung wurden jedoch die Scanner von *Cruse* und *Homrich* sehr viel besser gerecht. Ihre Bilder haben einen erheblich höheren Schärfegrad. Hier geht es nicht primär um Menge, sondern um Qualität. Der von *Homrich* bereitgestellte Scanner hatte vor allem deshalb die Nase vorne, weil er sehr vielseitig einsetzbar und mit sehr wenig Aufwand für andere Anwendungen umbaubar war. Der Wechsel der Vorlagengröße geschah per Knopfdruck. Für großformatige Aufnahmen konnte der Tisch versenkt werden. Auch für die schonende Aufnahme von einzelnen Dias und Photoplatten wäre dieses System geradezu ideal gewesen. Die einzige nennenswerte Schwachstelle war die anspruchsvollere Bedienung. Mehr machen zu *können* wurde hier zu dem Preis erkauf, mehr machen zu *müssen*.

6 Der *Warenkorb* an Testvorlagen

Bereits im Herbst 2004 ist von Frau Dr. Happ unter Einbeziehung mehrerer mit dem Landesarchiv Baden-Württemberg und der Universität Stuttgart verbundener Einrichtungen ein *Warenkorb* an Vorlagen zusammengestellt worden. Dabei wurde großer Wert darauf gelegt, mit den Vorlagen ein möglichst breites Anforderungsspektrum abzudecken. Die Seitenformate variieren von kleiner DIN A5 bis hin zu Großformaten größer DIN A0. Der Warenkorb enthält Dias, Planvorlagen und Bücher. Bei der Auswahl wurde berücksichtigt, daß die Digitalisierung der vorgesehenen Materialien einen unmittelbaren Nutzen für die Eigner des Bestandes bringt. Auf diese Weise konnten verschiedene Aspekte des Workflows mit vielfältigen Anforderungen aus der Praxis abgeglichen werden.

6.1 Staatsarchiv Ludwigsburg

Bestand EL 221/8, Württembergisches Staatstheater Stuttgart, Dekorationsmappen

Inhalt: Dekorationsmappen mit Bühnen- und Kostümentwürfen der 1950er bis 1970er Jahre

Umfang: Der Bestand umfaßt ca. 300 Mappen mit jeweils 2 bis 41 Blättern (ø 15-20 Blätter) mit einem Gesamtvolumen von ca. 4500-6000 Vorlagen. Digitalisiert wurden 59 Mappen mit insgesamt rund 1500 Aufnahmen.

Vorlagengröße: Kleiner DIN A 5 bis DIN A 2, vereinzelt bis DIN A 0

Findbuch: online vorhanden, wurde für das Projekt erstellt
(<https://www2.landesarchiv-bw.de/ofs21/olf/struktur.php?bestand=17906>)

Bemerkung: Ein Teil dieser durch das Projekt gewonnenen Digitalisate konnte für die Ausstellung *Theaterbilder – Bildertheater* des Staatsarchivs Ludwigsburg (22. September 2005 bis 31. Januar 2006) genutzt werden.

Scanner: HIT Vario XL, überwiegend 1 Shot, 5440x4080 Bildpunkte, 48 Bit Farbtiefe



Abbildung 6.1: Die Digitalisierungsarbeiten im Institut für Erhaltung in Ludwigsburg

6.2 Generallandesarchiv Karlsruhe

Bestand GLA 67/1058, Pfälzisches Lehenbuch Friedrichs I.

Inhalt: Lehenbuch Friedrichs I., zwischen 1538 und 1542 angelegt

Umfang: ca. 500 Blatt

Vorlagengröße: etwas größer als DIN A 2

Findbuch: nicht vorhanden

Bemerkung: Das Buch wurde zeitgleich im Institut für Erhaltung von Archiv- und Bibliotheksgut restauriert, so daß die Einzelblätter aus der Bindung gelöst und plan aufgenommen werden konnten. Insgesamt wurden 530 Aufnahmen gemacht. Auf die Digitalisierung des ursprünglich ebenfalls für den *Warenkorb* vorgesehenen 400 Blatt umfassenden Pfälzischen Lehenbuchs Ludwigs V. (ca. 1470) wurde aus Gründen der Bestandserhaltung verzichtet.

Scanner: HIT Vario XL, 16 Shot, 10880x8160 Bildpunkte, 48 Bit Farbtiefe

6.3 Hauptstaatsarchiv Stuttgart

Bestand H 52, Illuminierte Urkunden

Inhalt: Zusammenstellung von Urkunden aus verschiedenen Beständen

Umfang: 37 Urkunden, ca. 70 Aufnahmen

Vorlagengröße: Die größte Urkunde mißt 66x79 cm

Findbuch: online vorhanden, wurde für das Projekt erstellt

(<https://www2.landesarchiv-bw.de/ofs21/olf/struktur.php?bestand=5147>)

Scanner: HIT Vario XL, 16 Shot, 10880x8080 Bildpunkte, 48 Bit Farbtiefe

Bestand J 56, Nachlass des württembergischen Offiziers Christian von Martens, (1793–1882)

Inhalt: Handschriften, Bilderalben, Tagebücher, Ölbilder, Aquarelle, Fotografien

Umfang: Weniger als 1 lfdm; 21 Archivalieneinheiten

Vorlagengröße: einzelne große Fotos und Gemälde bis DIN A 2, Alben bis ca. DIN A
4

Findbuch: online vorhanden

(<https://www2.landesarchiv-bw.de/ofs21/olf/struktur.php?bestand=5528>)

Bemerkung: Der Bestand enthält zahlreiche Bücher, die erhöhte technische Ansprüche an den Umgang mit den Materialien erforderten. Insgesamt wurden ca. 2000 Aufnahmen gemacht.

Scanner: HIT Vario XL, Überwiegend 1 Shot, 4 Shot (Aquarelle, Zeichnungen) 5440x4040 Bildpunkte, 48 Bit Farbtiefe

6.4 Universitätsbibliothek Stuttgart

Kartensammlungen

Inhalt: Pläne von Stuttgart, dem Neuen Schloß Stuttgart, vom Schloß Solitude, den Schloßanlagen und der Stadt Ludwigsburg, Entwürfe von Badeanlagen; teilweise von Balthasar Neumann; 18. & 19. Jh.

Umfang: ca. 20 Karten und Pläne

Vorlagengröße: verschieden, kleine Größen z.B. 29,5x77,2 cm oder 42,5x56,5 cm, Maximalgröße: 135,5x88,5 cm (ca. DIN A 0)

Findbücher: nicht erschlossen. Einzelne Pläne sind in der Literatur beschrieben.¹

Scanner: HIT Vario XL, 16 Shot, 10880x8160 Bildpunkte, 48 Bit Farbtiefe

Digitale Veröffentlichungen

Behrendt, Frank: *Experimentelle und numerische Beschreibung katalytischer Zünd- und Verbrennungsprozesse*. Stuttgart, Uni., Habil.-Schrift, 1999
(<http://elib.uni-stuttgart.de/opus/volltexte/1999/560/>)

Schultz, Anne-Catrin: *Der Schichtungsprozess im Werk von Carlo Scarpa*. Stuttgart, Univ., Diss., 1999
(<http://elib.uni-stuttgart.de/opus/volltexte/1999/514/>)

Seegers, Ulrike: *Transformatio energetica – Hermetische Kunst im 20. Jahrhundert : Von der Repräsentation zur Gegenwart der Hermetik im Werk von Antonin Artaud, Yves Klein und Sigmar Polke*. Stuttgart, Univ., Diss., 2002
(<http://elib.uni-stuttgart.de/opus/volltexte/2003/1284/>)

Bemerkung: Die Texte liegen auf dem OPUS-Server der Universitätsbibliothek Stuttgart im PDF-Format vor. Vor der ARCHE-Bildaufbereitung mußten sie mit den unter Linux verfügbaren Tools *pdftops* und *pstoimg* mit dem Umweg über das Postscript-Format in das PNG-Bildformat umgewandelt werden.

6.5 Universitätsarchiv Stuttgart

Nachlaß Georg Grube

Inhalt: Dokumente zu Lebensstationen, Vortrags- und Redemanuskripte, wissenschaftliche Dokumente, Fotos (u.a. zwei Farbfotos um 1912)

Umfang: 0,4 lfdm., ca. 500 Aufnahmen

Vorlagengröße: verschieden, DIN A 4, Folio, mehrere Urkundenplakate ca. DIN A 2

Findbuch: Findliste in FAUST-Datenbank

Scanner: HIT Vario XL, überwiegend 1 Shot, 5440x4080 Bildpunkte, 48 Bit Farbtiefe

¹ Siehe [12, Karten Nr. 28, 29, 30, 31, 54] und [13, Abbildung Nr. 27 und 28].

6.6 Institut für Leichtbau, Entwerfen und Konstruieren der Universität Stuttgart (ILEK)

**Kleinbilddias vom Deutschen Pavillon auf der Weltausstellung Montreal
1967 (Signaturen: 1.29-1.31.30, 1.62-1.66)**

Inhalt: Dias von der Planung, vom Aufbau und von der Ersteröffnung des Deutschen Pavillons auf der Weltausstellung Montreal 1967 und des Prototyps ILEK, Pfaffenwaldring 14

Umfang: ca. 700 Kleinbilddias

Scanner: Nikon Super Coolscan LS-4000 ED Diascanner, ca. 4000x3000 Bildpunkte, 48 Bit Farbtiefe

—

Insgesamt wurden ca. 6000 Aufnahmen gemacht und Daten mit einem Gesamtvolumen von 1,2 TByte gesammelt. Wie schon zuvor bei Material und Räumlichkeiten, wurde das ARCHE-Projekt während der umfangreichen Aufnahmearbeiten vom Staatsarchiv Ludwigsburg mehrere Monate lang auch durch Personal unterstützt.

7 Entwicklungsschritte der Bildaufbereitungssoftware

Innerhalb des Projekts konnte ein Prototyp der Bildaufbereitungssoftware entwickelt werden. Die Entwicklung der Software geschah nach dem *Wasserfallmodell*. Die Anforderungen wurden in Zusammenarbeit mit dem Landesarchiv und der Universitätsbibliothek Stuttgart erarbeitet. Die Implementierung wurde vom technischen Projektmitarbeiter der Universitätsbibliothek Stuttgart geleistet.

7.1 ArcheShell

In der ersten Entwicklungsphase wurde mit *ArcheShell* ein Kommandozeilentool programmiert. Es enthielt bereits den vollen bis dahin konzipierten Funktionsumfang. Es war aus einer Linux-Shell zu starten und benötigte eine JAVA-Laufzeitumgebung.

```
/ARCHEWSOFT/archeshell - Befehlsfenster - Konsole
Sitzung Bearbeiten Ansicht Lesezeichen Einstellungen Hilfe

Basic Film@STOPPED>
Basic Film@STOPPED>?
? - Hilfe
q - Programm beenden
----- Plugin -----
Bezeichnung: Basic Film
Beschreibung: Erzeugt einen Film mit einer Beschreibung, einem Inhaltsverzeichnis
und der Bildfolge.
Kommandos:
config <file> setzt die Konfigurationsdatei
prepare - bereitet alle Bestandsverzeichnisse auf das Digitalisieren vor
start - startet/setzt fort das Erzeugen der Bildfolge
stop - hält das Erzeugen der Bildfolge an
reset - setzt das Bestandsverzeichnis zurück
index - erzeugt eine Index-Datei in dem Ausgabeverzeichnis
vp - zeigt den Fortschritt
vo - zeigt aktuelle Einstellungen
Basic Film@STOPPED>
```

Abbildung 7.1: ArcheShell

7.1.1 Funktionsweise der ArcheShell

Nach dem Start von archeshell erhält der Benutzer eine Kommandozeile über die das Programm bedient werden kann. Zuerst muß die Projektdatei dem Programm übergeben werden. Die Projektdatei enthält die Pfade zu den einzelnen Bestandsverzeichnissen, die im bevorstehendem Digitalisierungsprojekt bearbeitet werden müssen sowie diverse Konfigurationsparameter, die das Aussehen der Frames beeinflussen. In jedem Bestandsverzeichnis erwartet das Programm mindestens zwei Dateien: eine Textdatei und eine CSV-Datei, welche die Metadaten des Bestandes beinhalten. Nach dem Setzen der Projektdatei können die Bestandsverzeichnisse mit dem Befehl **prepare** vorbereitet werden. Dabei erzeugt das Programm in jedem Bestandsverzeichnis eine Verzeichnisstruktur, die die Digitalisate speichern soll. Nach dem Einreihen der Digitalisate kann das Rendern mit dem Befehl **start** begonnen werden. Dabei wird eine Reihe von Frames und eine Indexdatei im Ausgabeverzeichnis erzeugt, die an den Belichter übergeben werden können. Das Rendering kann angehalten und das Programm beendet werden, um z.B. den Datenträger zu wechseln. Nach dem erneuten Programmstart und Setzen der Projektdatei kann der Prozess fortgesetzt werden.

7.1.2 Primäre Befehle der ArcheShell

config

Mit dem Befehl **config** wird die Projekt-Datei gesetzt.

prepare

Bereitet alle in der Projekt-Datei angegebenen Bestandsverzeichnisse auf das Digitalisieren vor.

Pre-Condition: Die Bestandsverzeichnisse enthalten lediglich je eine CSV-Datei, die den Bestand beschreibt.

Post-Condition: Die Bestandsverzeichnisse enthalten zusätzlich den kompletten Verzeichnisbaum, der gemäß der CSV-Datei aufgebaut ist. Die Digitalisate sollen während der Digitalisierung in die erzeugten Verzeichnisbäume eingereiht werden.

start

Startet das Erzeugen der Filmframes. Die Bestände werden nacheinander abgearbeitet. Die Reihenfolge, in der die Bestände abgearbeitet werden, gibt die Projekt-Datei an.

Pre-Condition: Die Bestandsverzeichnisse enthalten je eine CSV-Datei und je einen Verzeichnisbaum mit den Digitalisaten.

Post-Condition: Alle Bestände sind abgearbeitet. In dem Ausgabeverzeichnis befinden sich die Frame-Bilder und die dazugehörigen Vorschaubilder. Die Index-Datei ist erzeugt (oder aktualisiert).

Bemerkung: Wenn die Bestandsverzeichnisse noch nicht abgeschlossen sind, werden sie zuerst abgeschlossen.

stop

Hält das Erzeugen der Filmframes an. Dies kann notwendig sein um z.B. das Ausgabeverzeichnis zu leeren, um Platz für weitere Filmframes zu schaffen.

Pre-Condition: Die Verarbeitung läuft.

Post-Condition: Die Verarbeitung ist angehalten und kann jederzeit fortgesetzt werden. Index-Datei ist erzeugt (oder aktualisiert).

7.1.3 Sekundäre Befehle

Befehle, die normalerweise nicht benötigt werden, jedoch hilfreich bei der Fehlersuche sind.

finalize

Schließt alle in der Projekt-Datei angegebenen Bestandsverzeichnisse ab. Damit ist die Digitalisierung der Bestände abgeschlossen. Die nach dem Befehl "finalize" eingereihten Bilder werden nicht mehr berücksichtigt.

Pre-Condition: Die Bestandsverzeichnisse enthalten je eine CSV-Datei und je einen Verzeichnisbaum mit den Digitalisaten.

Post-Condition: Es wird eine xml-Datei generiert, wo die eingereihten Bilder mit den Metainformationen aus der CSV-Datei verknüpft werden.

ci (create index)

Erzeugt oder aktualisiert eine Index-Datei.

reset

Versetzt das Projekt in den Zustand vor **finalize**.

Post-Condition:

- Inhalt des Ausgabeverzeichnis ist gelöscht
- meta.xml's sind gelöscht
- Sämtliche Zähler sind auf 0 zurückgesetzt

7.2 ArcheGui

In einer zweiten Entwicklungsphase wurde zur leichteren Bedienung ein grafisches Frontend zur ArcheShell entwickelt. *ArcheGui* hatte zunächst den gleichen Funktionsumfang wie ArcheShell, wurde später aber kontinuierlich erweitert.

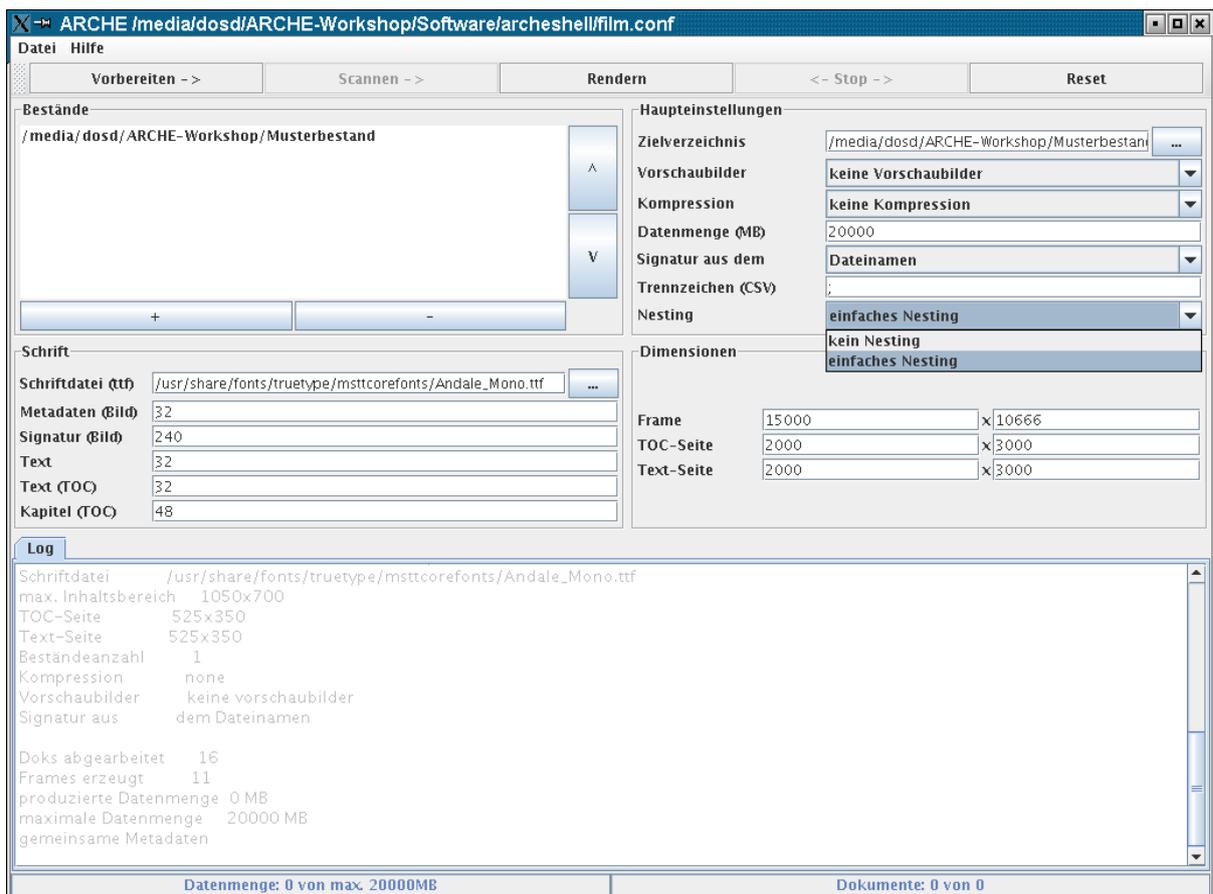


Abbildung 7.2: ArcheGUI

7.3 Metainput

Um dem Bearbeiter bereits bei der Digitalisierung die Aufnahme von bildbezogenen Metadaten zu ermöglichen, wurde ein flexibel konfigurierbarer Editor entwickelt. *Metainput* speichert die Benutzerdaten gleich in einer für die ARCHE-Bildaufbereitung notwendigen XML-Struktur.



Abbildung 7.3: Metainput

7.4 ArcheFilmcreator

Der *ArcheFilmcreator* stellt ein vollkommenes Redesign der ARCHE-Bildaufbereitungssoftware dar, welches größtenteils in der Projektverlängerung (Mai bis September 2006) ausgeführt wurde. Er vereint sämtliche Funktionalitäten von ArcheGui und Metainput und wurde darüberhinaus erheblich erweitert.

7.4.1 Funktionsweise des ArcheFilmcreators

Die Software ArcheFilmcreator arbeitet in drei Arbeitsschritten.

1. Bildbestände in das Programm importieren:
Es können mehrere Verzeichnisse in das Programm importiert werden. Die importierten Verzeichnisse können vom Nutzer sortiert werden, wodurch die Belichtungsreihenfolge für die Bestände festgelegt wird. Die Verzeichnisse werden rekursiv nach den Bilddateien durchsucht. Die gefundenen Bilder werden in der sogenannten „*natural order*“-Sortierreihenfolge verarbeitet. Die Importierung und Sortierung der Verzeichnisse erfolgt über die Werkzeugleiste des Navigationsbereiches.
2. Verarbeitungsparameter einstellen:
Die Verarbeitungsparameter werden über einen mehrstufigen Konfigurationsdialog eingestellt. Die einzelnen Parameter sind in einer Online-Hilfe ausführlich beschrieben.
3. Verarbeitung starten:
Die Verarbeitung wird über die Werkzeugleiste gestartet. Während des Vorgangs

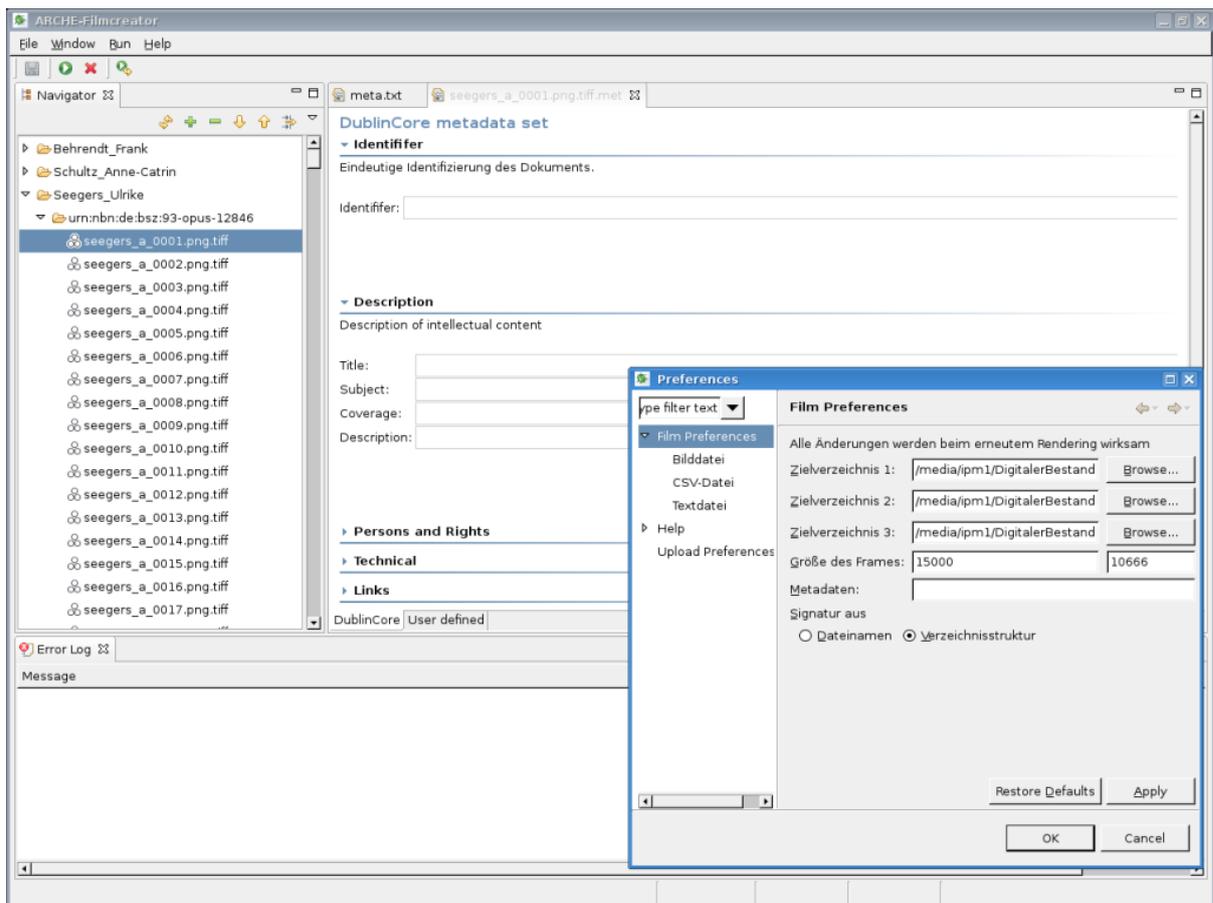


Abbildung 7.4: ArcheFilmcreator

werden die Warnungen und Fehlermeldungen in einem unteren Bereich des Programmfensters ausgegeben. Der Vorgang kann mit dem Cancel-Button des Fortschritts-Dialogs abgebrochen werden. Dabei wird die Verarbeitung nicht sofort abgebrochen, sondern erst nachdem die Generierung des aktuell bearbeiteten Frames abgeschlossen ist. Nach einem Abbruch kann die Verarbeitung jedoch wieder fortgesetzt werden. Die Werkzeugleiste hält auch Möglichkeiten zur Säuberung der Ausgabeverzeichnis bereit.

8 Zeitplan des Teilprojekts

August 2004 Einstellung der Projektmitarbeiterin Dr. Sabine Happ (inhaltliche Betreuung).

Oktober 2004 Einstellung des Projektmitarbeiters Wladimir Schwitin (technische Realisierung).

23./24. November 2004 Projekttreffen in der Universität Frankfurt. Persönliche Vorstellung aller Projektpartner, Festlegung der Rahmenbedingungen für die Konzeption und Entwicklung der einzelnen Projektbausteine.

Dezember 2004 Einstellung des Projektmitarbeiters Klaus Wendel (inhaltliche Betreuung, Nachfolge von Dr. Sabine Happ).

Oktober bis März 2005 Konzeption der ARCHE-Software und des Workflows

Januar 2005 Einrichtung eines geschlossenen Internetforums für die Projektkommunikation. Im Juni 2005 wurde dieses in ein „Wiki“ mit angeschlossenem Fileserver umgewandelt und kontinuierlich erweitert.

Februar 2005 bis Mai 2006 Parallel zum Projekt wurde kontinuierlich die Digitalisierung der Vorlagen betrieben. Wobei der Großteil der Digitalisierungsarbeiten erst ab **Juni 2005** durchgeführt wurde, nachdem in einem Raum des Instituts für Erhaltung eine Digitalisierungswerkstatt eingerichtet und der Reproscanner der Firma Homrich installiert worden war. Von **Februar bis Mai 2005** wurden vor allem die verschiedenen Scannersysteme getestet und zahlreiche praktische Versuche zur Foto- und Digitalisierungstechnik und dem Farbmanagement durchgeführt.

24./25. März 2005 Planungstreffen mit Projektpartnern im IPM, Freiburg.

Juni 2005 Bereitstellung einer ersten Version der *ArcheShell*.

August 2005 Entwicklung eines Prototyps der komfortableren Arbeitsumgebung *ArcheGui* und des Metadateneingabewerkzeugs *Metainput*. Zusätzlich wurde ein für den Workflow nicht relevantes, jedoch die Funktionalität der Scannersoftware *CaptureShop* erweiterndes Tool (*ShadingHack*) entwickelt, wodurch ein wichtiger Kalibrierungsprozess bei der Digitalisierungsarbeit erheblich beschleunigt werden konnte. Danach kontinuierliche Erweiterung des Funktionsumfangs von *ArcheGui* und *ArcheShell*.

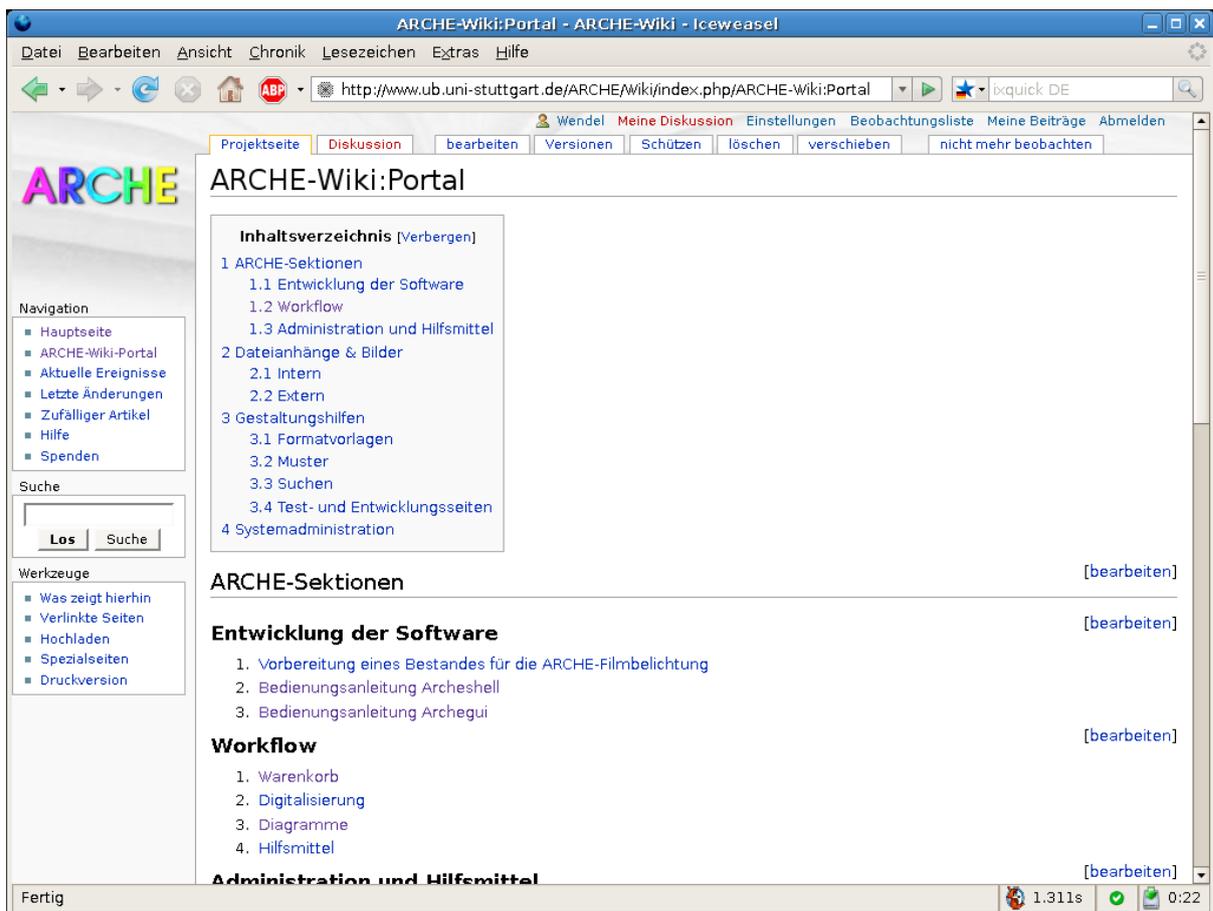


Abbildung 8.1: Das ARCHE-Wiki als zentraler Informationsspeicher

15./16. September 2005 Teilnahme an der *International Conference on Preservation of Digital Objects* (iPRES) in Göttingen.

27.-30. September 2005 Am Stand des Landesarchivs Baden-Württemberg wurde auf der dem 75. Deutschen Archivtag in Stuttgart angeschlossenen Fachmesse ARCHIVISTICA 2005 mit Plakaten und einer Powerpoint-Präsentation über das Projekt informiert.

7./8. Dezember 2005 Vorstellung der ARCHE-Bildaufbereitung auf einem projektinternen Workshop im IPM, Freiburg.

21./22. Februar 2006 Präsentation der ARCHE-Bildaufbereitung auf einem Workshop vor möglichen Kunden und Interessenten der ArchiveLasers[®] im IPM, Freiburg.

13.-15. März 2006 Präsentation der ARCHE-Bildaufbereitung am Stand des Industriepartners ArchiveSystems auf der Messe Cebit 2006 in Hannover.

24. März 2006 Vortrag über das ARCHE-Projekt am 95. Bibliothekartag in Dresden.¹

März-Juni 2006 Entwicklung zur eigenständigen Software *ArcheFilmcreator*.

6. Mai 2006 Der bundesweit durchgeführte Tag der Archive lockt zahlreiche Besucher in das Staatsarchiv Ludwigsburg und das Institut für Erhaltung. Die Digitalisierungswerkstatt des ARCHE-Projekts sowie das ARCHE-Konzept zur Bestandserhaltung erfreute sich sehr regen Interesses.

27. Juni 2006 Abschlußkolloquium des ARCHE-Projekts vor ca. 150 Interessenten in der Universitätsbibliothek Stuttgart.²

Mai-August 2006 Testphase des *ArcheFilmcreators*. Die funktionalen Tests, in denen der zuvor eingescannte Bestand verarbeitet wurde, wurden erfolgreich abgeschlossen und die vorbereiteten Bestände an Fraunhofer IPM zur Belichtung übergeben.

August/September 2006 Der Prototyp der Software wurde in Absprache mit der Firma MAS getestet; hierbei wurde die Software einem von MAS genannten potentiellen Anwender zur Verfügung gestellt, und es wurden Fehlermeldungen und Änderungsvorschläge eingearbeitet.

30. September 2006 Offizielles Ende des Projektes.

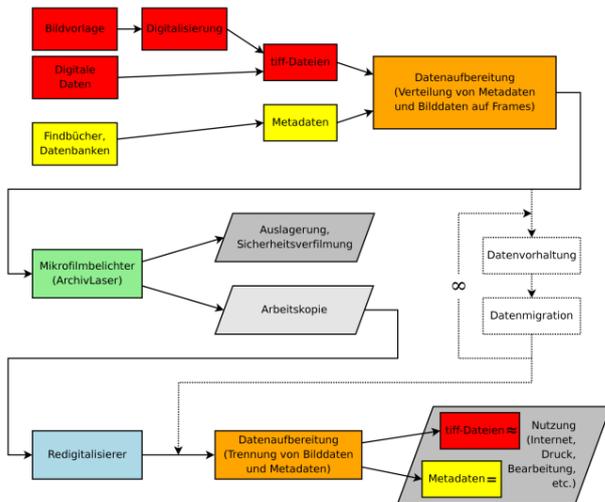
¹ Siehe [9].

² Siehe [10].



Abbildung 8.2: Das erste ARCHE-Plakat (2005), Vorderseite

ARCHE



Die Partner

Das Projekt ARCHE ist ein Gemeinschaftsunternehmen des Landesarchivs Baden-Württemberg, der Universitätsbibliothek Stuttgart, der Firma MicroArchive Systems und weiteren Industriepartnern unter der Leitung des Fraunhofer-Instituts für Physikalische Messtechnik IPM in Freiburg. Es wird vom Bundesministerium für Wirtschaft und Arbeit über das Programm "Förderung von innovativen Netzwerken" (InnoNet) gefördert.

Technische Daten des ArchivLasers

Framegröße	32 x 45 mm (35 mm Film)
Pixelgröße	3 µm (160 Linienpaare / mm)
Pixel / Frame	10,666 x 15,000 Pixel
Farbtiefe	36 (3 x 12) Bit
Filmmaterial	S/W oder Color-Mikrofilm (Ilford, Kodak, Agfa)
Belichtungszeit	40 sec / Frame
Filmlänge	≤ 600 m (13,000 Frames)
Zoomfaktor	34 (1 DIN A0 Blatt / Frame)
Kapazität	16 Blätter DIN A4 / Frame bzw. 1 GB / 25 cm Film

Das Projekt ARCHE

Zielstellung ist die Entwicklung eines Workflow und der dazugehörigen technischen Lösung für die Ausbelichtung digitaler Dokumente auf alterungsbeständigem und hochauflösendem Farbmikrofilm sowie die Redigitalisierung verfilmter Daten.

Unterstützt werden dadurch Fachaufgaben wie die Bestandserhaltung und die verbesserte Nutzung von Kulturgut sowie die Langzeitarchivierung digitaler Dokumente in Archiven, Bibliotheken und anderen Institutionen.

Die Prozesskette zeigt die vielfältigen Aspekte des ARCHE-Projekts von der Digitalisierung, über die lasergestützte Ausbelichtung, bis hin zur Redigitalisierung und Nutzung der gespeicherten Informationen.

Funktionsweise des ArchivLasers

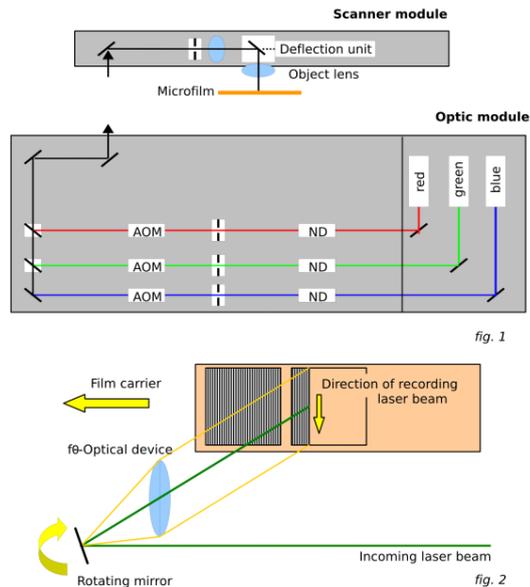
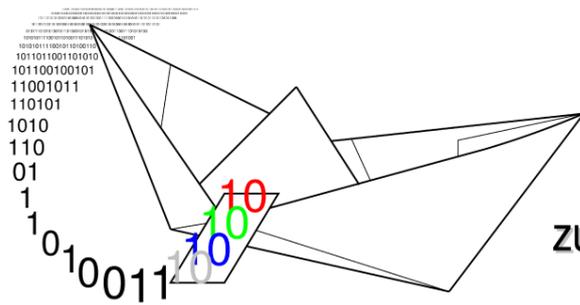


Abbildung 8.3: Das erste ARCHE-Plakat (2005), Rückseite



Eine **ARCHE** zur Rettung digitalen Kulturguts

Stuttgarter Kolloquium des ARCHE-Projekts

am
27. Juni 2006, 13:30 bis 16:30 Uhr
im
Vortragssaal der Universitätsbibliothek Stuttgart

Die Veranstaltung dokumentiert die im Rahmen des Projekts geleistete Entwicklung eines Laserbelichters zur Langzeitarchivierung digitaler Bilddaten auf alterungsbeständigem Farbmikrofilm mit sehr hoher Genauigkeit und Farbtreue.

Programm

Begrüßung durch Werner Stephan, Direktor der Universitätsbibliothek Stuttgart

Dr. Gerald Maier, Landesarchiv Baden-Württemberg
„Strategische Ziele des ARCHE-Projekts für Archive, Bibliotheken und Museen“

Andreas Hofmann, Fraunhofer Institut für Physikalische Messtechnik
„ArchiveLaser®: Technik der Laserbelichtung von Farbmikrofilm mit Vorstellung von Belichtungsbeispielen“

Klaus Wendel, Universitätsbibliothek Stuttgart
„Der Workflow zur Kombination digitaler Bild- und Metadaten auf einem optischen Medium“

David Gubler, MicroArchive Systems
„Möglichkeiten der Nutzung des ArchiveLaser®s“

Diskussion

www.ipm.fraunhofer.de • www.landesarchiv-bw.de/arche • www.ub.uni-stuttgart.de/arche • www.microarchive.de



Abbildung 8.4: Das zweite ARCHE-Plakat (2006)

Literaturverzeichnis

- [1] *DIN-Taschenbuch 153 – Publikation und Dokumentation 1 : Gestaltung von Veröffentlichungen, Terminologische Grundsätze, Drucktechnik, Alterungsbeständigkeit von Datenträgern : Normen.* Berlin : Beuth, ⁴1996
- [2] *DIN-Taschenbuch 154 – Publikation und Dokumentation 2 : Erschließung von Dokumenten DV-Anwendungen in Information und Dokumentation, Reprographie, Photographie, Mikrofilmtechnik, Bibliotheks- und Verlagstechnik : Normen.* Berlin : Beuth, ⁴1996
- [3] HARTMUT WEBER (Hrsg.): *Bestandserhaltung in Archiven und Bibliotheken.* Stuttgart: Kohlhammer, 1992 (Werkhefte der Staatlichen Archivverwaltung Baden-Württemberg : Serie A, Landesarchivdirektion ; H. 2)
- [4] HARTMUT WEBER (Hrsg.): *Digitale Archive und Bibliotheken : neue Zugangsmöglichkeiten und Nutzungsqualitäten.* Stuttgart : Kohlhammer, 2000 (Werkhefte der Staatlichen Archivverwaltung Baden-Württemberg : Serie A, Landesarchivdirektion ; H. 15)
- [5] GERALD MAIER / THOMAS FRICKE (Hrsg.): *Kulturgut aus Archiven, Bibliotheken und Museen im Internet : neue Ansätze und Techniken.* Stuttgart : Kohlhammer, 2004. (Werkhefte der Staatlichen Archivverwaltung Baden-Württemberg : Serie A, Landesarchivdirektion ; H. 17)
- [6] STEFAN ROHDE-ENSLIN: *Nicht von Dauer : Kleiner Ratgeber für die Bewahrung digitaler Daten in Museen.* Berlin, 2004 (Materialien aus dem Institut für Museumskunde ; Sonderheft 2)
- [7] Produktseite des Fraunhofer Instituts für Physikalische Messtechnik (IPM)
http://www.ipm.fraunhofer.de/fhg/ipm/anwendungen_maerkte/laserbelichtung/archivlaser/index.jsp
- [8] Internetseite der Firma MicroArchive Systems GmbH
<http://www.microarchive.com/>
- [9] Eine ARCHE zur Rettung digitalen Archivguts - Langzeitarchivierung digitaler Daten auf laserbelichtetem Farbmikrofilm (Vortrag gehalten auf dem 95. Deutschen Bibliothekartag in Dresden am 24.3.2006)
<http://elib.uni-stuttgart.de/opus/volltexte/2006/2738/>

- [10] Eine ARCHE zur Rettung digitalen Kulturguts - Stuttgarter Kolloquium des ARCHE-Projekts am 27. Juni 2006
<http://elib.uni-stuttgart.de/opus/volltexte/2006/2731/>
- [11] KLAUS WENDEL: *Eine ARCHE zur Rettung digitalen Archivguts*. In: Archivnachrichten / Landesarchiv Baden-Württemberg, Nr. 32 (2006), S. 18f
- [12] *Stuttgart im Spiegel alter Karten und Pläne : Ausstellung des Hauptstaatsarchivs Stuttgart ; Katalog / bearb. von Jürgen Hagel. Hrsg. vom Württembergischen Geschichts- und Altertumsverein*. Stuttgart, 1984
- [13] *Die Gärten der Herzöge von Württemberg im 18. Jahrhundert : Katalog zur Ausstellung in Schloss Ludwigsburg vom 4.7. bis 1.11.1981 / bearbeitet von Andrea Berger-Fix und Klaus Merten*. Worms, 1981