

Gesamtschweizerische Strategie zur dauerhaften Archivierung von Unterlagen aus elektronischen Systemen (Strategiestudie)

Anhang 3: Standards

Vorbemerkung: Dieses Dokument stellt ein Nebenprodukt der Arbeiten an der Strategiestudie dar. Es ist unvollständig, wurde nicht lektoriert und soll periodisch aktualisiert werden.

A) Anwendungsstandards

Die Anwendungsstandards entsprechen den bereits in Kapitel 3.1.3 der Strategiestudie angeführten Problembereichen. Im Vordergrund stehen

- Verzeichnungsstandards;
- Präsentationsstandards;
- Übernahmestandards;
- Aufbewahrungsstandards;
- Schriftgutverwaltungsstandards;
- Retrievalstandards sowie
- Standards für Spezialgebiete.

Nachstehend werden die insgesamt wichtigsten Standards zusammengestellt:

Standard Akronym	Vollst. Titel	Verantwortlich	Beschreibung / Bemerkung
ISAD(G)	International Standard Archival Description, General Rule	International Council on Archives (ICA)	Standard zur Verzeichnung von Archivgut (-Metadaten)
ISAAR+(CPF)	International Standard Archival Authority Records: Corporate Bodies, Persons, Families	International Council on Archives (ICA)	Standard zur Verzeichnung von Sekundärwortschätzen zu Archivgut (primär: unterlagenbildenden Stellen, evt. auch betroffenen Personen, Organisationen usw. in Registern)
DC	Dublin Core Metadata for Simple Resource Discovery	Online Computer Library Center (OCLC) and IETF	Metadaten Set zur Beschreibung von Informationsquellen im Internet und anderen Netzwerken. 15 Elemente: TITLE, CREATOR, SUBJECT, DESCRIPTION, PUBLISHER, CONTRIBUTOR, DATE, TYPE, FORMAT, IDENTIFIER, SOURCE, LANGUAGE, RELATION, COVERAGE RIGHTS Internet RFC 2413 als Implementation dieses Standards

Standard Akronym	Vollst. Titel	Verantwortlich	Beschreibung / Bemerkung
EAD	Encoded Archival Description	Network Development and MARC Standards Office of the Library of Congress (LC) in partnership with the Society of American Archivists	Darstellungs-Standard und –Format für Archiv-Findmittel (Finding Aids) auf der Basis von ursprünglich SGML sowie heute XML 1992–1994 an der Berkeley University entwickelt; verwendbar für die Präsentation von Archivgut-Metadaten, beruhend auf ISAD(G)
OAIS	Open Archiving Information System	Consultative Committee for Space Data Systems (CCSDS)	Referenzmodell für die technische Implementierung von Systemen zur Verwaltung von Archivobjekten beliebiger Beschaffenheit, v.a. in Bezug auf Long-Term Archival Information Package (AIP) Preservation Vorprodukt eines ISO-Standard; ursprünglich entwickelt für Archivdaten von Raumfahrtbehörden, heute für alle Interessierten offen
RMA	Design Criteria Standard for Electronic Records Management Software Applications	U.S. Department of Defense Records Management Task Force (RMTF)	Ergebnis in DoD 5012.2 Design Criteria Standards for Electronic Records Management Software Applications
ISO-15489	Standard for Managing Business Records	International Organization for Standardization (ISO)	Allgemeine internationale Norm für eine ordnungsgemäße aktengestützte Geschäftsführung, bestehend aus einem allgemeinen und einem technischen Teil (freigegeben 2002)
AS4390	Australian Standard for Records Management	IT/21 Committee of Standards Australia	National standard for management of physical and electronic records; Basis für und Vorläufer von ISO-15489

Verzeichnungsstandards – ISAD(G) / ISAAR(CPF) / DC

Für das Verzeichnen von Archivgut (d.h. die Verwaltung von Metadaten oder archivierte Unterlagen) stehen im Vordergrund

- der in Anlehnung an den bibliothekarischen Katalogisierungsstandard ISBD geschaffene und 1994 in Kraft gesetzte internationale Archivgut-Verzeichnungsstandard **ISAD(G)** („International Standard Archival Description, General Rule“);
- der zu ISAD(G) komplementäre, 1996 erlassene Standard **ISAAR(CPF)** („International Standard for Authority Records: Corporate Bodies, Persons, Families“) für die Ansetzung von Referenzdateien zu den Urhebern von Unterlagen (die mit Einträgen zu Regel 3.2.1 von ISAD(G) übereinstimmen sollen);
- der **DC** („Dublin Core“) als Minimal-Set bibliografischer Angaben zu beliebigen Informationsressourcen im Internet;
- der **EAD**-Standard („Encoded Archival Description“) für die Präsentation von Archivfindmitteln im Internet (vgl. unten), dessen Gliederung teilweise mit derjenigen des ISAD(G) übereinstimmt, aber mehrere (maximal 9) Verzeichnungsstufen umfassen kann.

Ein zentrales Merkmal des ISAD(G)-Regelwerks ist der Grundsatz der mehrstufigen Verzeichnung, während der ISAAR(CPF) die entstehungskontextbezogenen „access points“ bereitstellt. Andere „access points“ beziehen sich auf die Inhalte von Archivgut und können beispielsweise Schlagwörter oder Stichwörter („keywords“, „index terms“) zu Personen und Personengruppen, Orts- oder Sachbegriffen usw. enthalten.

Präsentationsstandard – EAD

Aus der Welt der (oft in Bibliotheken angesiedelten) Handschriftensammlungen stammt der *EAD*-Standard („Encoded Archival Description“), der zu Beginn der 1990er Jahre als Nachfolger des für mehrstufige Beschreibungen mangelhaft geeigneten MARC-Standards (Machine Readable Cataloguing) an der Berkeley University entwickelt wurde. Wegweisend für diese Norm wurde der technologische Ansatz, Textdokumente mittels einer Textbeschreibungssprache (zuerst SGML, heute eher XML) zu strukturieren. EAD ist somit ein Verzeichnungsstandard (der bis zu einem gewissen Grad den ISAD(G) konkurrenziert, siehe oben), primär aber ein Präsentationsstandard, da mit seiner Hilfe herkömmliche Archivfindmittel („Finding Aids“) unabhängig vom jeweiligen Texteditor mittels Stylesheets abgebildet werden können. XML beschränkt sich überdies längst nicht mehr auf Textdokumente, sondern wird auch zur Präsentation von Datenkonstrukten aus Datenbank-Applikationen oder als Übernahmestandard beim Austausch zwischen verschiedenen Applikationen verwendet.

Vergleichbare Präsentationsstandards gibt es für wissenschaftliche Dokumente (TED), aber auch zum Beispiel für Vermessungsdaten (in der Schweiz mittels der Beschreibungssprache INTERLIS). Es ist davon auszugehen, dass weitere derartige Standards (nicht nur auf der Basis von XML, sondern auch anderer vergleichbarer Sprachen wie z.B. TEX/LATEX usw., vgl. Kap. 3.1.4.2: Textformate) für besondere Textgruppen entstehen werden.

Ob sich das EAD-Format auch zur Darstellung elektronisch erzeugter Unterlagen eignet, ist noch ungewiss. Wegweisend ist in jedem Fall die verwendete Technik: Mittels Document Type Definitions (DTDs) lassen sich bei Bedarf weitere, unterschiedlich strukturierte Findmittel oder Metadatensammlungen zu Archivobjekten definieren. Mittels XML-Schemata ist es heute zusätzlich möglich, nicht nur die logische Struktur, sondern auch die zugelassenen Inhalte (Orthografie, Semantik, Syntax usw.) von Findmitteln festzulegen. Nebst einer DTD kann ein XML Schema beliebige Stylesheets zur Speicherung der Präsentation von Daten am Bildschirm oder auf Ausdrucken usw. enthalten.

Übernahmestandards

Es fällt umso einfacher, gegenüber den Unterlagenproduzenten die Berücksichtigung archivi-scher Bedürfnisse zu verlangen, je klarer die Übernahmekriterien definiert sind. Im OAIS-Modell (vgl. unten: Aufbewahrungsstandards) enthält eine sogenannte „Submission Information Package“ (SIP) diejenigen Kontext-Metadaten, die ein Produzent von Archivgut bei dessen Übergabe an ein Archivmanagementsystem mitliefern soll. Konkret implementierte Übernahmestandards gibt es heute erst vereinzelt und ansatzweise. Dies hängt nicht zuletzt damit zusammen, dass die wenigsten Archive bisher über eine technische Ausstattung verfügen, die eine regelmässige Übernahme von Archivgut überhaupt gestattet.

In der angelsächsischen Archivwelt zielte die Strategie des „non-custodial archiving“ (vgl. Kapitel 3.1.6.1 der Strategiestudie) auf eine zeitlich unbeschränkte Lagerung von Archivdaten in den jeweiligen operativen Systemen. Erst seit kurzer Zeit hat auch in diesen Ländern ein Umdenken stattgefunden.

Einen sehr einfachen, örtlich beschränkten und nicht verallgemeinerbaren Übernahmestandard stellt die Spezifikation des DOMEA-Aussonderungsmoduls, im Hinblick auf Datenübergaben an das deutsche Bundesarchiv, dar. In der Schweiz bestehen verschiedene Lösungsansätze, die im Projekt ARELDA des Bundesarchivs (Kapitel 4.5.4 der Strategiestudie), im Rahmen der scopeArchiv User Group (Kapitel 4.6.4 E) sowie zwischen den Produkten Konsul und Star II (Kapitel 4.6.4 D) entweder konzipiert, spezifiziert oder bereits realisiert worden sind. In all diesen Fällen kann noch nicht von einem Standard gesprochen werden.

Schriftgutverwaltungsstandards

Verhältnismässig weit fortgeschritten sind Standards im Bereich des Records Management / Recordkeeping (Schriftgutverwaltung, gestion des documents):

- Auf der Basis des UBC-Projekts (vgl. Kapitel 3.1.3) erliess das US-Verteidigungsministerium 1997 den **DoD 1505-2-Standard**, der die wichtigsten Anforderungen an das Design von Applikationen im Bereich des Records Management beschreibt. Auf dessen Grundlage wurden seither zahlreiche Produkte zertifiziert.
- **GEVER** („Anforderungen an Geschäftsführungssysteme“) ist eine Strategie und ein Vorgehenskonzept zur Förderung aktengestützter Geschäftsführungssysteme in der Schweizerischen Bundesverwaltung (vgl. Kapitel 4.4.2, 4.4.6).
GEVER wird normativ durch eine Reihe von gesetzlichen bzw. in Verordnungen und Weisungen festgehaltenen Bestimmungen abgestützt (vgl. Kapitel 4.2.4). Auf seiner Grundlage wurde das Projekt GBL (GEVER-Basislösung) durchgeführt. Dieses verfolgte das Ziel, ein eGovernment-taugliches System möglichst flächendeckend einzuführen. 2001 wurden die GEVER-Anforderungen zu einem produkteneutralen Standard erweitert.
- **DOMEA** ist ein Standard für die deutsche Bundesverwaltung zur elektronischen Unterstützung von Verwaltungsvorgängen gemäss den Schriftgutverwaltungsnormen der GGO. DOMEA wurde in Zusammenarbeit mit dem deutschen Bundesarchiv und dem Institut für Verwaltungsinformatik an der Universität Koblenz erarbeitet. Es beruht auf Akten, die einem Ordnungssystem (Registraturplan) zugeordnet werden müssen und aus Vorgängen bestehen, denen wiederum Dokumente zugeordnet werden. Zu DOMEA gehört ferner ein Aussonderungskonzept für die Übergabe von Metadaten und Dateien an das System des Bundesarchivs. Der Standard wird inzwischen auch in mehreren Bundesländern als verbindlich anerkannt. Rund ein halbes Dutzend Produkte haben sich bisher die DOMEA-Zertifizierung erhalten.
- **ELAK** („Elektronischer Akt“) ist ein Anforderungskatalog an Schriftgutverwaltungssysteme, der im Rahmen des österreichischen landesweiten eGovernment-Programms erarbeitet und 2001 verabschiedet wurde.
Die Verantwortlichen für die drei deutschsprachigen Standards DOMEA, ELAK und GEVER haben 2001 damit begonnen, ihre Aktivitäten zu koordinieren.

- **NOARK** ist ein Schriftgutverwaltungs-Standard für die norwegische Reichsverwaltung, der erstmals 1984 erlassen wurde und inzwischen in vierter Version vorliegt. Dieser sieht vier unterschiedliche Ausbaustufen vor. Die Ablagestruktur kennt, wie DOMEA und GEVER (Kap. 4.4.6), die Komponenten „Ordnungssystem“, „Akte“, „Dokument“ sowie weitere transaktionsabhängige Kontext-Metadaten.
- Der australische Records management-Standard **AS 4390** wurde 1998 in Kraft gesetzt. Er besteht aus sechs Teilen (1: General, 2: Responsibilities, 3. Strategies, 4: Control, 5: Appraisal and Disposal, 6: Storage) und diente als Ausgangspunkt für weitergehende Standardisierungen in Australien – vor allem der Recordkeeping Metadata Standard for Commonwealth Agencies (1999) und das Recordkeeping Metadata Scheme (RKMS) – sowie den ISO Standard 15'489.
- Der **ISO 15'489**-Standard der International Standards Organization wurde 1998 auf Initiative australischer Records Manager initiiert, dann aber stark auf die zum Teil erheblich anders gelagerten Bedingungen europäischer Schriftgutverwaltungen angepasst. Er besteht aus einem allgemeinen und einem technischen Teil, der verschiedene Beispiele als Codes of best practice für verschiedene Länder enthält.
- Der von einem Gremium aus EU-Ländern als Bestandteil von Folgeaktivitäten des DLM-Forums erarbeitete **MOREQ-Standard** spezifiziert technische Anforderungen an Electronic Records Management Systems (ERMS). Da verschiedene nationale Traditionen unterstützt werden müssen, bleiben die Metadaten-Anforderungen in gewissen grundlegenden Fragen (aufgabenbezogener hierarchischer Aktenplan oder inhaltsorientierter Thesaurus) etwas unbestimmt. Davon abgesehen bildet der Katalog aber eine nützliche Ergänzung der zuvor genannten Normenwerke.

Insgesamt ergibt sich aus den besprochenen Standards, zu denen auch der GEVER-Standard des schweizerischen Bundesarchivs zu zählen wäre, ein hoher Grad an Konvergenz. Es könnte nützlich sein, eine abgestufte Synthese für verschiedene Anspruchsniveaus in einer möglichst einfachen Formulierung für schweizerische Bedürfnisse zusammenzustellen.

Im Fluss befindet sich die Frage, wie weit die Aufzeichnungen von Geschäftsprozessen, die durch unterschiedlichste Technologien (Datenbankapplikationen, Workflow, Groupware Document Management Systems usw.) unterstützt werden, systematisch in solche recordkeeping systems übernommen werden können (Problem des „records capturing“ nach D. Bearman).

Metadaten-Standards

Metadaten sind Daten, die andere Daten beschreiben. Einen Versuch, sämtliche zur Beschreibung elektronischer Unterlagen relevanten Metadaten in ihren gegenseitigen Beziehungen und Abhängigkeiten zu beschreiben, stellt das australische *Recordkeeping Metadata Scheme* (RKMS) dar. Dieses beruht einerseits auf der australischen Archivverzeichnungs-Tradition des „Series“-Prinzips; andererseits abstrahiert es verschiedene bestehende Metadaten-Sammlungen wie zum Beispiel Dublin Core, ISAD (G), EAD, AGLS¹, Pittsburgh Requirements usw.

¹ GILS („Government Information Locator System“ und AGLS („Australian Government Locator System“) sind Verzeichnisdienste für den Zugriff auf öffentliche Informationsressourcen unterschiedlicher Art. Die in diesem

Das RKMS beschreibt die Anforderungen an den Nachweis von

- Aufgaben und Geschäftsprozesse bis zu Systemtransaktionen („business“, „functions“ etc.)
- Unterlagen (Dokument-Kollektive, Dokumente, Dokument-Teile; „records“);
- Akteure (Organisationen bis zu Einzelpersonen; „actors“)
- Schriftgutverwaltungsdaten („business recordkeeping entities“).

Diese vier Entitäten müssen in unbeschränkter Verschachtelungstiefe und in beliebigen gegenseitigen Beziehungen abgebildet werden können. Der RKMS-Standard wurde in den Staaten Victoria (vgl. VERS) und New South Wales weiter verfeinert (vgl. Anhang 2).

Dokumentaustausch-Standards – Edifact / ODA-ODIF

EdiFact und ODA-ODIF repräsentieren internationale Normen für den Austausch digitaler Dokumente, die noch vor dem Durchbruch des Internets entwickelt wurden und vor allem im kommerziellen Geschäftsalltag Verwendung finden sollten. Sie haben sich jedoch nicht in grösserer Breite durchgesetzt.

Im Augenblick macht es den Anschein, als ob sich XML-Schemata als geschäftsspezifische Austauschformate profilieren könnten.

Aufbewahrungsstandards – OAIS

Das zeitlich unbeschränkt haltbare digitale Archivformat ist noch nicht geboren. In der Praxis werden technische Standards (vgl. Kap. 3.1.4.2) für Aufbewahrungszwecke vor einem Zeithorizont von maximal 10 bis 15 Jahren verwendet. Was darüber hinausreicht, ist vorerst noch Gegenstand von Forschungsprojekten, die für den Alltag der Archive bis auf weiteres nicht relevant sein können.

Einem interessanten Vorgehensansatz folgt das seit 1997 vom Consultative Committee for Space Data Systems (CCSDS) entwickelte *Reference Model for an Open Archiving Information System* (OAIS), das gemäss bekannt gegebenem Fahrplan bis September 2001 zu einem ISO-Standard erhoben wird.² Massgeblich für diese Initiative war zunächst der Bedarf der amerikanischen Luftfahrtsbehörde (NASA) nach raumbezogenen Archivdaten; die Standardisierungsarbeit wurde aber von früh an auf beliebige Unterkategorien auszuweiten versucht. Im Rahmen einer solchen Erweiterung fanden mehrere Workshops mit der NARA (National Archives and Records Management Administration), der Research Library Group (RLG) und weiteren Gremien statt. Auch die Metadata Requirements for Evidence des Pittsburgh Project wurden berücksichtigt.³

Rahmen verwendeten Metadata-Standards (oft DC, im AGLS in Verbindung mit RKMS) sind für den Aufbau von e-government-Wissensstrukturen unentbehrlich.

² Vgl. das Redbook vom Juli 2001 unter der Bezeichnung CCSDS 650.0-R-1; <http://www.ccsds.org/documents/pdf/CCSDS-650.0-R-2.pdf>

³ Das OAIS-Konzept beruht auf einem Umgebungsmodell, das zwischen Produzenten, Konsumenten und Managern von Archivgut unterscheidet. Um Missverständnisse beim Gebrauch archivischer Zentralbegriffe wie „ar-

Aufgrund der verwendeten Methodologie erscheint das OAIS-Modell wie geschaffen für eine technische Implementierung von Emulations-Modellen (obwohl das Dokument selbst darauf keinerlei Bezug nimmt). Jedenfalls wird es von den bibliothekarischen „digital preservation“-Projekten so verstanden (vgl. z.B. CAMILEON, NEDLEB). Der generische OAIS-Ansatz ist aber auch für klassische Migrations-Verfahren (vgl. Kap. 3.1.6.1) verwendbar.

Offen bleibt in jedem Fall bis auf weiteres, wie die von OAIS verwendeten Begriffe mit den in IT-Projekten gängigen Terminologien in Deckung gebracht werden können. Hier wären zuerst einmal konkrete Erfahrungen zu sammeln. Ein administrativer Geschäftsprozess lässt sich wohl nicht ohne weiteres mit dem Produktions- und Weiterverarbeitungsprozess einer elektronischen Publikation vergleichen.

Technische Standards

Im Bereich der technischen Standards sind im Rahmen dieser Studien von Bedeutung:

- Charactersets
- Dokument-, Audio- und Bildformaten
- Datenträgern
- Information Retrieval und Abfragesprachen

Hinsichtlich Formate gelten die folgenden Anforderungen:

- plattformunabhängige Lesbarkeit (keine Abhängigkeit von bestimmten Hard- oder Software-Typen)
- offen gelegte Formate; vollständig zugängliche Dokumentationen

Auf Text basierende Informationen sollten wenn möglich in ASCII oder allenfalls ISO-8859-1 abgelegt werden. Der ASCII-Zeichensatz ist standardisiert und kann auf allen Systemen gelesen werden. Probleme ergeben sich, wenn der Text formatiert oder Sonderzeichen (Umlaute) enthält. In diesem Fall müssen die Formatierungen und die Sonderzeichen im ASCII-Text kodiert werden.

chive“, „record“, „document“ usw. zu vermeiden, werden möglichst neutrale Termini wie „data object“, „information object“, „information package“, „packaging information“, „submission information package“ (SIP = entgegengenommenes Informationspaket), „archival information package“, (AIP = archiviertes Informationspaket) oder „disseminated information package“ (DIP = zugänglich gemachtes Informationspaket) verwendet. Der Prozess, der aus einer SIP eine AIP macht, wird als „Ingest“, derjenige, der die AIP in eine DIP überführt, als „Access“ bezeichnet.

Das logische Modell für Archiv-Informationen gliedert, mittels United Modelling Language (UML), Informationsobjekte in Datenobjekte plus Darstellungs-Informationen („representation informations“), während die Datenobjekte aus physischen (nicht-digitalen) Objekten oder Daten-Objekten und diese letztlich aus Bits bestehen. Eine Darstellungs-Information enthält Struktur- und Semantik-Informationen. Weiter kann sich ein Informationsobjekt auf Inhalts-, Aufbewahrungs-, Verpackungs-, Beschreibungs-Informationen usw. beziehen und beliebigen weiteren Informationsobjekten zugeordnet sein. Nur die Inhalts-Information betrifft die eigentlichen zu archivierenden Inhalte; alle übrigen zugeordneten Informationsobjekte sind als generisch aufgebaute Metadaten-Konstrukte zu betrachten. Die Verpackungs-Information („packaging information“) umhüllt in diesem Sinn analog Zwiebelschalen mehrfach eine Inhalts-Information.

Textformate

Auf die dauerhafte Archivierung von Texten, deren Informationen in binärer Form kodiert sind, wie dies für die meisten Textverarbeitungssysteme zutrifft, soll verzichtet werden. Geeignete Formate sind:

- **Text-Datei (Plain text):**
Der Text wird direkt als ASCII-Text abgelegt. Formatierungsinformationen (z.B. Fettschrift, Kursivschrift, fonts) gehen in diesem Fall verloren, Umlaute sind nicht darstellbar. Die Strukturierung des Texts in Abschnitte bleibt aber in der Regel erhalten.
- **Formate auf der Basis von Textkodierungssprachen (SGML, HTML, XML usw.):** Auf der Basis der Textkodierungssprache SGML zur Kodierung sind mehrere Standards entstanden, welche ausschliesslich den ASCII-Zeichensatz verwenden:
 - **HTML:** das Format, das heute hauptsächlich im Bereich von WEB-Applikationen verwendet wird
 - **XML:** ein flexibler Standard zur plattformunabhängigen Speicherung von strukturierten Daten; eine vollständige XML-Datei besteht aus einer Beschreibung der Struktur (DTD, Document Type Definition; künftig: Schema) mit der Definition der verwendeten Textmarken (Tags) und dem eigentlichen Datenteil
 - **TeX/LaTeX:** ein Format aus dem wissenschaftlichen Umfeld, welches die Strukturierung von und Formatierung fachspezifischer Texte unterstützt
 - **RTF:**
Beim RichTextFormat handelt es sich um ein proprietäres Format mit offen gelegtem Standard. RTF wird von vielen Textverarbeitungssystemen (MSWord, Write) interpretiert.
 - **Postscript, PDF:**
Postscript und PDF beinhalten ebenfalls ein proprietäres Format mit offen gelegtem Standard. Es existieren unabhängige Implementationen von Postscript-Interpretern.

Bildformate

Von der grossen Anzahl unterschiedlicher Dateiformate für digitale Bilder eignen sich für die Langzeitarchivierung nur wenige:

- **TIFF:**
ein plattformunabhängiges Format mit offen gelegtem Standard. Alle gängigen Bildbearbeitungsprogramme unterstützen TIFF
- **JPEG:**
Ein weit verbreitetes Format mit offen gelegtem Standard; für die Langzeitarchivierung wegen seiner Kompression mit Verlust nur bedingt geeignet. Das Format eignet sich eher als Arbeitskopie für digitale Bilder
- **Übersicht**

Nachstehend aufgeführt sind Standards und deren Entwickler (resp. verantwortliche Institutionen), die durch verschiedene nationale und internationale Standardkomitees erstellt worden („de jure“-Standards) oder aber als Industriestandard, aufgrund der Zusammenarbeit verschiedener Firmen und Konsortien, zu Stande gekommen sind („de facto“-Standard, z.B. TIFF und Postscript).

Character sets

Standard		Verantwortlich	Bemerkung
Ascii	American Standard Code for Information Interchange	American National Standards Institute (ANSI)	7bit Characterset
EBCDIC	Extended Binary Coded Decimal Interchange Code	Proprietary specification developed by IBM	8bit Characterset für IBM Computer
ISO 8859	ISO/IEC 8859: 8-bit single-byte coded graphic character sets	ISO	umfasst Spezialzeichen aus europäischem Sprachraum (ISO-8859-1)
OCR	Optical Character Recognition	ISO	
ISO 10646	ISO 10646: Universal Multiple-Octet Coded Character Set (UCS)	ISO	Multilingual character set covering all major trading languages Unicode 3.0 to be based on 2nd Edition of ISO 10646

Dokument-, Audio- und Bildformate

Standard		Verantwortlich	Bemerkung
HTML	HyperText Markup Language	World Wide Web Consortium (W3C)	
PDF	Portable Document Format	Proprietary standard developed by Adobe Systems Inc.	Derivat von Postscript
RTF	Rich Text Format	Proprietary standard developed by Microsoft Corporation	
SGML	Standard Generalized Markup Language	ISO/IEC	
XML	The eXtensible	World Wide Web	

Standard		Verantwortlich	Bemerkung
	Markup Language	Consortium (W3C)	
XSL	The eXtensible Stylesheet Language	World Wide Web Consortium (W3C)	Transformation and presentation of XML documents
Postscript		Proprietary standard developed by Adobe Systems Inc	Basis von PDF Industriestandard
UPF	Universal Preservation Format	WGBH (Boston)	Kein eigentliches Format sondern Richtlinien mit Schwerpunkt auf digitalen Videoformaten
MPEG-2 Audio / MP3	Coding of Moving Pictures and Associated Audio for Digital Storage Media	ISO (Moving Pictures Expert Group)	Teil 3 von MPEG-2
WAVE	Waveform Audio File Format	Proprietary standard developed by Microsoft and IBM	
TIFF	Tag Image File Format	Proprietary format developed by Aldus Corporation (now owned by Adobe) and Microsoft	Industriestandard
GIF	Graphic Interchange Format	Proprietary specification developed CompuServe Inc.	Standard offen gelegt, enthält aber proprietären Algorithmus zur verlustfreien Kompression, welcher nur nach Lizenzierung verwendet werden darf
G and SPIFF	Joint Photographic Experts Group standard and Still Picture Interchange File Format	ISO	Komprimierte Bilddaten

Datenträger

Standard		Verantwortlich	Bemerkung
ISO-9660		ISO	Volume and File Structure of CD-ROM for Information Interchange kann auf allen gängigen Systemen gelesen werden
UDF	Universal Disc Format		
DVD	Digital Video Disc / Digital Versatile Disc	The DVD Video Group	
PHOTO-CD	Photo Compact Disk	Proprietary standard patented by Eastman Kodak Company	

Information Retrieval und Abfragesprachen

Standard		Verantwortlich	Bemerkung
Information Retrieval		ANSI	ANSI/NISO Z39.58-1992 Common Command Language for On-Line Interactive Information Retrieval
SQL	Structured Query Language	ISO	Datenbank-Definitions (DDL Data Definition Language) und -Abfragesprache (DML Data Manipulation Language)
Z39.50			