

*„Electronic Archives
are the Memory of the
Information Society.“*

Erkki Liikanen, EU-Kommissar für die Informationsgesellschaft
(bis 2004)

Erkki Liikanen, EU-Kommissar für die Informationsgesellschaft (bis 2004)



digitales

STAATSARCHIV

Digitale Langzeitarchivierung von elektronischen Akten der Bundesverwaltung

15.11.2007

Uni Wien

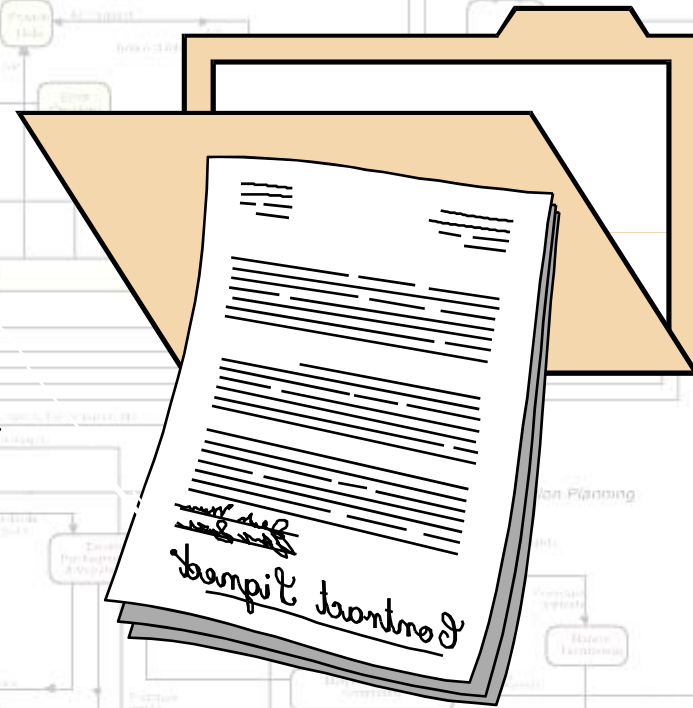
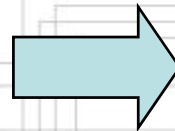
Thomas Callsen-Rauer, PL



Historisches

manuell

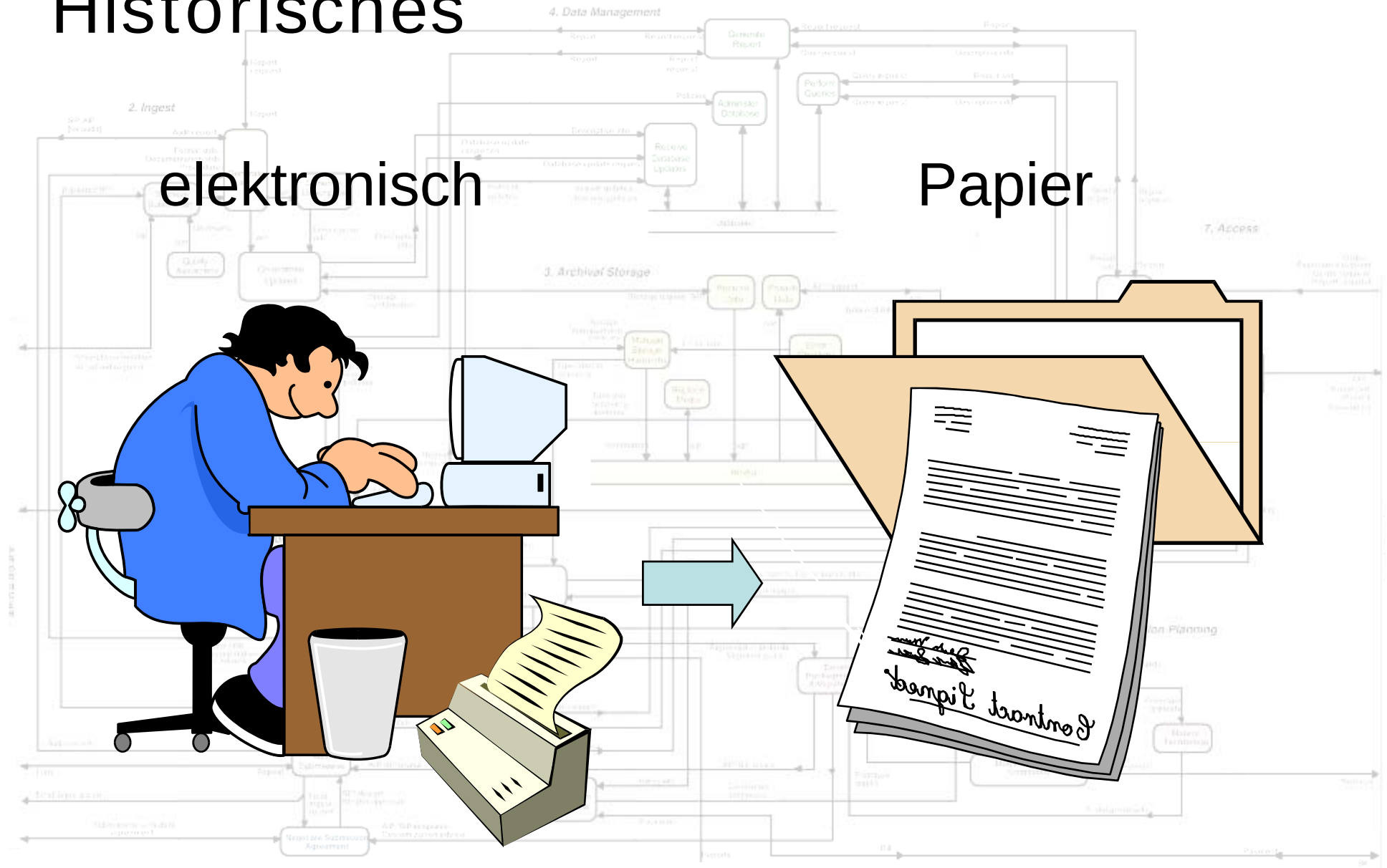
Papier



Historisches

elektronisch

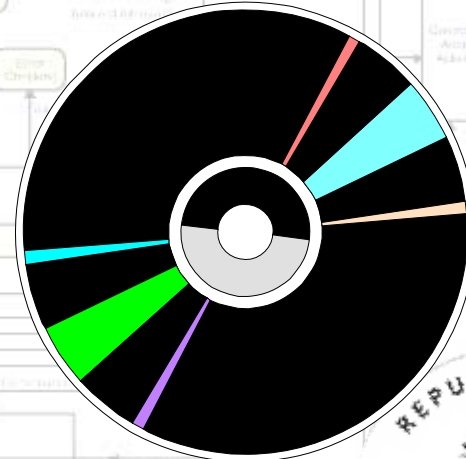
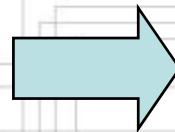
Papier



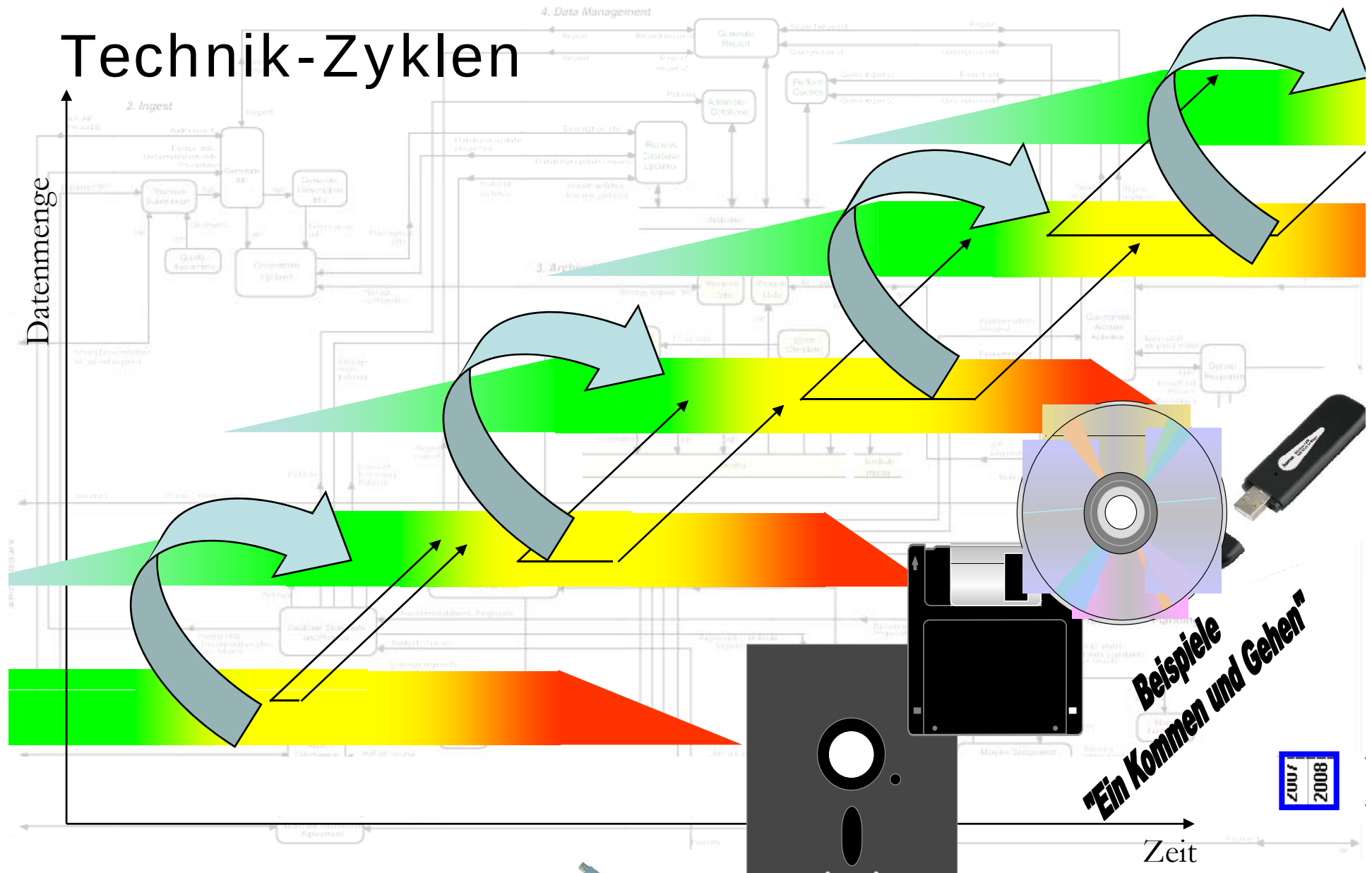
Historisches

elektronisch

elektronisch

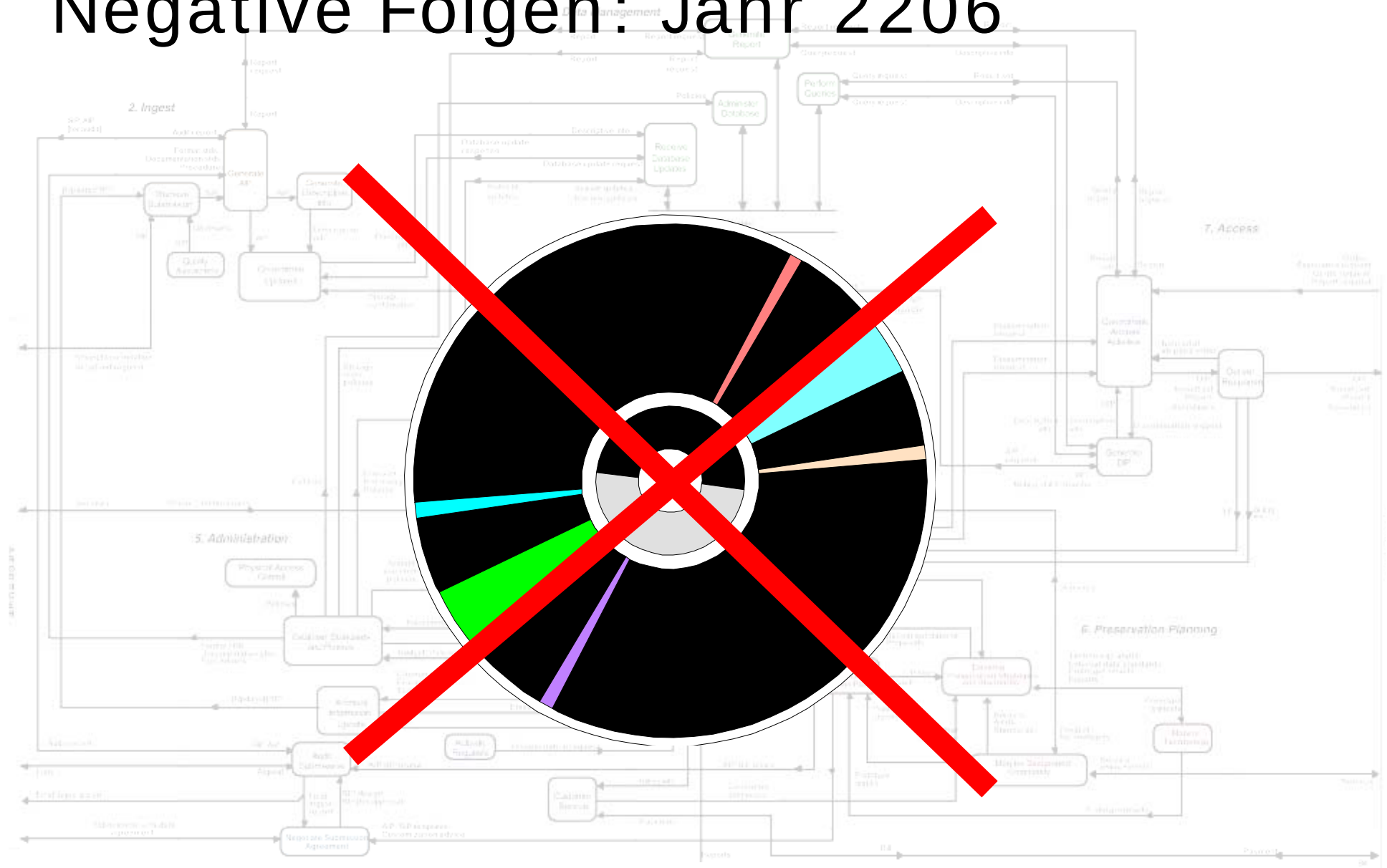


Technik-Zyklen



**Beispiele
"Ein Kommen und Gehen"**

Negative Folgen: Jahr 2206



Ziel



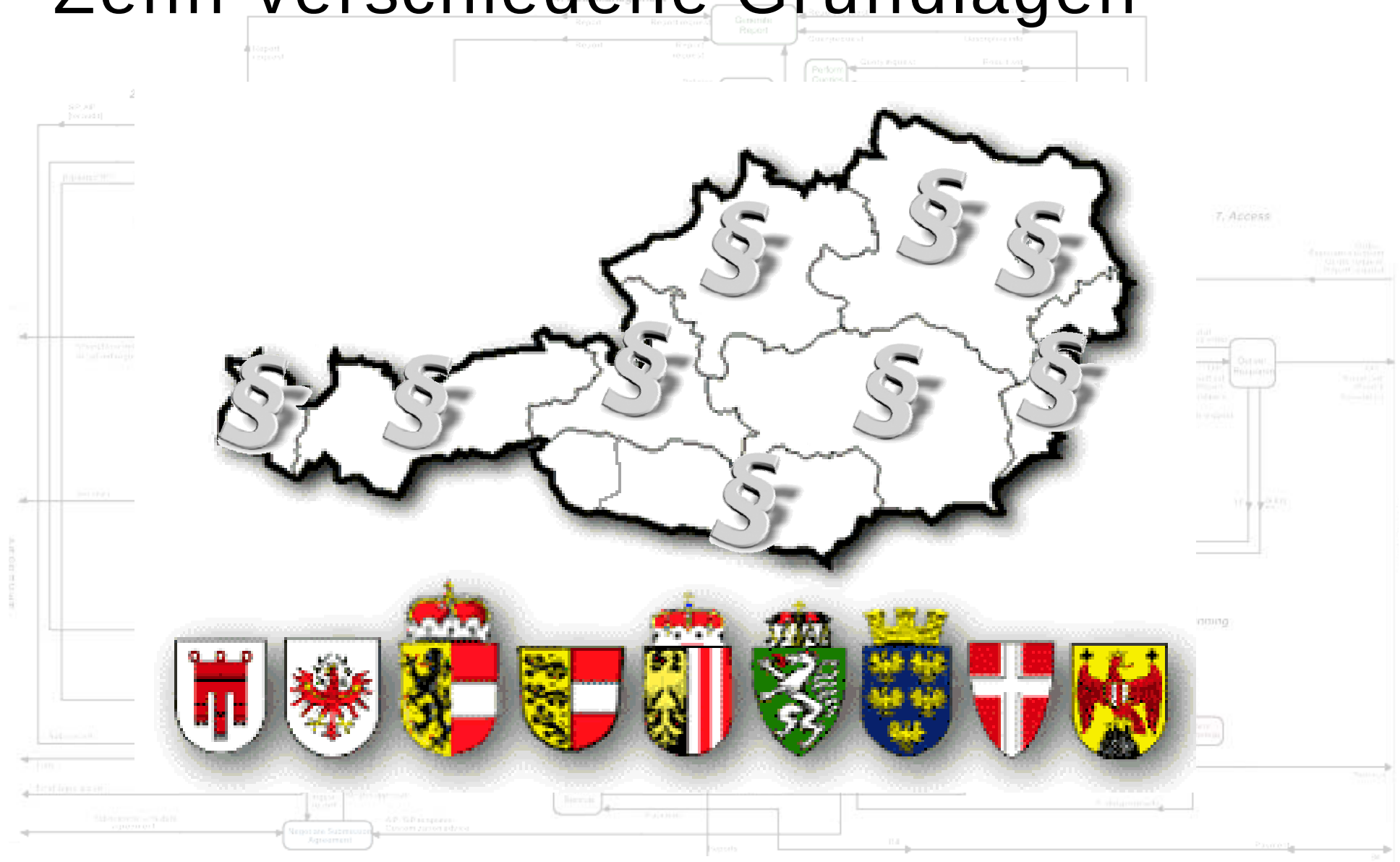
***lesbare* digitale
Informationen
auch in
200+ Jahren**



BLSG Arbeitsgruppe



Zehn verschiedene Grundlagen



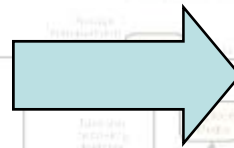
4. Data Management



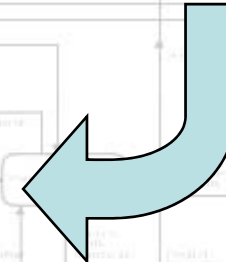
Konkretisierung für den Bund

Bundesministerien

Österreichisches
Staatsarchiv



St. Johann



Technischer Hintergrund

Bundesministerien

Österreichisches
Staatsarchiv



St. Johann



Mehrere Quellen

Bund



Bundesländer



Europäische Union



St. Johann



Österreichisches
Staatsarchiv



Speicherformate

PDF/A

ISO-Standard ISO/CD 19005-1

„Document management
– Electronic document file format
for long-term preservation
Part 1: Use of PDF (PDF/A)“

EDIAKT II

BLSG Empfehlung; siehe

<http://reference.e-government.gv.at>



Mailingliste:

<https://labs.cio.gv.at/mailman/listinfo/ediakt2>

Verwaltungsinformationen

Inhalte

EDIAKT – XML Paket

- Metadaten
- Prozessdaten
- Inhaltsdaten
- Fachdaten

Sender → Empfänger



Sender

EDI AKT – XML Paket

Metadaten

Prozessdaten

Inhaltsdaten

Fachdaten

Empfänger



EDI AKT Viewer

<http://www.bka.gv.at/ediakt>

The screenshot shows the EdiaktViewer V1.0 application window. The title bar reads "EdiaktViewer V1.0". The menu bar includes "Datei", "Bearbeiten", and "Hilfe". The toolbar contains icons for file operations. The left pane displays a tree view of a document structure:

- Ressort_A-123.456/0001-XY/2005
 - Ressort_A-123.456/0001-XY/2005 (GF)
 - Interne Beilagen (Gstk)
 - Belegaufstellung (D)
 - Kanzlei u. Votum (D)
 - Erledigung an den Turnverein (Gstk)
 - Standarderledigung Ressort_A-Logo/08.03..
 - Standarderledigung Ressort_A-Logo/08.03..
 - Ressort_A-047011/2005 (Gstk)
 - Turnverein; Förderabrechnung 2004 (D)

The right pane shows the document "Bsp-Gericht_2006-10-23.xml". The document title is "Übergeordneter Betreff Förderungen der Antragsteller". The organization unit is "Ressort_A - XY (Förderkontrolle)".

Status

- Geschlossen ☒ [x]
- Abgebrochen ☐ []

Rollen

Icon	Name	Role
	Förderkontrolle	(Ersteller)
	KANZL Anton	(Letzte Änderung von)
	MUSTERMEIER Max	(Erstellt von)
	ZANGEL Maria	(Genehmigender)

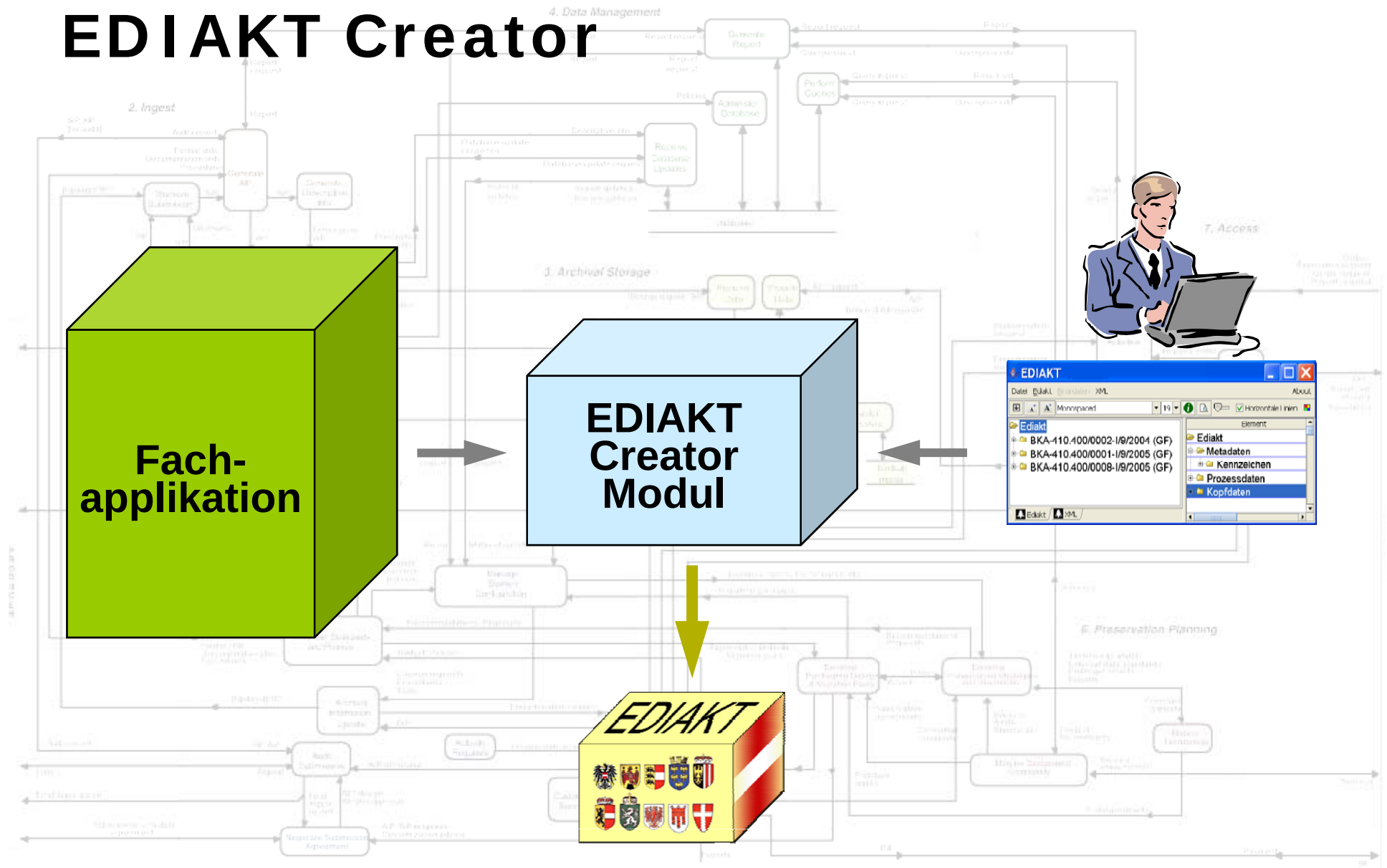
Zeiten und Fristen

- Letzte Änderung 2005-03-16T17:40:00

Versand

Navigation tabs at the bottom: Deckblatt (Geschäftsfall) | Metadaten | Prozessdaten | Referenzen

EDIAKT Creator



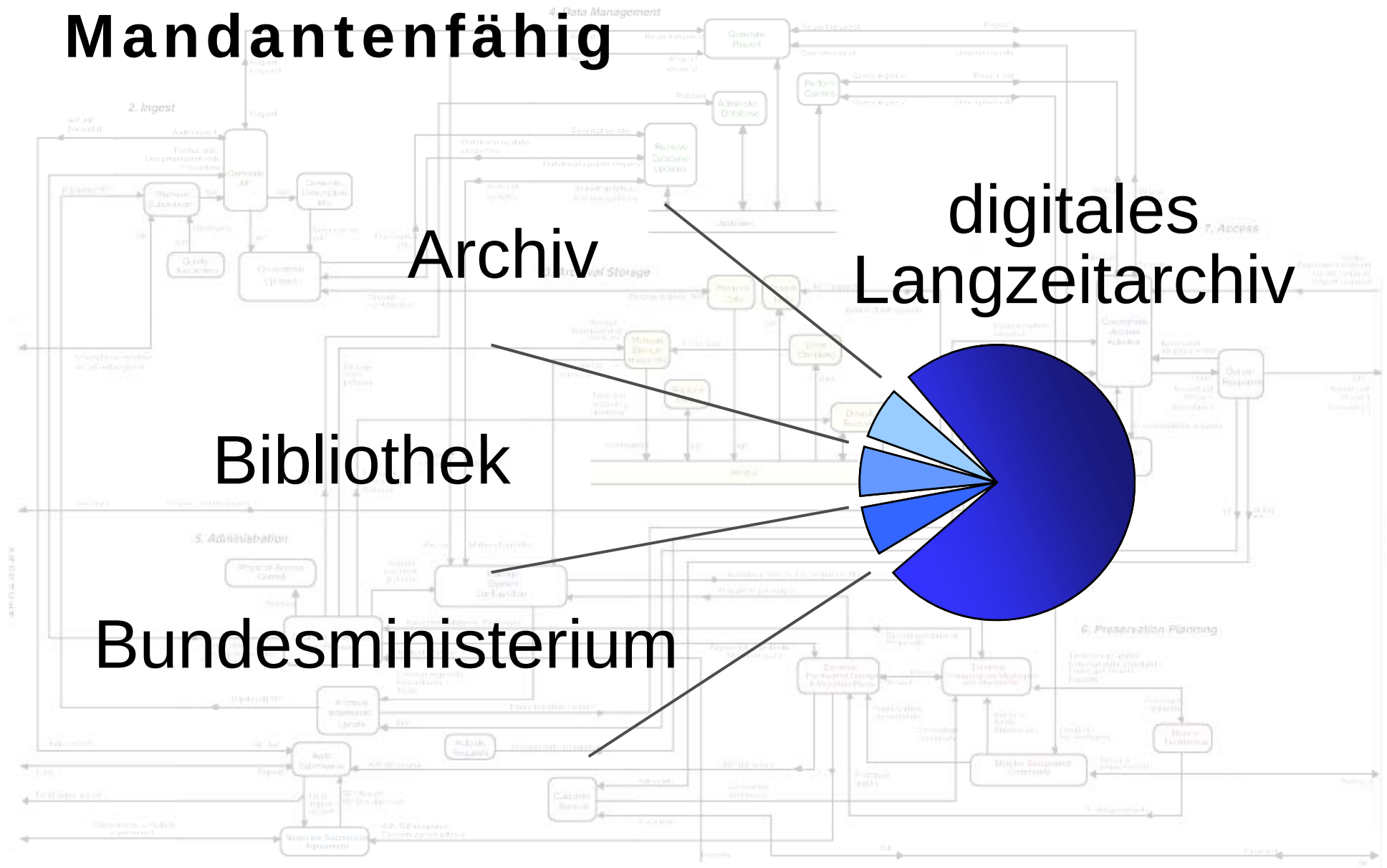
Mandantenfähig

Archiv

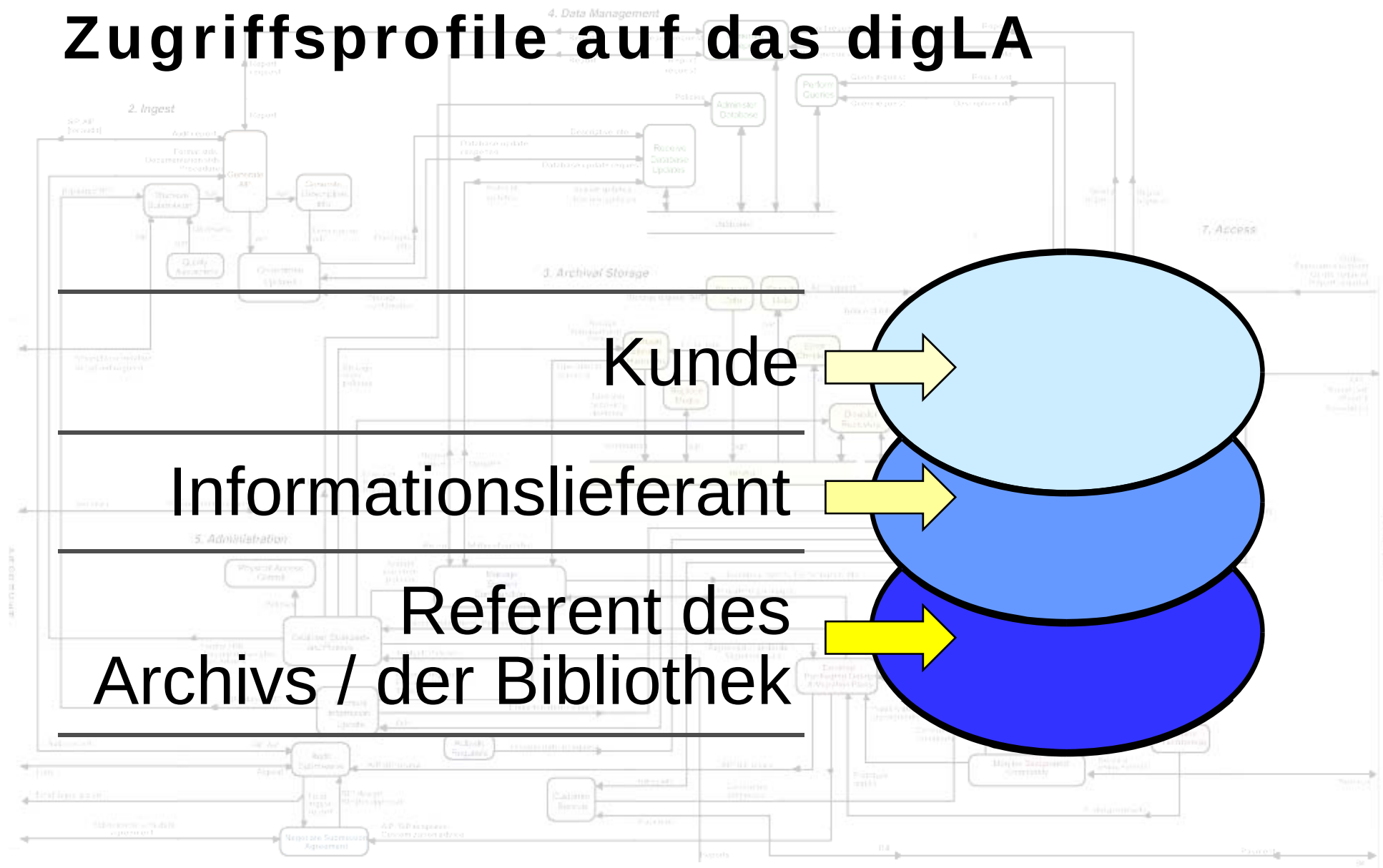
digitales Langzeitarchiv

Bibliothek

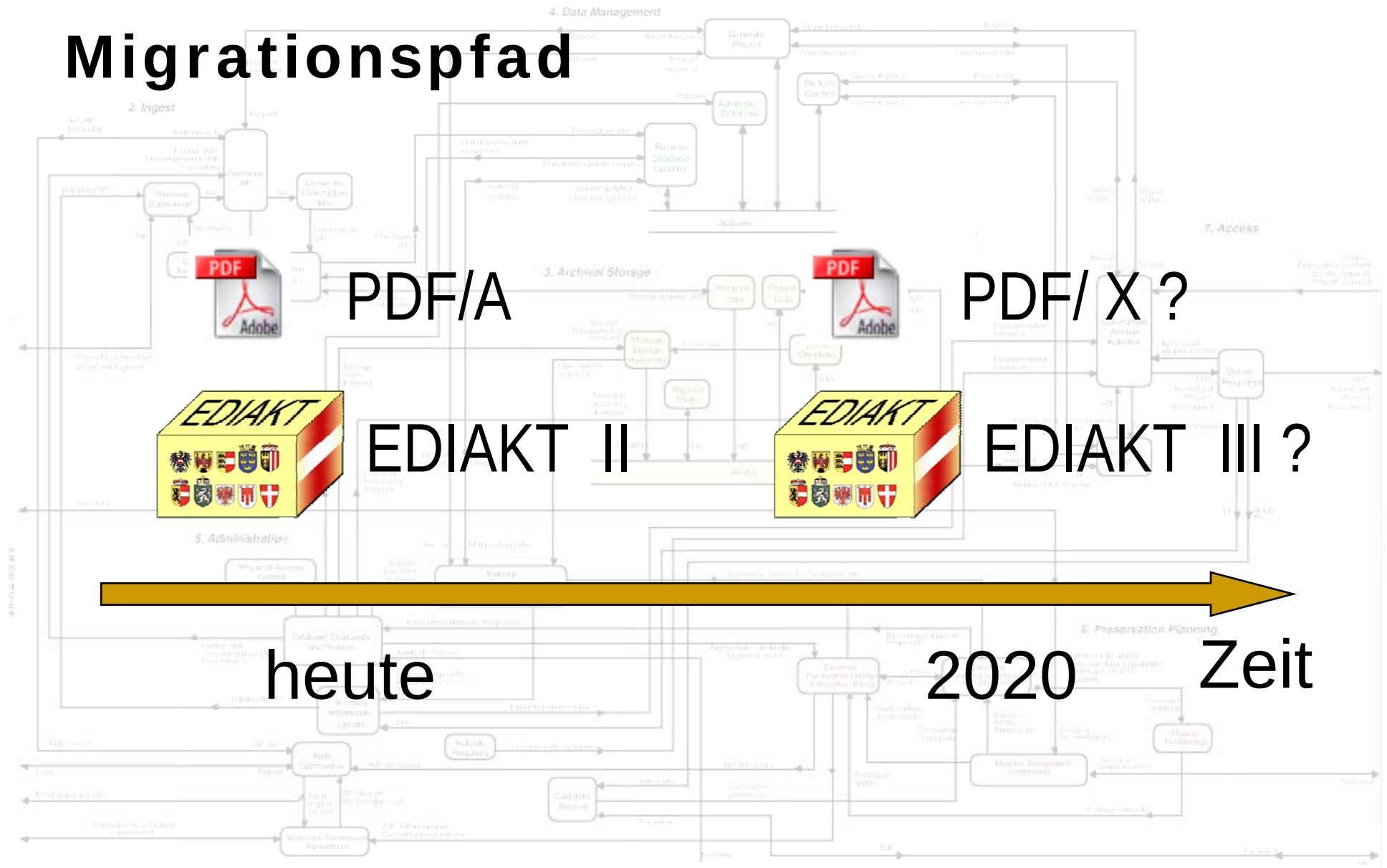
Bundesministerium



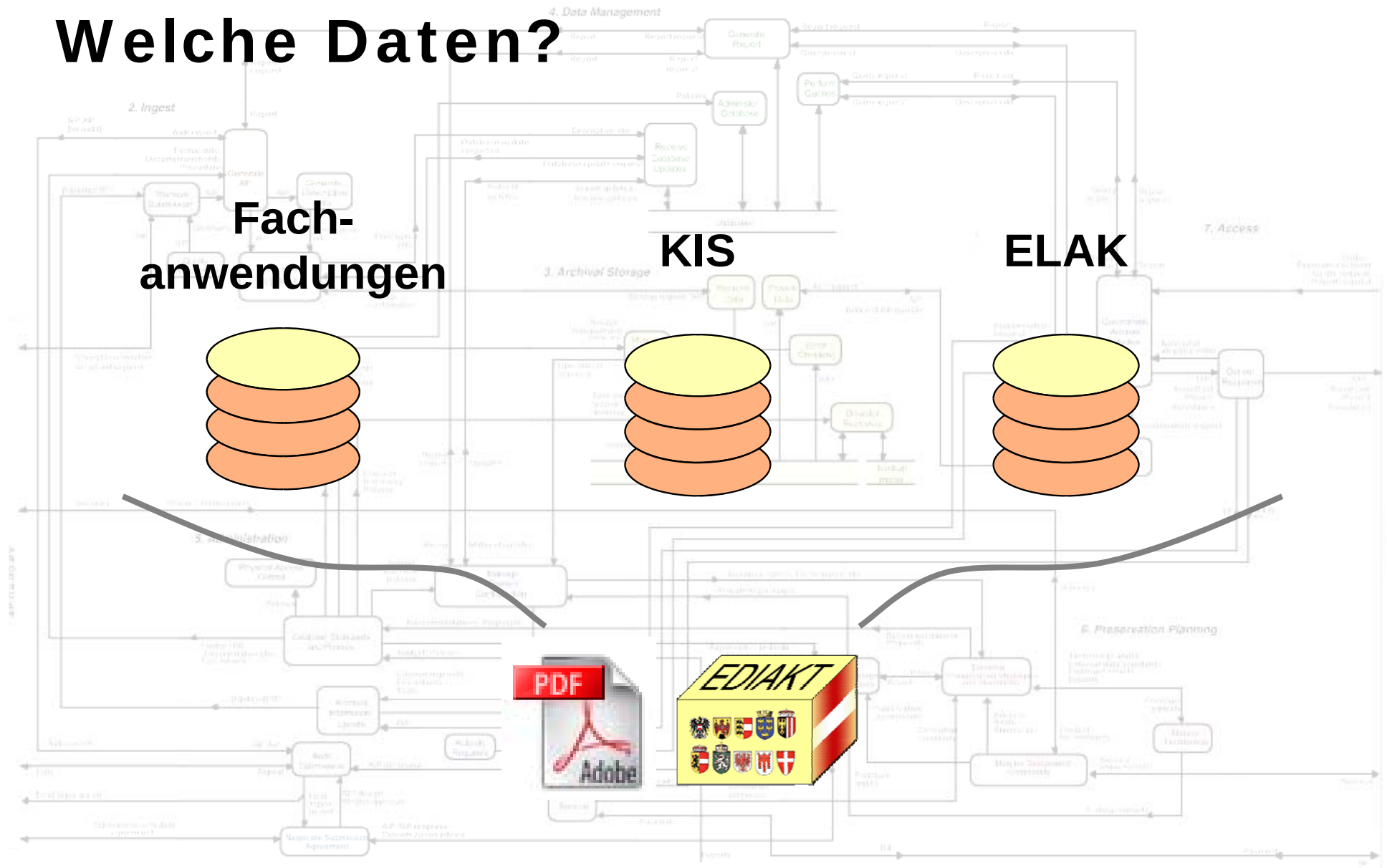
Zugriffsprofile auf das digLA



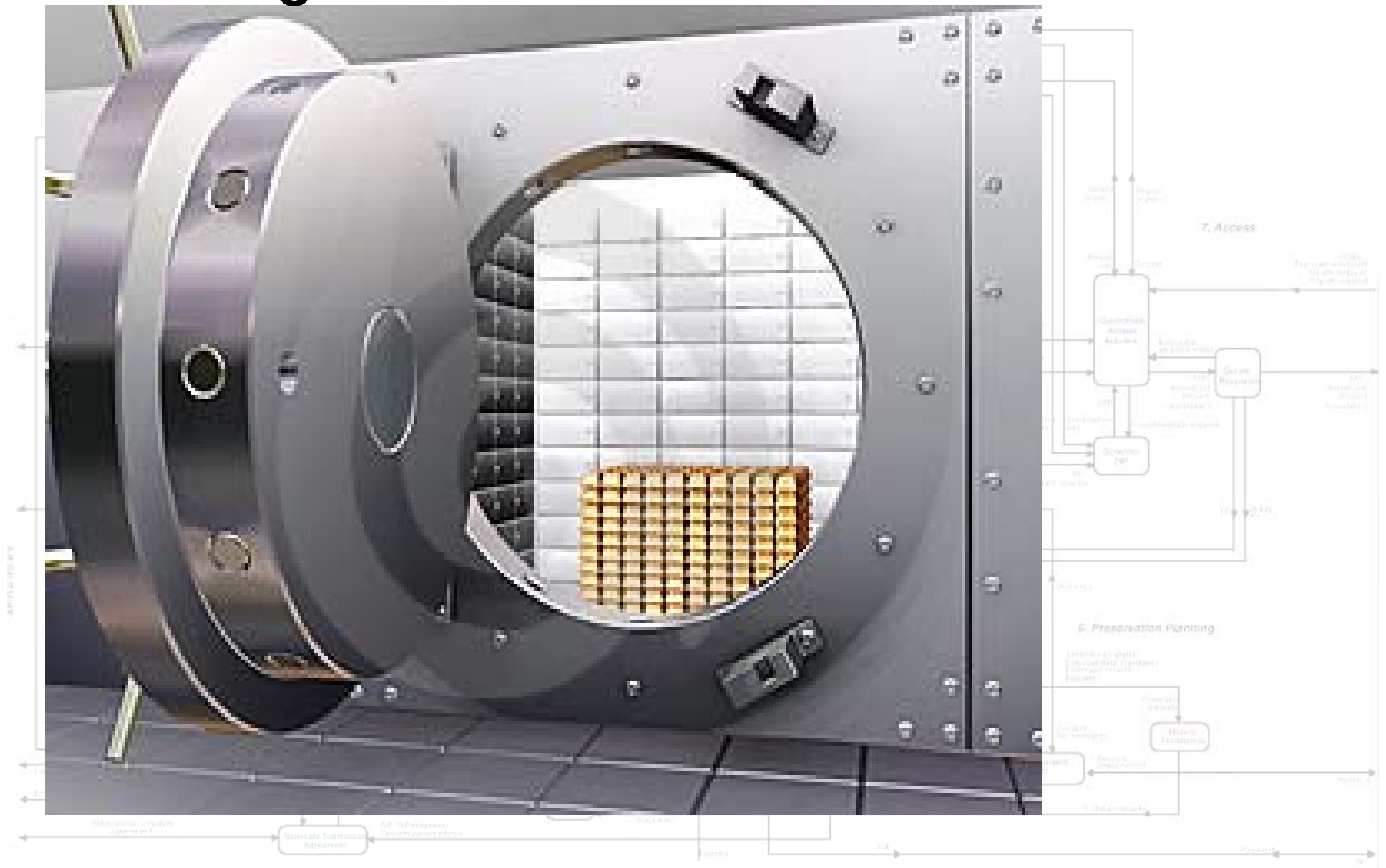
Migrationspfad



Welche Daten?



Wo liegen die Daten?



Zusätzliche Vorteile des Standortes

- befindet sich in einem militärischen Hochsicherheitsbereich
- damit verbunden ein extrem hoher Objekt- und Personenschutz
- Höchste Redundanzen bei Infrastruktur (Klima, Energieversorgung, USV etc.)
- Zero – Risk - Umgebung

The diagram illustrates the OAIS Reference Model, a framework for digital preservation. It is organized into seven functional areas, each with specific components and interactions:

- 1. Ingest:** This area is responsible for receiving digital objects. It includes the **Ingest Manager**, which interacts with the **Archival Storage Manager** and the **Preservation Planning Manager**. It also involves **Physical Access Control** and **Access Information Update**.
- 2. Archival Storage:** This area is responsible for storing digital objects. It includes the **Archival Storage Manager**, which interacts with the **Ingest Manager** and the **Preservation Planning Manager**. It also involves **Physical Access Control** and **Access Information Update**.
- 3. Administration:** This area is responsible for managing the digital objects. It includes the **Administration Manager**, which interacts with the **Ingest Manager** and the **Preservation Planning Manager**. It also involves **Physical Access Control** and **Access Information Update**.
- 4. Data Management:** This area is responsible for managing the data. It includes the **Data Manager**, which interacts with the **Ingest Manager** and the **Preservation Planning Manager**. It also involves **Physical Access Control** and **Access Information Update**.
- 5. Access:** This area is responsible for providing access to the digital objects. It includes the **Access Manager**, which interacts with the **Ingest Manager** and the **Preservation Planning Manager**. It also involves **Physical Access Control** and **Access Information Update**.
- 6. Preservation Planning:** This area is responsible for planning the preservation of digital objects. It includes the **Preservation Planning Manager**, which interacts with the **Ingest Manager** and the **Archival Storage Manager**. It also involves **Physical Access Control** and **Access Information Update**.

The diagram shows a complex network of interactions between these functional areas, ensuring the long-term preservation and access of digital objects.

Ziele

■ Vortrag OAIS Modell

- Verständnis der Aufgabenbereiche
- Verständnis des OAIS Modell als Referenzmodell

■ Ermitteln

- welche Teile sind vorhanden, in welcher Form
- welche Teile sind zu schaffen
- wie spielen die einzelnen Komponenten zusammen

Überblick

- Prinzipien des OAIS Modells
- Technischer Überblick
- Funktionaler Überblick
- Informationsmodell
- Zusammenfassung:
wichtige Komponenten -> DigLAimBund

OAIS & Rolle der NASA

- **National Space Science Data Center**

- NASA's erstes digitales Archiv

- hat viele Technologiewechsel seit 1966 durchlebt

- **Consultative Committee for Space Data Systems**

- Internationale Gruppe von Space Agencies

- Entwickelte eine Reihe von Disziplin-unabhängigen Standards

- Wurde zur Arbeitsgruppe für ISO TC 20/ SC 13 um 1990

- TC20: Aircraft and Space Vehicles

- SC13: Space Data and Information Transfer Systems

Was ist ein Referenz Modell

■ Ein Framework

- um die signifikanten Beziehungen zwischen signifikanten Entitäten in einer Umgebung zu verstehen, und
- für die Entwicklung von konsistenten Standards oder Spezifikationen zur Unterstützung dieser Umgebungen.

■ Ein Referenzmodell

- basiert auf einer kleinen Anzahl vereinigender Konzepte
- stellt eine Abstraktion der Kernkonzepte, deren Beziehungen, und deren Schnittstellen sowohl zueinander als auch zur externen Umgebung dar
- kann als Basis zum Training und zur Erklärung eines Standards dienen.

OAIS und DigLAimBund

- OAIS ist ein Referenzmodell
- keine Umsetzungsanleitung, kein Datenmodell, keine Spezifikation
- beschreibt jene Elemente, Konzepte, die für das Projekt relevant sind
- Ziel: festzustellen, welche Teilbereiche des Referenzmodells welchen konkreten Systemen, Funktionen und Verantwortlichkeiten in der angestrebten Lösung entsprechen.

OAIS Informationsquellen

- Reference Model for an Open Archival Information System (OAIS), Blue Book, CCSDS 650.0-B-1, January 2002
- Folien basierend auf Blue Book und:
 - Don Sawyer, Lou Reich: ISO Reference Model for an Open Archival Information System (OAIS) Tutorial Presentation, LOC, 13. Juni 2003
 - größtenteils in Originalsprache belassen, um Kernkonzepte im internationalen Sprachgebrauch zu belassen

Überblick

- Prinzipien des OAIS Modells
- Technischer Überblick
- Funktionaler Überblick
- Informationsmodell
- Zusammenfassung:
wichtige Komponenten -> DigLAimBund

Open Archival Information System (OAIS)

■ Open

- Reference Model standard(s) are developed using a public process and are freely available

■ Information

- Any type of knowledge that can be exchanged
- Independent of the forms (i.e., physical or digital) used to represent the information
- Data are the representation forms of information

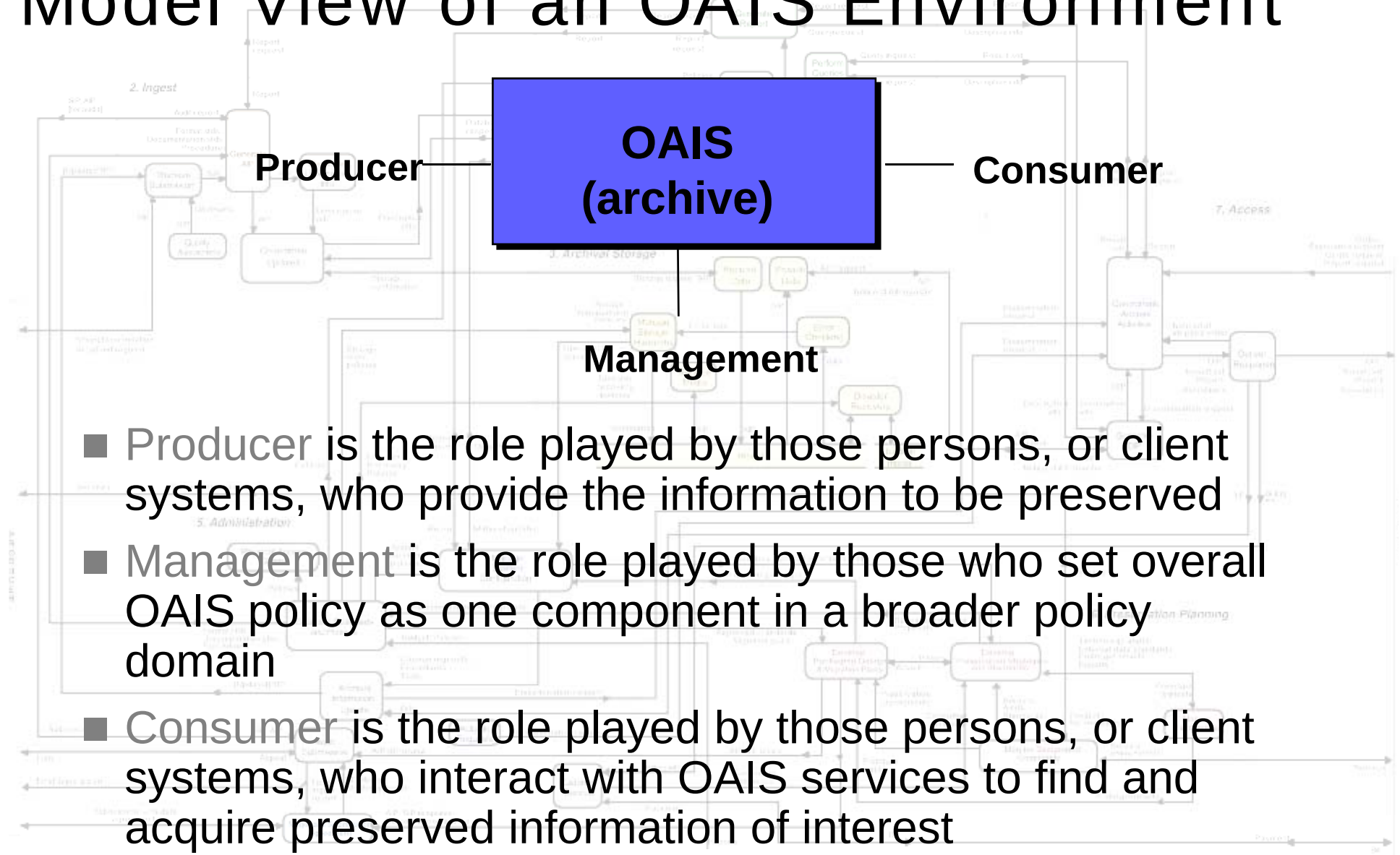
■ Archival Information System

- Hardware, software, and people who are responsible for the acquisition, preservation and dissemination of the information

Purpose, Scope, and Applicability

- Framework for understanding and applying concepts needed for long-term digital information preservation
 - Long-term is long enough to be concerned about changing technologies
 - Starting point for model addressing non-digital information
- Provides set of minimal responsibilities to distinguish an OAIS from other uses of 'archive'
- Framework for comparing architectures and operations of existing and future archives
- Basis for development of additional related standards
- Addresses a full range of archival functions
- Applicable to all long-term archives and those organizations and individuals dealing with information that may need long-term preservation
- Does NOT specify an implementation

Model View of an OAIS Environment



OAIS Information Definition

- Information is always expressed (i.e., represented) by some type of data
- Data interpreted using its Representation Information yields Information
- Information Object preservation requires clear identification and understanding of the Data Object and its associated Representation Information



Information Package Definition

Content
Information

Preservation
Description
Information

- An Information Package is a conceptual container holding two types of information
 - Content Information
 - Preservation Description Information (PDI)

Information Package Variants

■ **SIP:** Submission Information Package

- Negotiated between Producer and OAIS
- Sent to OAIS by a Producer

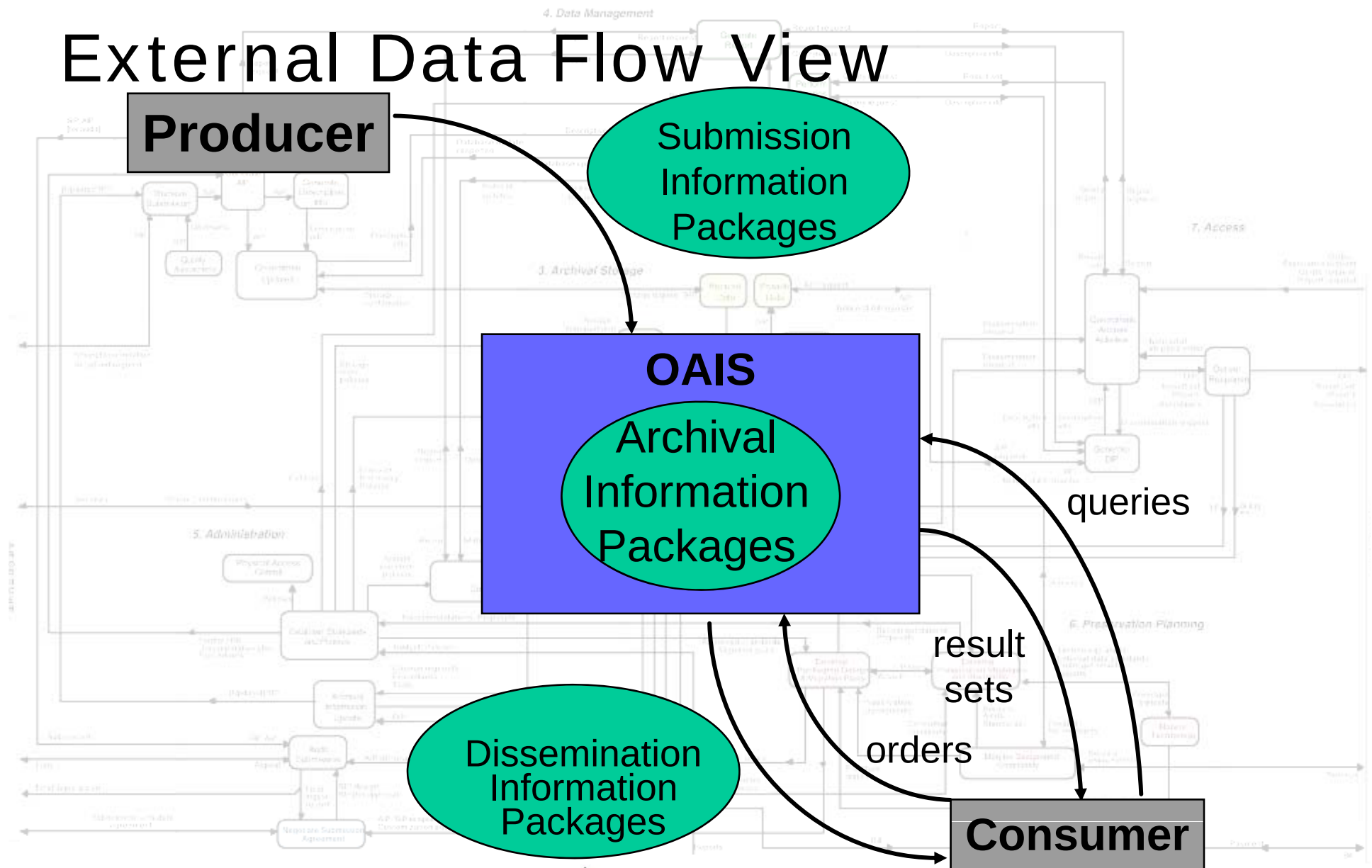
■ **AIP:** Archival Information Package

- Information Package used for preservation
- Includes complete set of Preservation Description Information (PDI) for the Content Information

■ **DIP:** Dissemination Information Package

- Includes part or all of one or more Archival Information Packages
- Sent to a Consumer by the OAIS

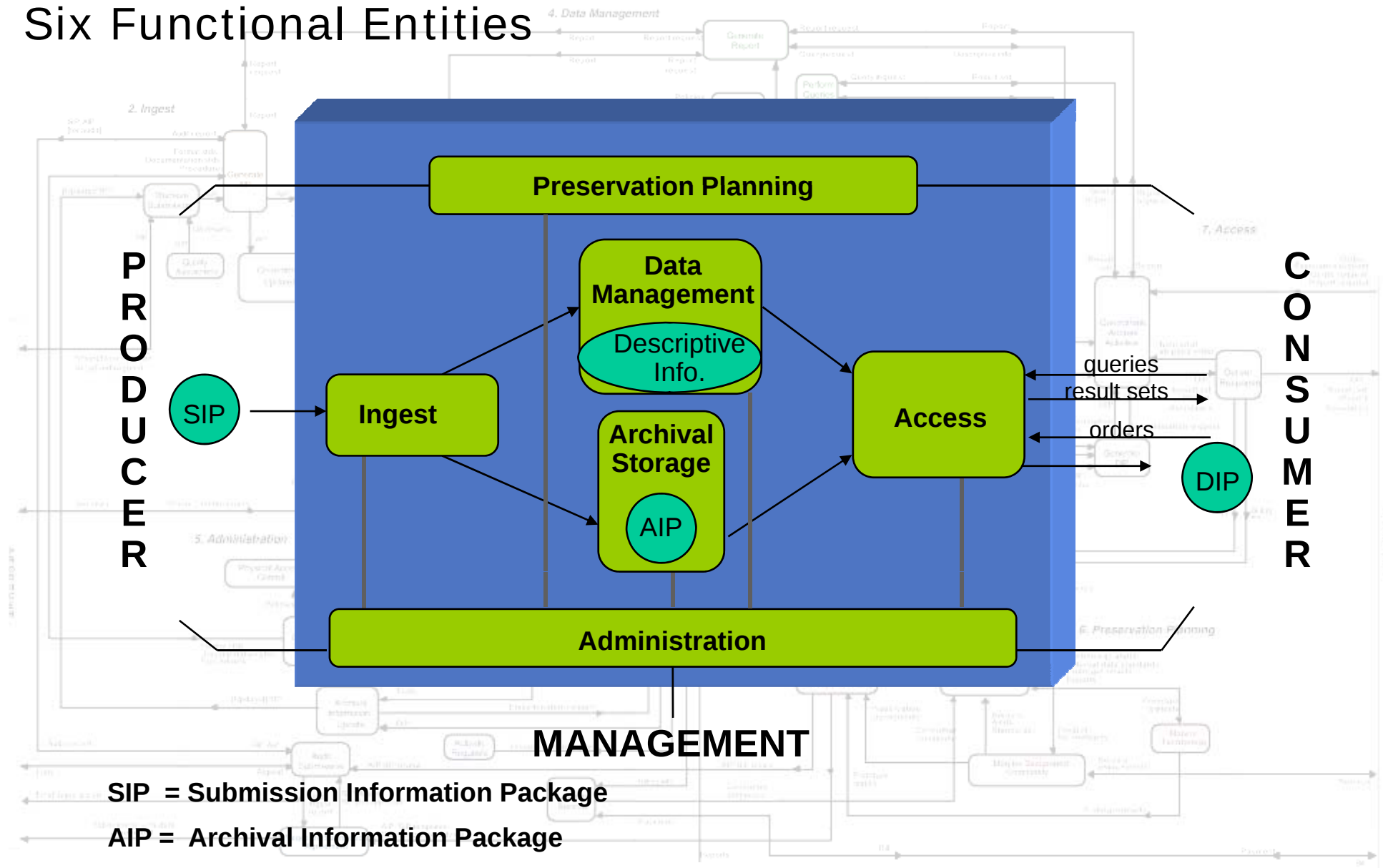
External Data Flow View



Überblick

- Prinzipien des OAIS Modells
- Technischer Überblick
- Funktionaler Überblick
- Informationsmodell
- Zusammenfassung:
wichtige Komponenten -> DigLAimBund

Open Archival Information System: Six Functional Entities



SIP = Submission Information Package

AIP = Archival Information Package

DIP = Dissemination Information Package

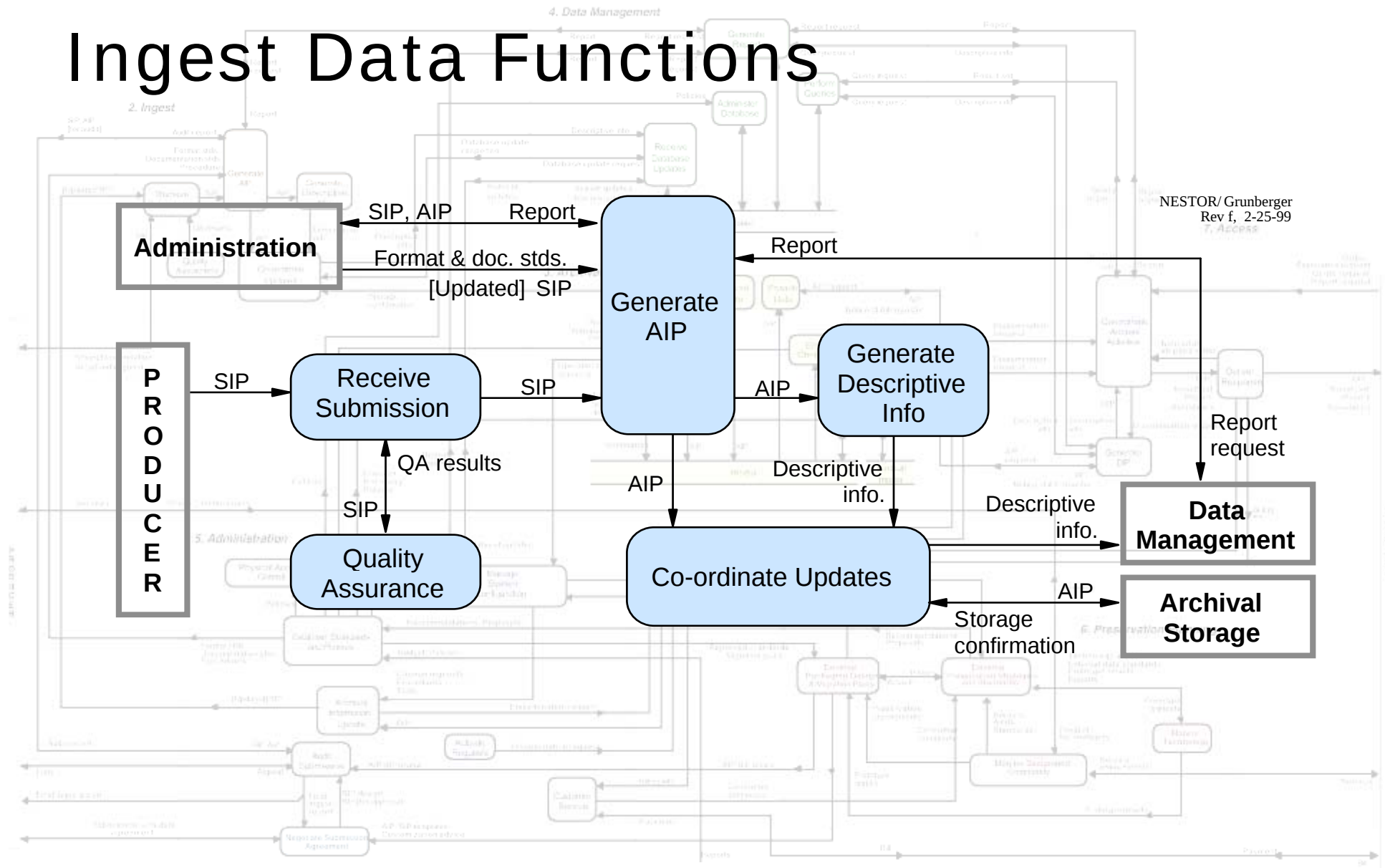
Functional Entities in an OAIS (1/2)

- **Ingest:** This entity provides the services and functions to accept Submission Information Packages (SIPs) from Producers and prepare the contents for storage and management within the archive
- **Archival Storage:** This entity provides the services and functions for the storage, maintenance and retrieval of Archival Information Packages
- **Data Management:** This entity provides the services and functions for populating, maintaining, and accessing both descriptive information which identifies and documents archive holdings and internal archive administrative data.

Functional Entities in an OAIS (2/2)

- **Administration:** This entity manages the overall operation of the archive system
- **Preservation Planning:** This entity monitors the environment of the OAIS and provides recommendations to ensure that the information stored in the OAIS remain accessible to the Designated User Community over the long term even if the original computing environment becomes obsolete.
- **Access:** This entity supports consumers in determining the existence, description, location and availability of information stored in the OAIS and allowing consumers to request and receive information products

Ingest Data Functions



NESTOR/Grünberger
Rev f, 2-25-99
f, ACCESS

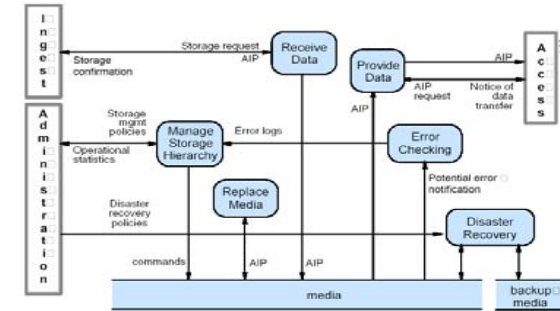
Ingest Data Functions

- Receive Submissions:
 - Zwischenspeicher (Staging Area) für Submissions
 - Bestätigung über Aufnahme in Zwischenspeicher
- Quality Assurance
 - Validierung der Submission (CRC, logs, identity checks, media)
- Generate AIP
 - Transformation von SIPs in AIPs entsprechend den Standards (Transformation, Migration, Umkodierung)
 - Weiterleitung der AIPs an Audit (Administration)
- Generate Descriptive Information
 - Sammlung bzw. Extraktion von deskriptiver Information zum AIP für Data Management und Access Aids
- Coordinate Updates
 - Transfer der AIPs zu Archival Storage
 - Bestätigung -> Deskriptive Information -> Data Management

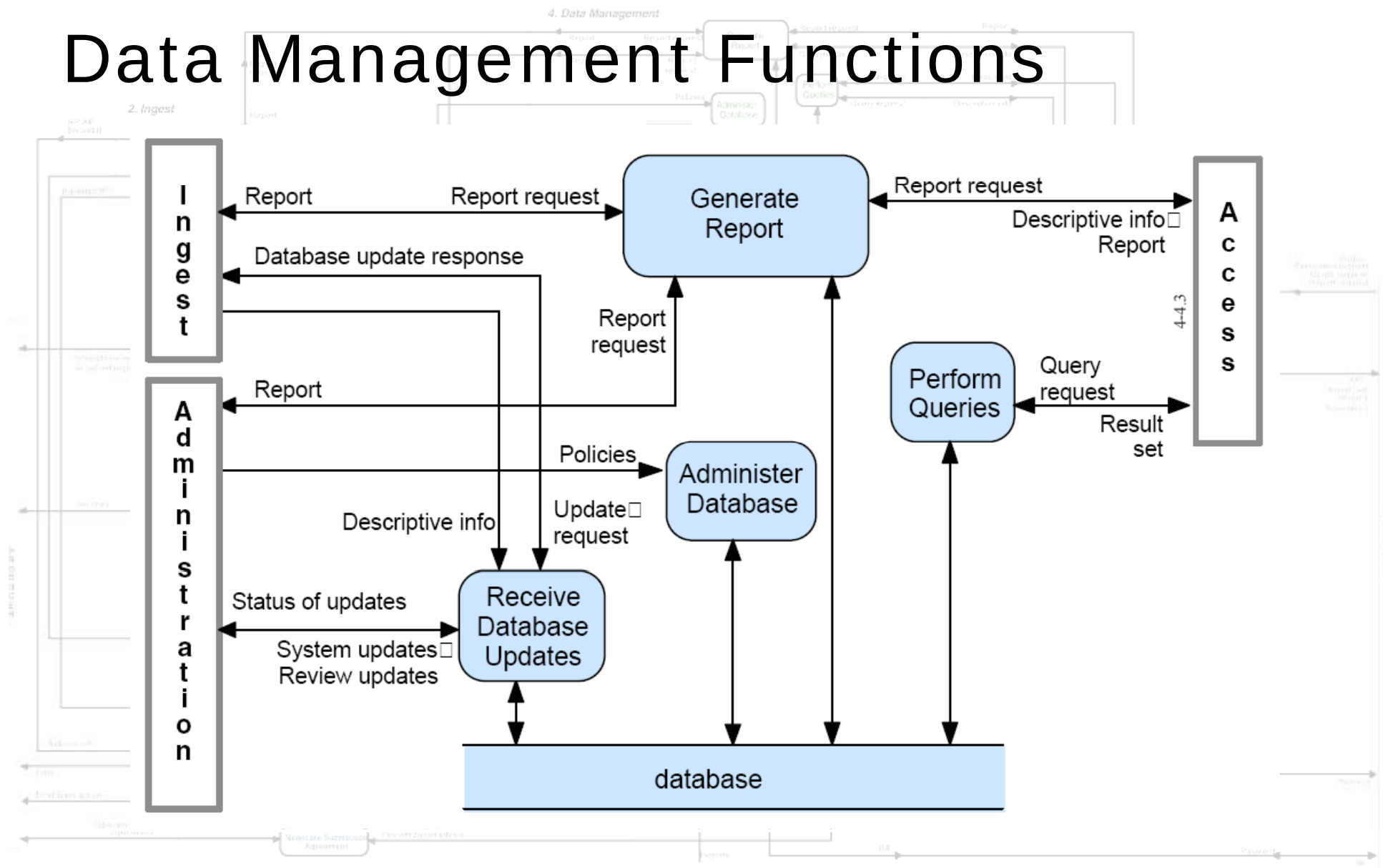
Archival Storage Functions

Archival Storage Functions

- **Receive Data:**
 - übernimmt Storage Request für AIP
 - entscheidet über Speicherort, Medien
 - retourniert Bestätigungsmeldung
- **Manage Storage Hierarchy**
 - Verwaltung entsprechend Policy
 - Überwachung von Fehlermeldungen, operationale Statistiken
- **Replace Media**
 - Reproduktion von AIPs über die Zeit (keine Änderung von Content oder Preservation Description Information, nur Packaging Information - andere Änderungen: Administration)
- **Error Checking**
 - PDI Fixity Information (CRCs, error-correcting codes, ...)
- **Disaster Recovery**
 - Duplizierung des Inhalts of Speichermedien
 - Transport an physisch getrennten Ort
- **Provide Data**
 - Kopien von AIPs für Access

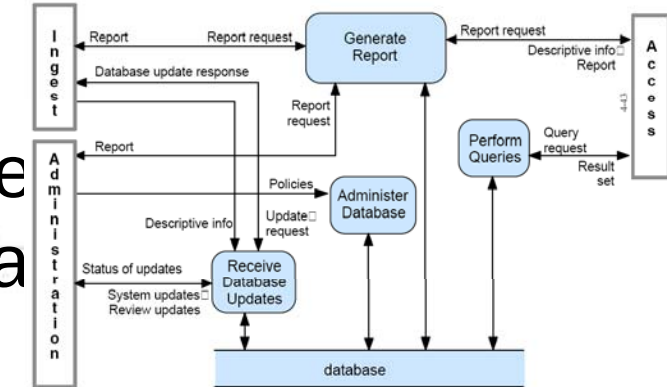


Data Management Functions



Data Management Functions

- Administer Database
 - Integrität der DB für Descriptive Information und Systeminformation
- Perform Queries
 - Bearbeitung von Anfragen durch Access
- Generate Reports
 - Berichte für Ingest, Access, Administration
- Receive DB Updates
 - Fügt hinzu/Löscht/Modifiziert Information in Management DB
 - Ingest: neue AIPs, Administration: updates



[illegible]

Administration Functions

- Negotiate Submission Agreement
 - Verträge mit Produzenten, Übergabe, Prozedere
- Manage System Configuration
 - Systementwicklung, Überwachung der Funktion
 - Informationen für Policies
- Archival Information Update
 - Aktualisiert den Inhalt des Archivs: Änderung der DIPs und Re-Submission -> Migration
- Establish Standards and Policies
 - Budget, Standards, Policies
- Audit Submission
 - Analyse ob SIPs und AIPs den Vorschriften entsprechen
 - Verifikation der Representation und Package Information

Administration Functions

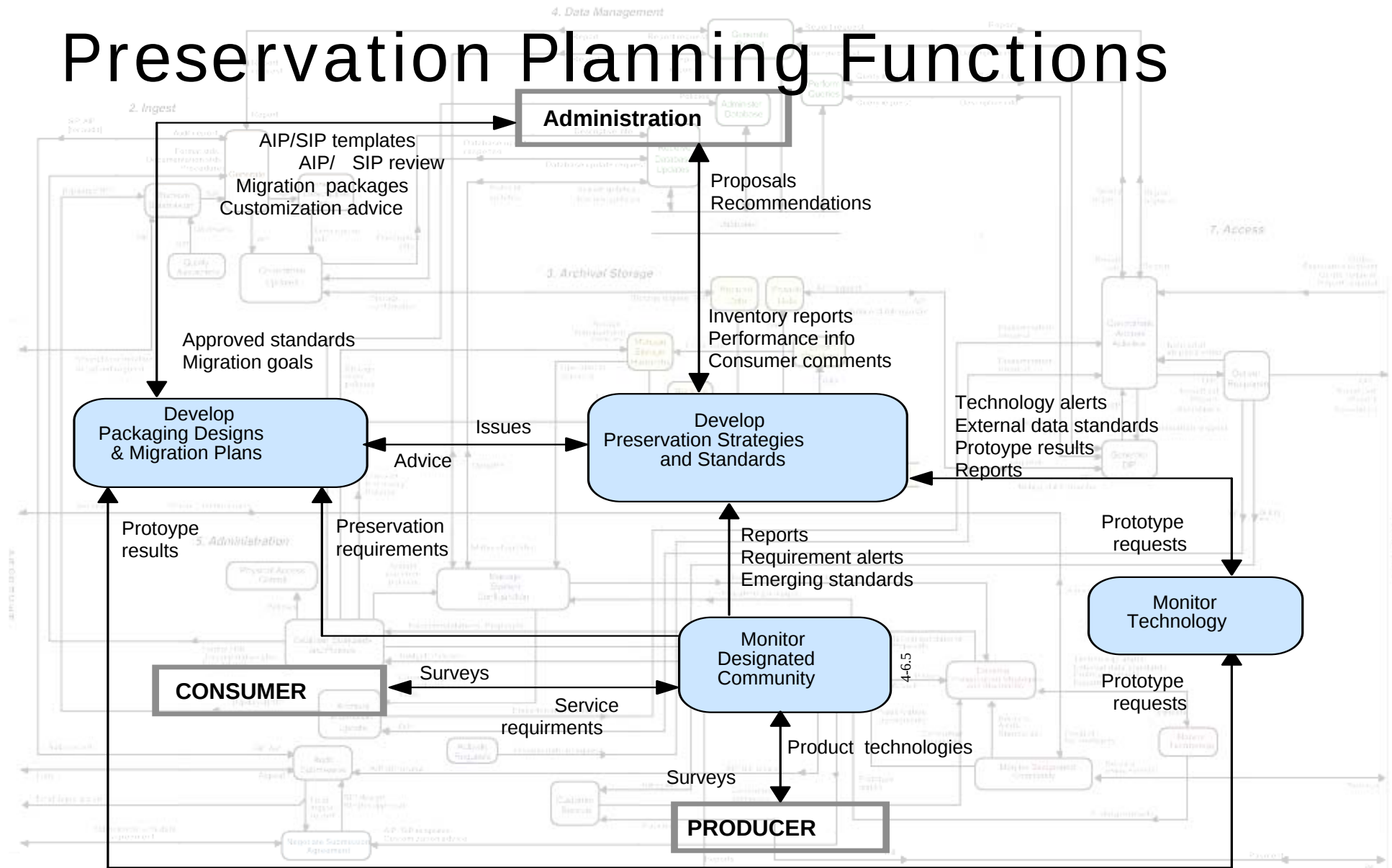
■ Activate Requests

- Aufzeichnungen über ereignisgesteuerte Abfragen
- regelmäßige Abfragen an Archiv um Vorhandensein der Daten zu verifizieren
- Bestellungen auf periodischer Basis

■ Customer Service

- Kundenaccounts
- sammelt Kosten von Access, erstellt Rechnungen für Kunden

Preservation Planning Functions



Preservation Planning Functions

- Monitor Designated Community
 - interaktion mit Produzenten und Kunden
- Monitor Technology
 - Technologieentwicklung: HW, SW, Formate
- Develop Preservation Strategies and Standards
 - Strategien, Entwicklungen, Vorhersage von Trends
- Develop Packaging Designs and Migration Plans
 - Migrationspfade, Tools
 - Erstellung von Preservation Description Information

Digital Migration

Digital Migration is defined to be the transfer of digital information, while intending to preserve it, within the OAIS.

- Focus on preservation of the full information content
- New information implementation replaces the old
- OAIS has full control and responsibility over all aspects of the transfer

Migration Motivators

■ Motivators driving digital migrations

– Media Decay

- Often this is superceded by escalating media drive maintenance costs

– Increased Cost Effectiveness

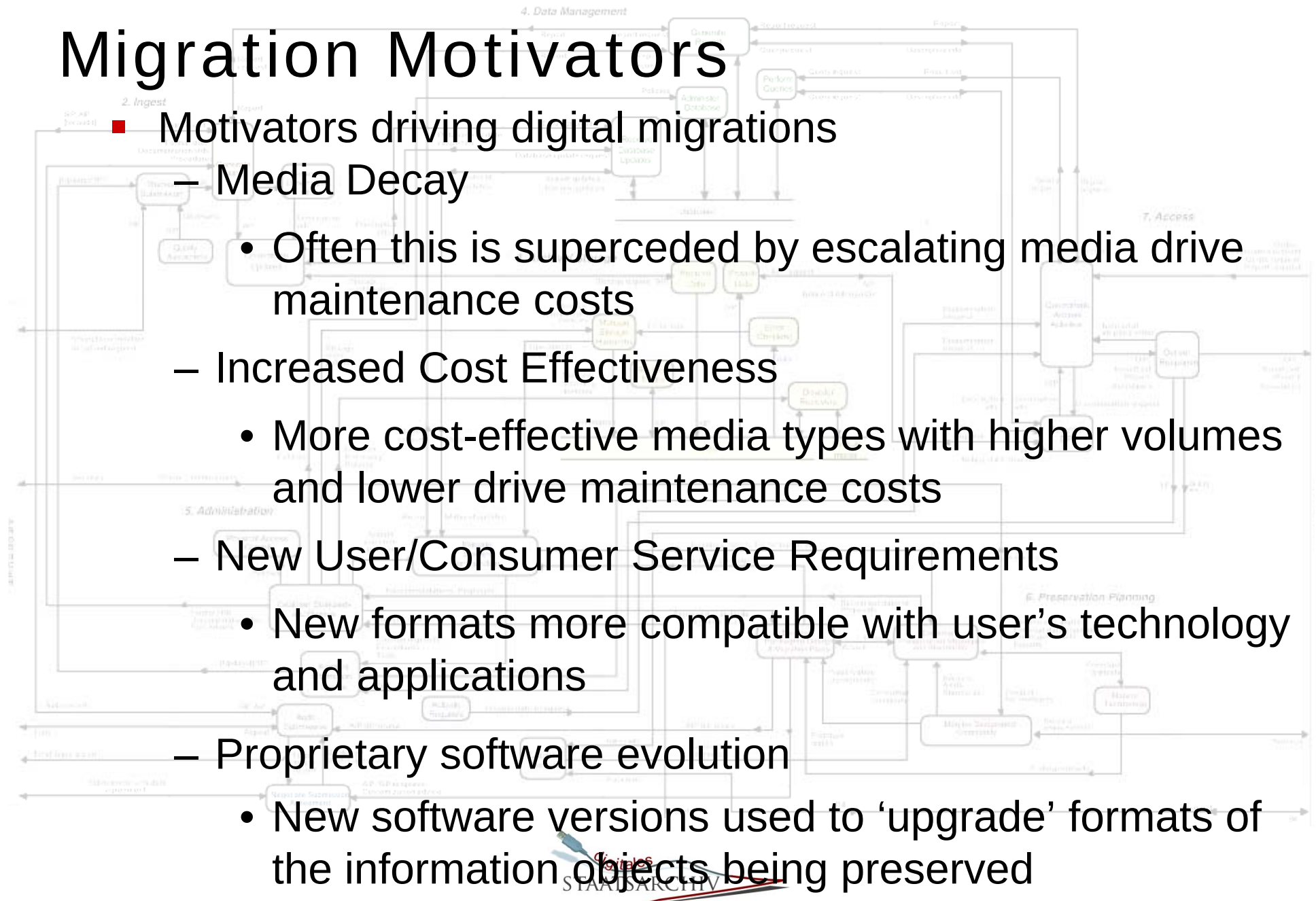
- More cost-effective media types with higher volumes and lower drive maintenance costs

– New User/Consumer Service Requirements

- New formats more compatible with user's technology and applications

– Proprietary software evolution

- New software versions used to 'upgrade' formats of the information objects being preserved



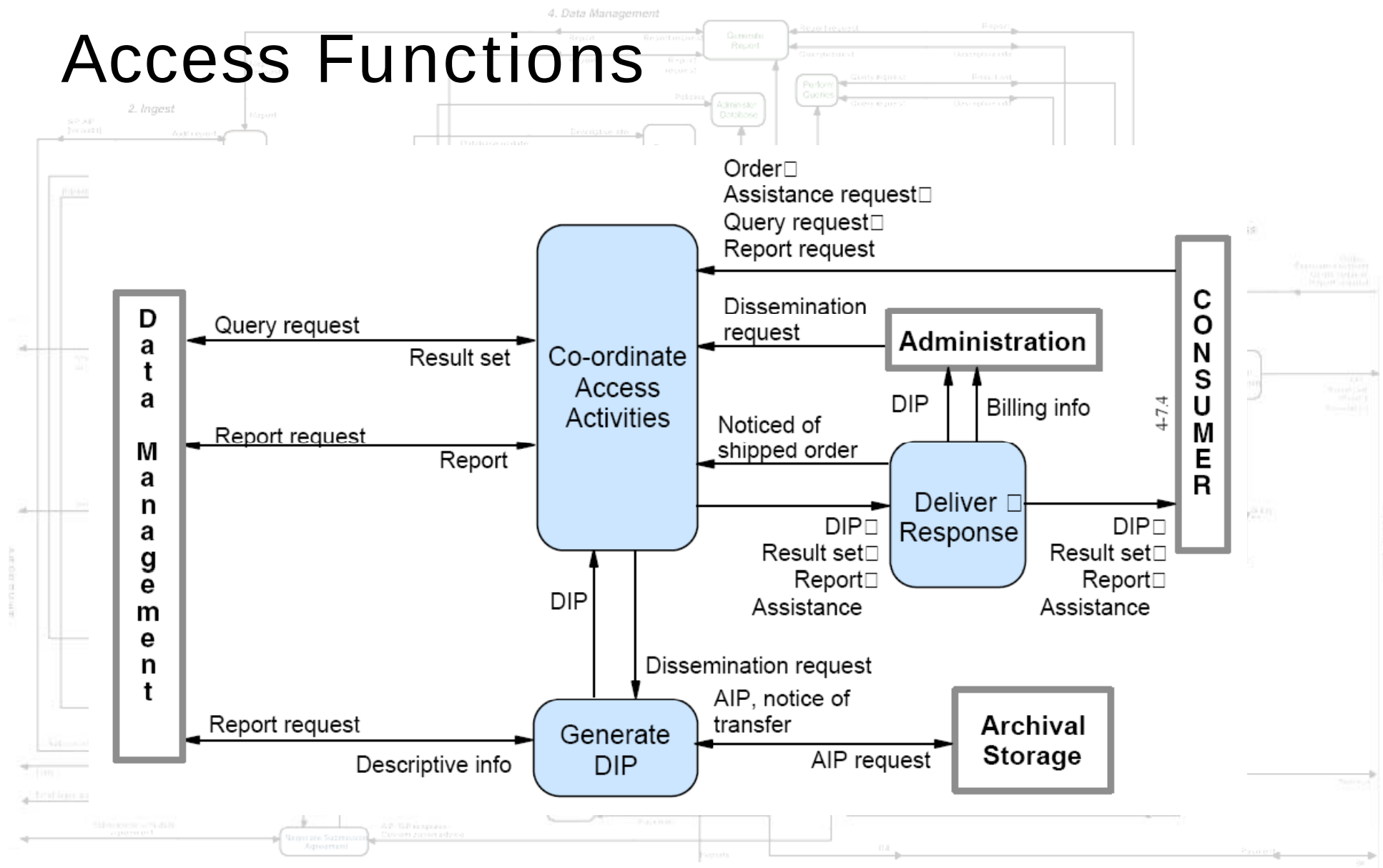
Digital Migration Approaches

- Four primary types of digital migration in response to motivators, ordered by increasing risk of information loss:
 - Refreshment
 - Media replacement with no bit changes
 - Replication
 - No change to Packaging Information or Content Information bits
 - Repackaging
 - Some bit changes in Packaging Information
 - Transformation
 - Reversible: Bit changes in Content Information are reversible by an algorithm
 - Non-reversible: Bit changes in Content Information are not reversible by an algorithm

Digital Migration and AIPs

- Unless migration involves transformation:
no new AIP version
- Transformation:
new **AIP Version**
- Upgrading or improvement of AIPs:
new **AIP Edition**
- Extracting or aggregating from multiple AIPs:
Derived AIP

Access Functions



Access Functions

■ Coordinate Access Activities

– Benutzerschnittstelle, Autorisierung

– 3 Arten von Requests:

- Anfragen für Data Management mit Result Set
- Bestellungen für Data Management und Archival Storage
- Dissemination Requests durch Administration für Archival Information Update

■ Generate DIP

- Holt Daten aus Archival Storage in Staging Area
- Holt Descriptive Information von Data Management
- Anwendung von Prozessen zur Transformatin des AIPs in ein geeignetes DIP für die jeweilige Anfrage

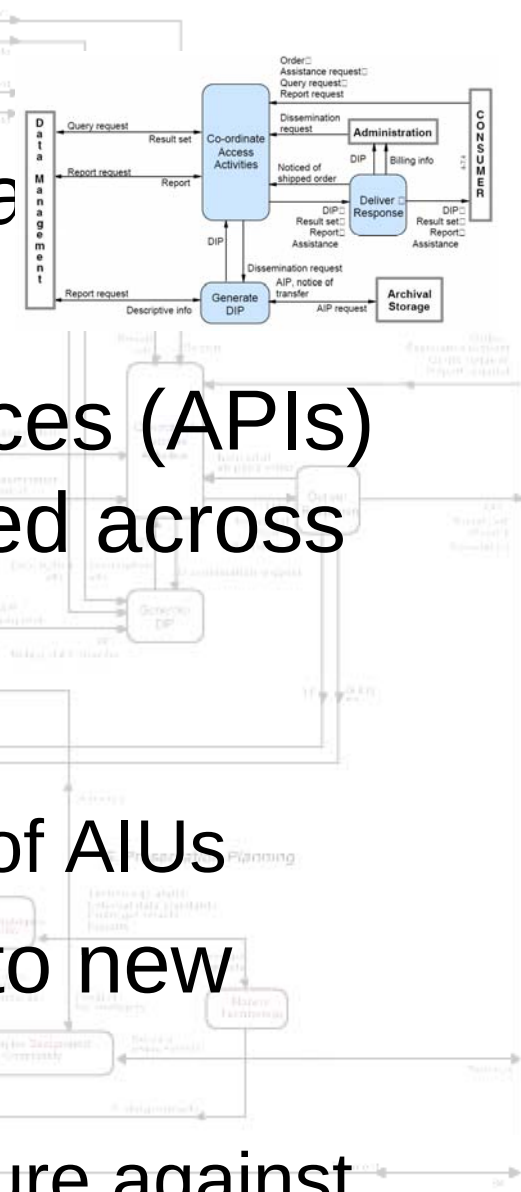
■ Deliver Response

- on-line und off-line Anfragen
- Weiterleitung des Ergebnisses

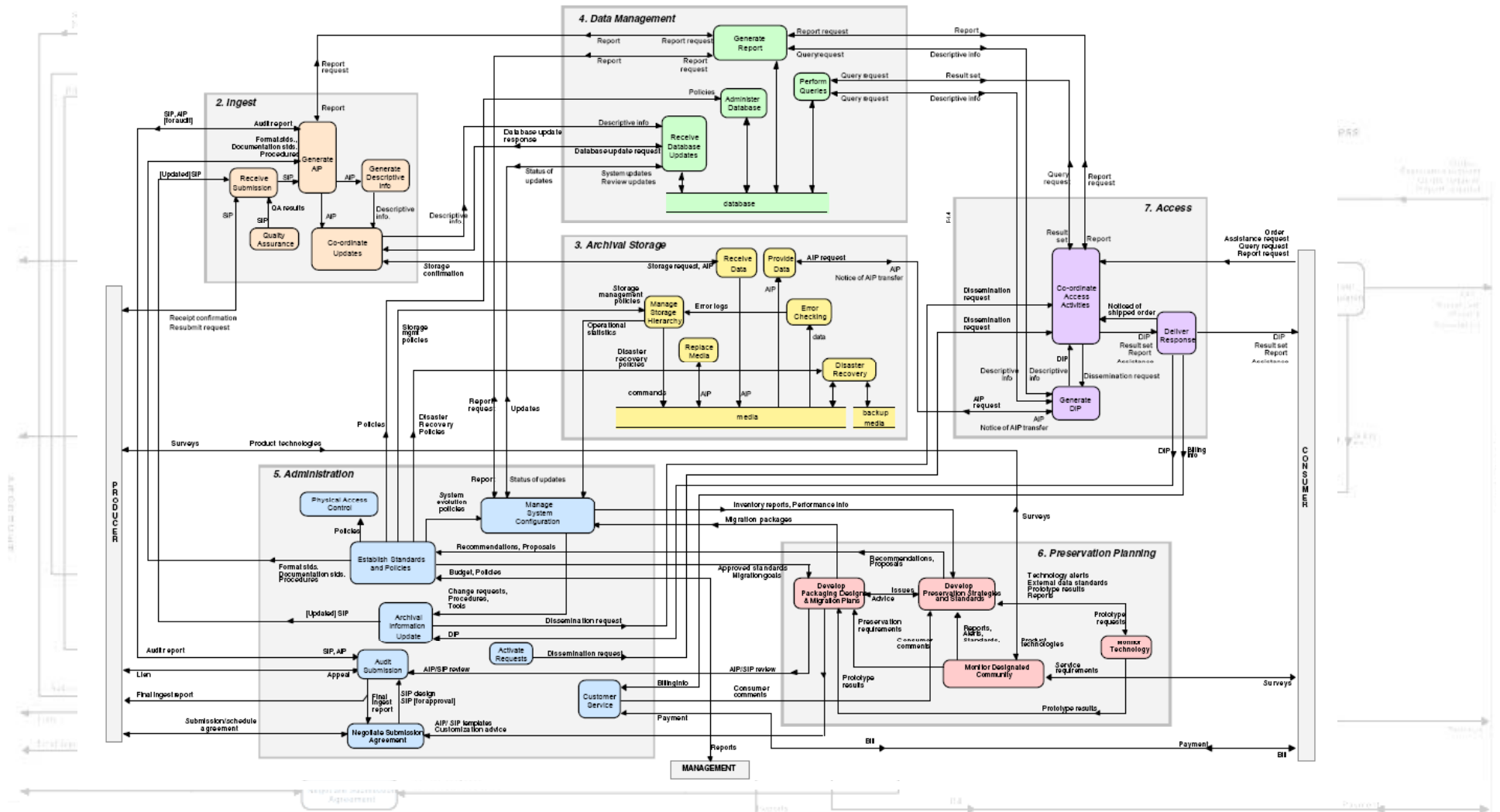


Access Preservation

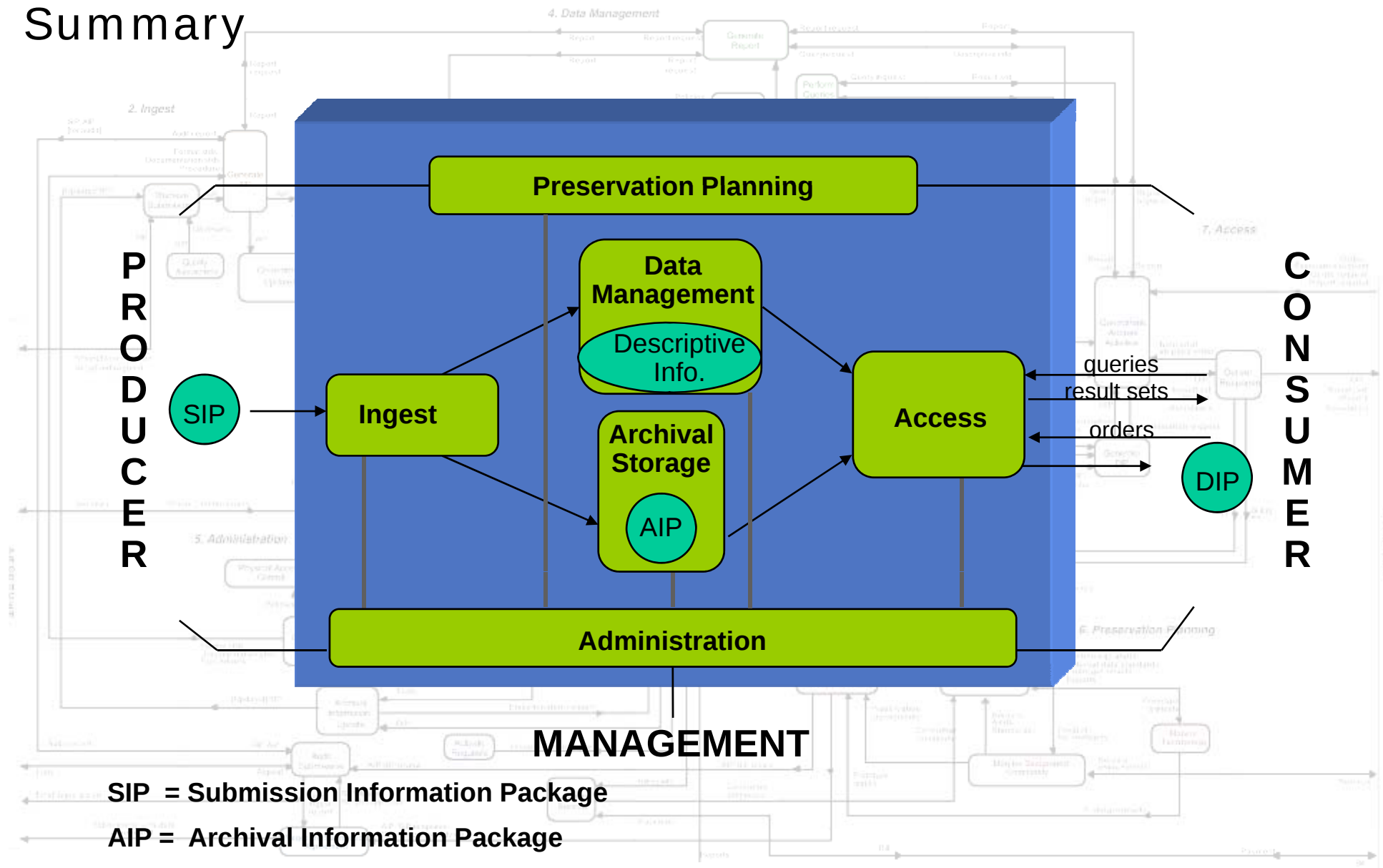
- Effective access to digital information requires the use of software
- Application Programming Interfaces (APIs) may be cost-effectively maintained across time by an OAIIS when:
 - API is not too complex
 - API is applicable to a wide variety of AIUs
- API source code may be ported to new environments
 - Extensive testing is needed to ensure against information loss



OAIS Composite Functional Entities



Open Archival Information System: Summary



SIP = Submission Information Package

AIP = Archival Information Package

DIP = Dissemination Information Package

Überblick

- Prinzipien des OAIS Modells
- Technischer Überblick
- Funktionaler Überblick
- Informationsmodell
- Zusammenfassung:
wichtige Komponenten -> DigLAimBund

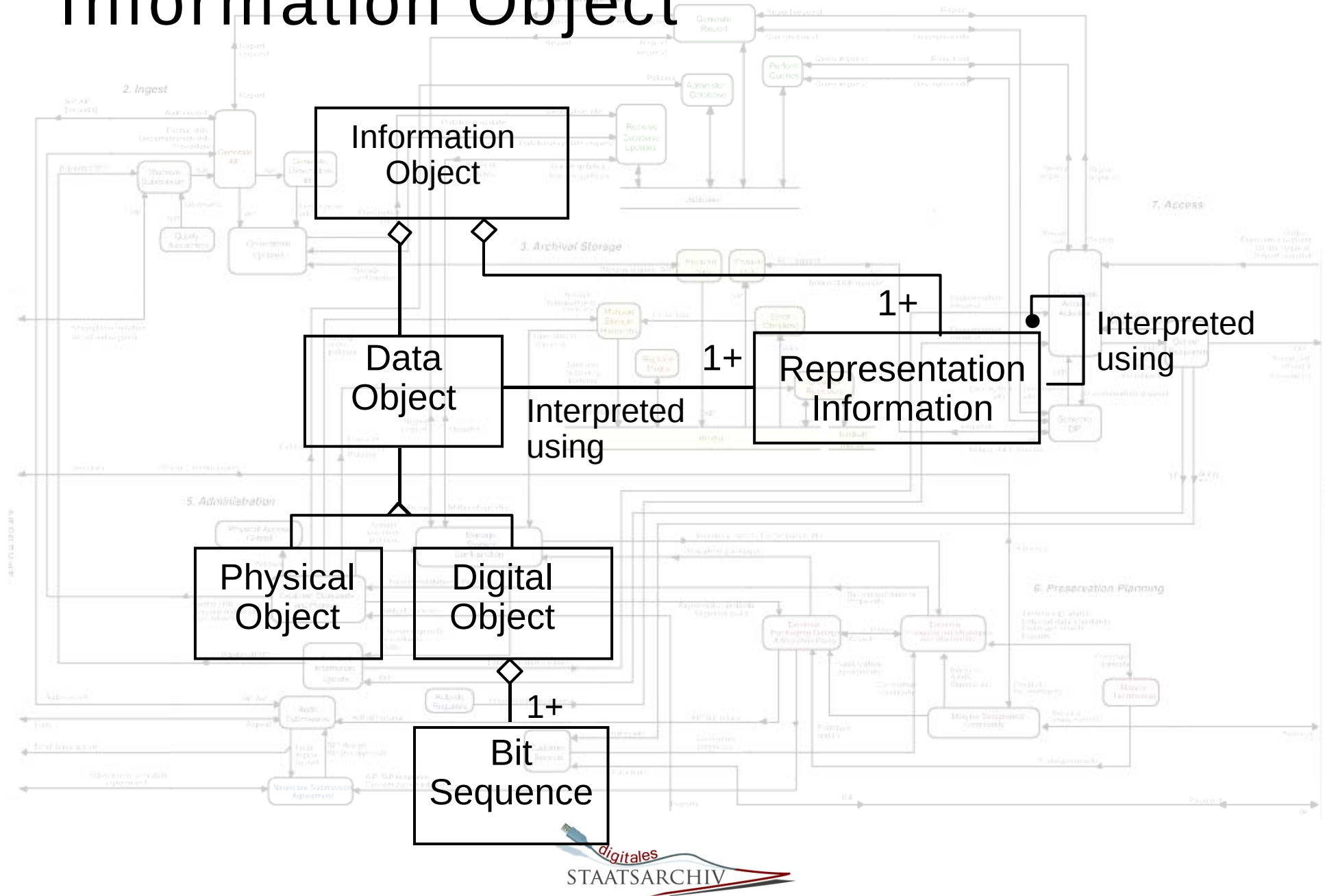
Information Package Definition

Content
Information

Preservation
Description
Information

- An Information Package is a conceptual container holding two types of information
 - Content Information
 - Preservation Description Information (PDI)

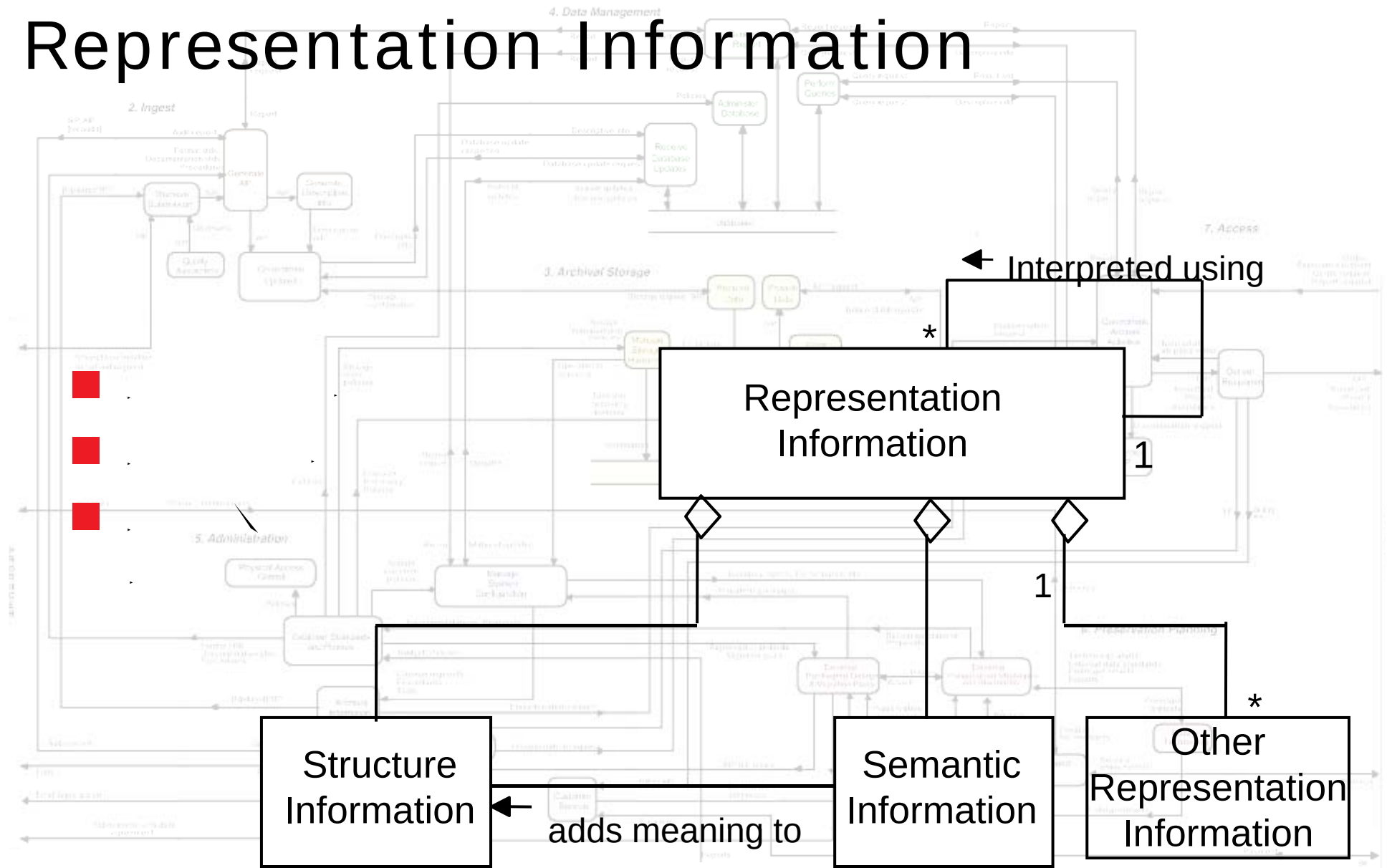
Information Object



Representation Information

- The Representation Information accompanying a physical object, like a moon rock, may give additional meaning
 - It typically is a result of some analysis of the physically observable attributes of the rock
- The Representation Information accompanying a digital object, or sequence of bits, is used to provide additional meaning.
 - It typically maps the bits into commonly recognized data types such as character, integer, and real and into groups of these data types.
 - It associates these with higher level meanings which can have complex inter-relationships that are also described

Recursive Nature of Representation Information



The diagram illustrates the relationships between various digital preservation components, organized into seven main sections:

- 1. Ingest**: Processes incoming data, including document generation and storage.
- 2. Archival Storage**: Manages the storage of digital objects, including metadata and access control.
- 3. Access**: Provides mechanisms for retrieving and viewing digital objects.
- 4. Administration**: Handles system management, including user roles and permissions.
- 5. Preservation Planning**: Develops strategies for long-term digital preservation.
- 6. Multimedia Document Representation Information**: The central focus, containing:
 - Multimedia Mapping Rules**: Rules for mapping multimedia content.
 - Additional Semantics**: Extra information about the content.
 - JPEG Representation Information**: Information about JPEG images.
 - HTML Representation Information**: Information about HTML documents.
 - JPEG Mapping Rules**: Rules for mapping JPEG content.
 - Other Representation Objects**: Additional objects for representation.
- 7. ASCII**: Provides a text-based representation of the digital objects.

The diagram also shows the flow of data between these components and the underlying systems, including databases, servers, and user interfaces.

The diagram illustrates the types of information used in OAI (Open Archival Information Interchange). It shows a central 'Information Object' (3. Archival Object) which is linked to several other information objects and processes. The main components are:

- Information Object (3. Archival Object)**: The central focus, with a triangle pointing to it from below.
- Content Information (5. Administration)**: A box on the left, linked to the central object.
- Preservation Description Information**: A box below the central object, linked to it.
- Packaging Information**: A box below the central object, linked to it.
- Descriptive Information**: A box on the right, linked to the central object.
- 2. Ingest**: A section on the top left, showing processes like 'Ingest request', 'Generate AIP', 'Generate Object', and 'Generate Object ID'.
- 7. Access**: A section on the top right, showing processes like 'Perform Access', 'Access request', and 'Access response'.
- 8. Preservation Planning**: A section on the bottom right, showing processes like 'Preservation planning', 'Preservation request', and 'Preservation response'.

The diagram uses arrows to show the flow of information between these objects and processes, indicating the relationships and data exchange within the OAI framework.

Content Information

- The information which is the primary object of preservation
- An instance of Content Information is the information that an archive is tasked to preserve.
- Deciding what is the Content Information may not be obvious and may need to be negotiated with the Producer
- The Data Object in the Content Information may be either a Digital Object or a Physical Object (e.g., a physical sample, microfilm)

Preservation Description Information

■ Provenance Information

– Describes the source of Content Information, who has had custody of it, what is its history

■ Context Information

– Describes how the Content Information relates to other information outside the Information Package

■ Reference Information

– Provides one or more identifiers, or systems of identifiers, by which the Content Information may be uniquely identified

■ Fixity Information

– Protects the Content Information from undocumented alteration

PDI Examples

Content Information Type	Reference	Provenance	Context	Fixity
Space Science Data	- Object identifier - Journal reference - Mission, instrument, title, attribute set	Instrument description Processing history Sensor description Instrument Instrument mode Decommulation map Software interface specification	Calibration history Related data sets Mission Funding history	CRC Checksum Reed-Solomon coding
Digital Library Collections	Bibliographic description Persistent identifier	For scanned collections: metadata about the digitisation process pointer to master version For born-digital publications: pointer to the digital original Metadata about the preservation process: pointers to earlier versions of the collection item change history	Pointers to related documents in original environment at the time of publication	Digital signature Checksum Authenticity indicator
Software Package	Name Author/Originator Version number Serial number	Revision history License holder Registration Copyright	Help file User guide Related software Language	Certificate Checksum Encryption CRC

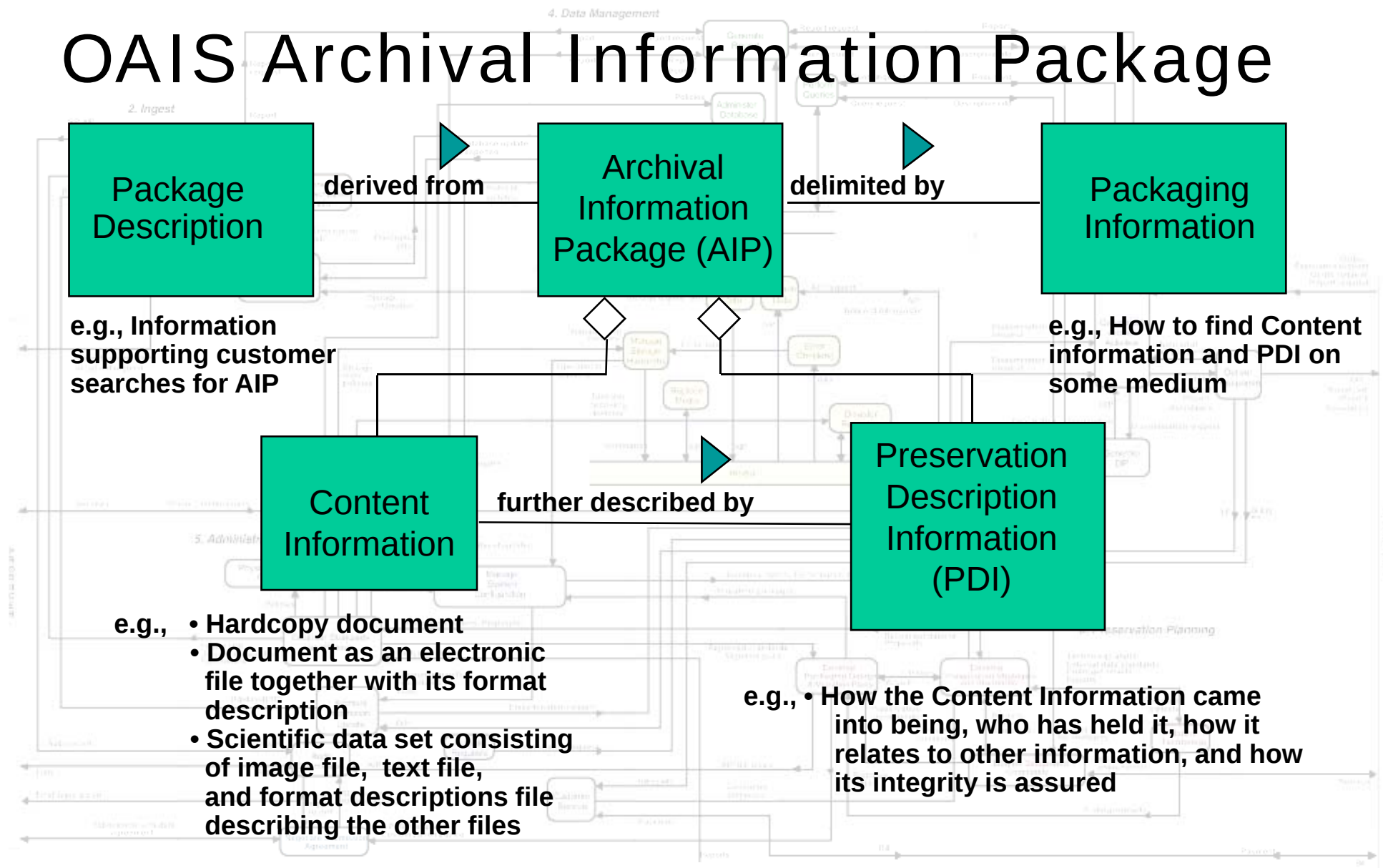
Descriptive Information

- Contain the data that serves as the input to documents or applications called Access Aids.
- Access Aids can be used by a consumer to locate, analyze, retrieve, or order information from the OAIS.

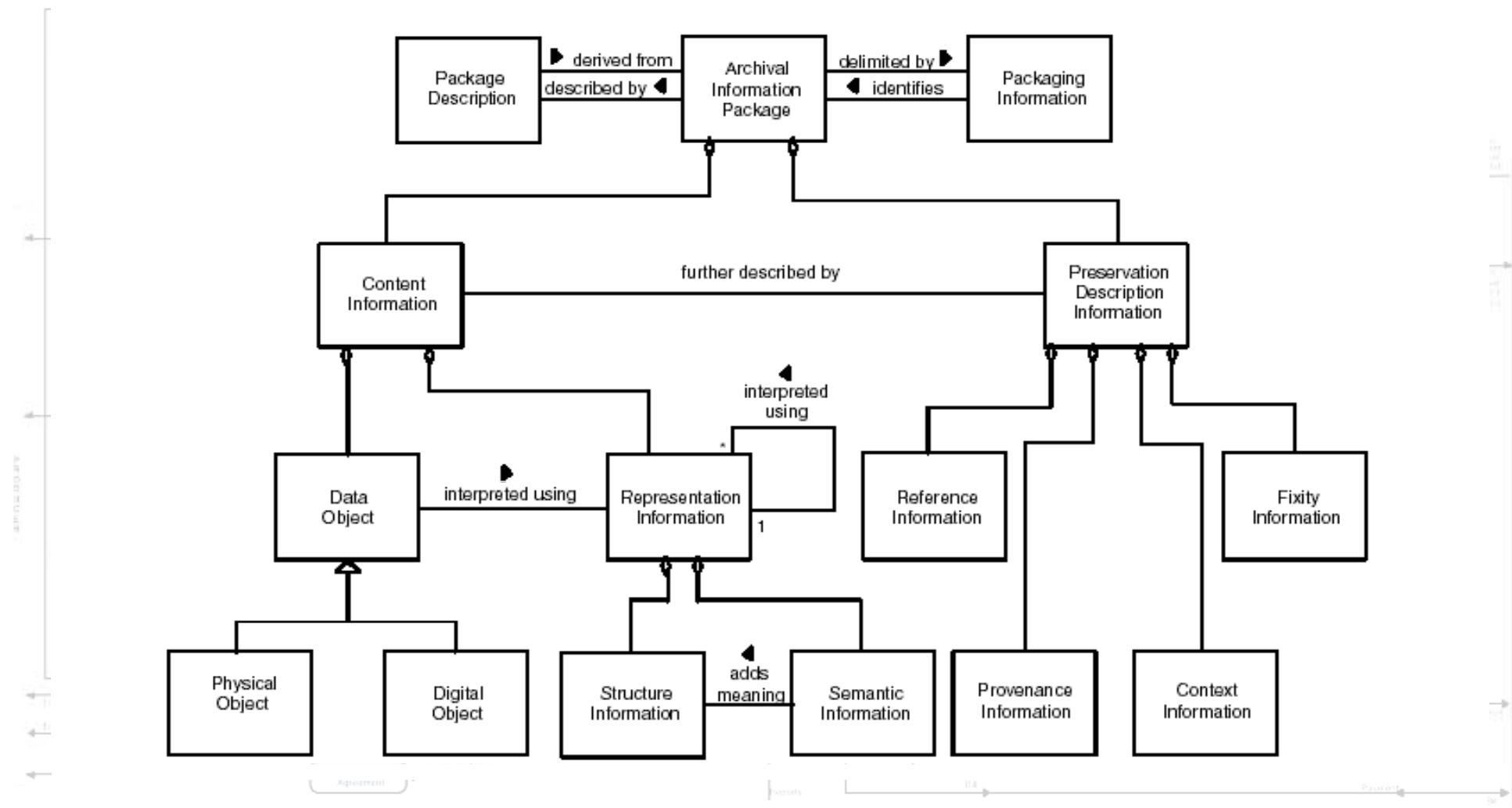
Packaging Information

- Information which, either actually or logically, binds and relates the components of the package into an identifiable entity on specific media
- Examples of Packaging Information include tape marks, directory structures and filenames

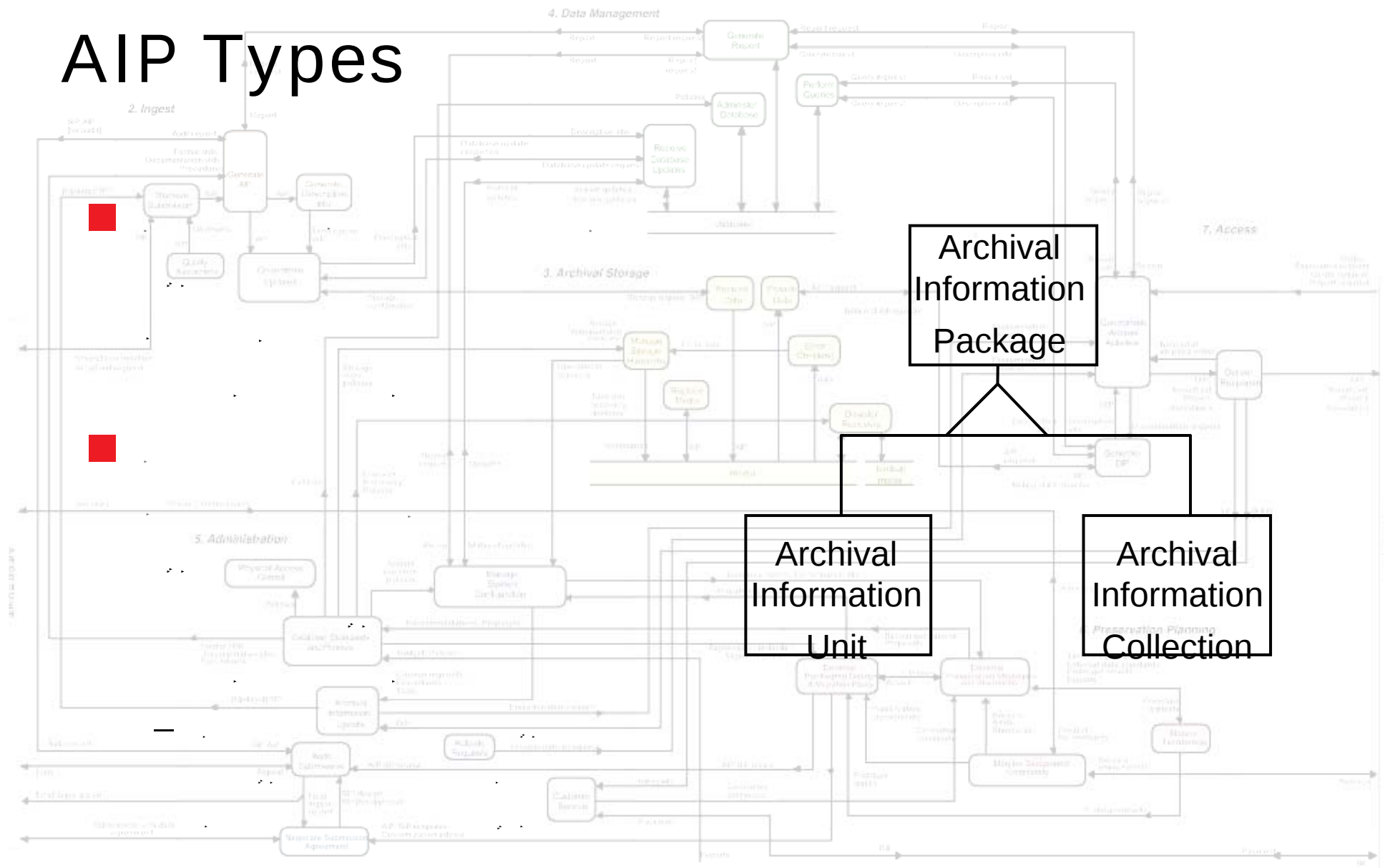
OAIS Archival Information Package



AIP detailed view



AIP Types



Package Descriptions and Access Aids

- Package Descriptions are needed by an OAIS to provide visibility and access to the OAIS holdings
- Package Descriptions contain 1 or more Associated Descriptions which describe the AIP Content Information from the point of view of a single Access Aid
- Some example of Access Aids Include:
 - Finding Aids - assist the consumer in locating information of interest

Information Model Summary

- Presented a model of information objects as containing data objects and representation objects
- Classified information required for Long-term archiving into 4 classes: Content Information, PDI, Packaging Information and Descriptive Information
- Described how these classes would be aggregated and related in an AIP to fully describe an instance of Content Information
- Presented information needed for Access,

Zusammenfassung

- OAIS ist ein Referenzmodell
- OAIS ist KEINE Umsetzungsanleitung
- Regelt Sprachgebrauch, Rollen, Aufgabenbereiche, Funktionalitäten
- Anwendbar für alle Archive, Organisationen, Individuen
- Anwendbar für alle Arten von Informationsobjekten