

Preservation Planning im Digitalen Archiv Österreich

Hannes Kulovits

11.3.2015

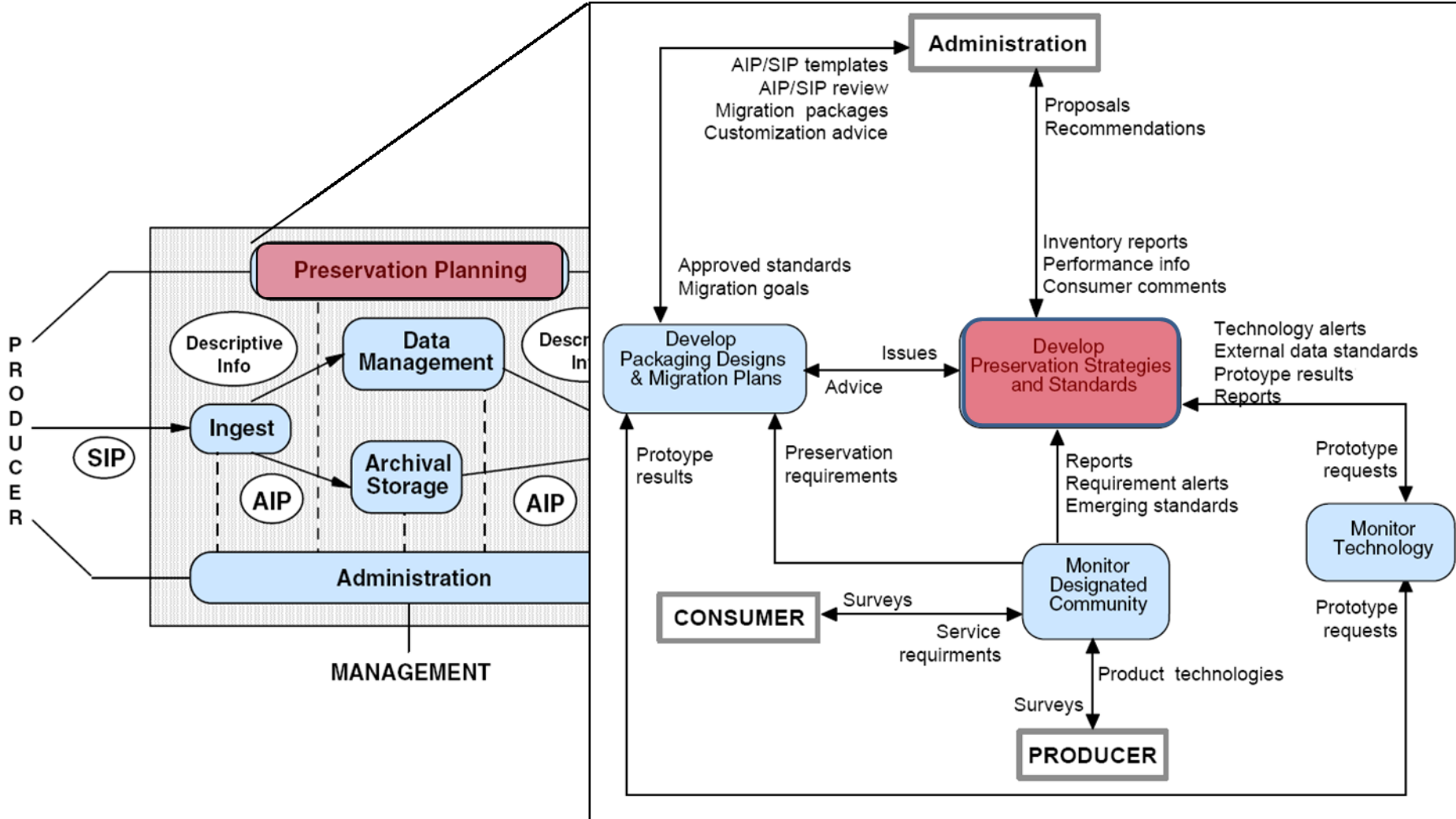
Agenda

- Preservation Planning/Operations
- Synchronisation mit AIS
- Showcase

Agenda

- Preservation Planning/Operations
- Synchronisation mit AIS
- Showcase

Preservation Planning im OAIS



Nestor Kriterienkatalog und PP

A2 Das digitale Langzeitarchiv ermöglicht seinen Zielgruppe(n) eine angemessene Nutzung der durch die digitalen Objekte repräsentierten Informationen.

[...] Die Nutzung setzt den Erhalt der digitalen Objekte und deren Verfügbarkeit sowie die Sicherstellung der Interpretierbarkeit voraus. Die Nutzung kann angemessen sein, obwohl [...] nicht alle Eigenschaften des Originals erhalten wurden

Nestor Kriterienkatalog und PP

B9.2 Das digitale Langzeitarchiv identifiziert, welche Eigenschaften der digitalen Objekte für den Erhalt von Information signifikant sind.

Bei der Entscheidung über den Umfang der zu bewahrenden Eigenschaften ist vor dem Hintergrund der eigenen Ziele zwischen den technischen Möglichkeiten sowie dem Aufwand für die Langzeitarchivierung einerseits und den Bedürfnissen der Zielgruppe(n) andererseits abzuwägen.

Es kann notwendig sein, die digitalen Objekte in mehreren Varianten vorzuhalten, um möglichst viele Eigenschaften zu erhalten.

TRAC and Preservation Planning

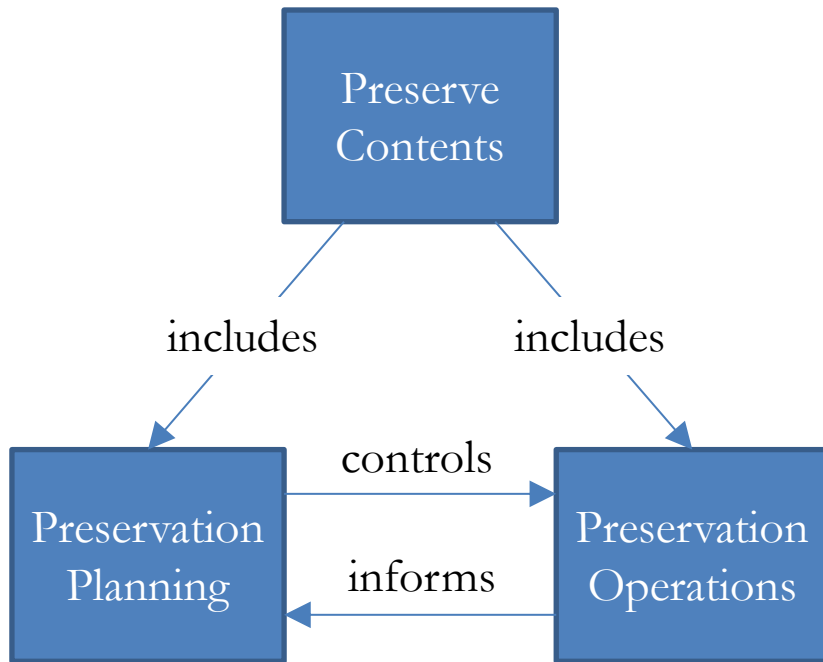
A3.6 Repository has a documented history of the changes to its operations, procedures, software, and hardware that, where appropriate, is linked to relevant preservation strategies and describes potential effects on preserving digital content.

TRAC and Preservation Planning

B3.1 Repository has documented preservation strategies.

B3.3 Repository has mechanisms to change its preservation plans as a result of its monitoring activities.

Preservation capabilities



- Analysis (“characterisation”)
 - Identification
 - Feature extraction
 - Validation
- Actions
 - Migration
 - Emulation
 - Others
- Quality Assurance
 - Authenticity
 - Significant properties
- Metadata
- Reporting

Welche zwei Dateien sind ähnlich?

- Nimmt man die folgenden drei Dateien: A, B,

	A	B	C
Format	PDF 1.2	PDF 1.2	PDF 1.4



Welche zwei Dateien sind ähnlich?

	A	B	C
Format	PDF 1.2	PDF 1.2	PDF 1.4
Page count	20	1.700	40
Encryption	Yes	No	Yes
File size	1MB	65 MB	2 MB
Valid	no	yes	No
Well-formed	Yes	yes	Yes
Digital signature	no	yes	no

... das Dateiformat ist lediglich eine Eigenschaft von vielen.

Eigenschaften von Bedeutung

- Format und Sub-Format
- Representation Instance Properties
- Information Properties

Formatrisiken

- Verfügbare freie und Open Source Software (Anzahl)
- Verfügbarkeit der Dokumentation
- Qualität der Dokumentation
- Standardisiert
- Identifikation (einfach/komplex)
- Validierung
- Lizenzkosten
- Browserunterstützung
- ...

Representation Instance Properties

- Well-formed
- Valid
- Dateigröße
- Kompression
- Dokument durchsuchbar
- Dokument maschinenlesbar
- Eingebettete Metadaten gültig
 - Z.B. EXIF

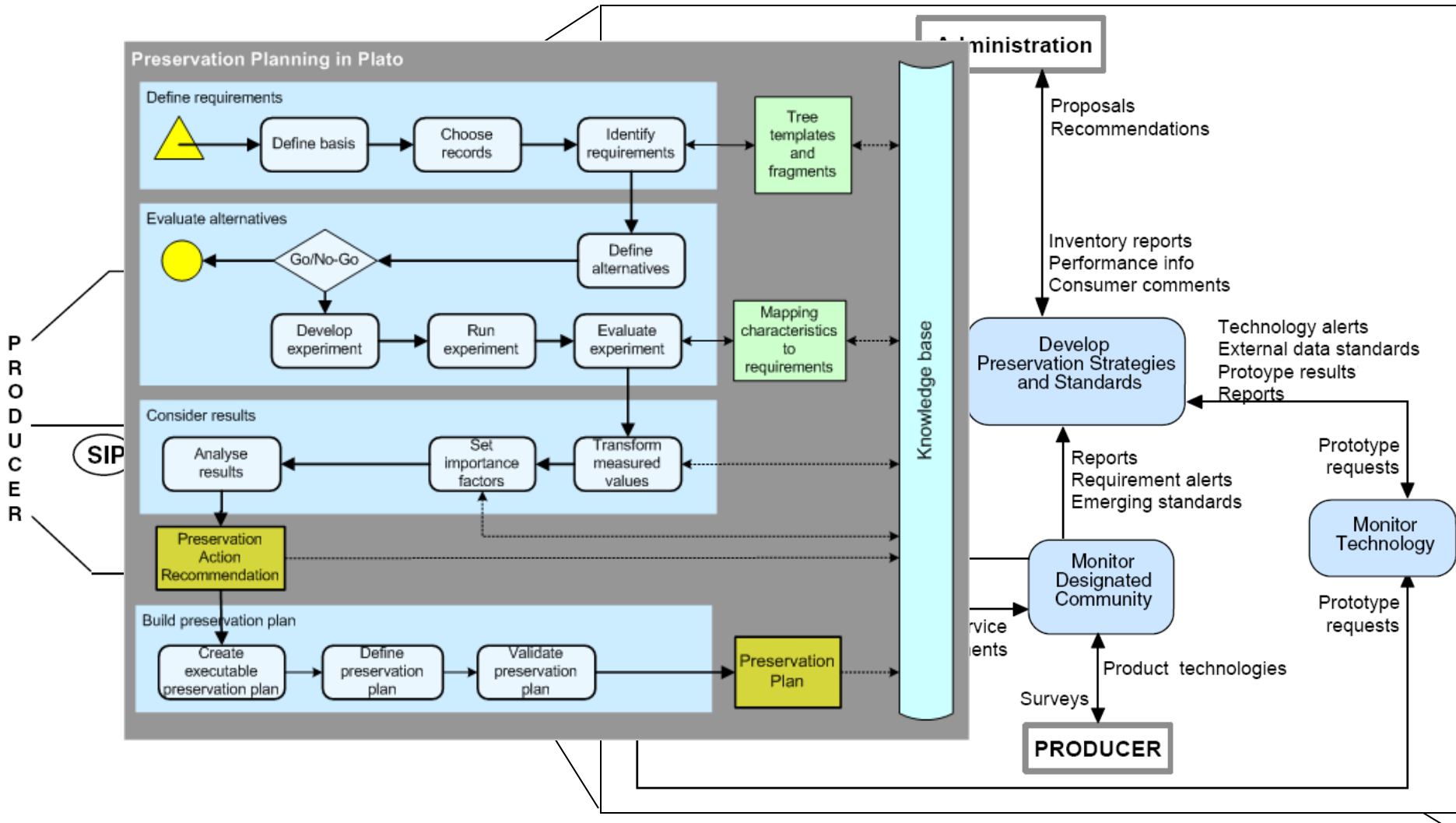
Information Properties

- Dokumente
 - Anzahl Seiten
 - Anzahl Zeichen
 - Kopfzeile, Fußzeile
 - Inhaltsverzeichnis
- Image
 - Höhe, Breite
 - Bits per sample
 - Farbraum
 - Farbprofil

Unterstützung im Digitalen Archiv

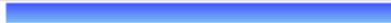


















- Analysis (“characterisation”)
 - Identification
 - Feature extraction
 - Validation
- Actions
 - Migration
 - Emulation
- Quality Assurance
 - Authenticity
 - Significant properties
- Metadata
- Reporting

Preservation Planning im OAIS



Results: Weighted multiplication















Result-Tree with all Alternatives, Aggregation method: Weighted multiplication

Node	Results
<input checked="" type="checkbox"/> Requirements	Archive to XML: 3.88  Keep original DB: 3.72  CSV export: 0.00
<input checked="" type="checkbox"/> Object characteristics	Archive to XML: 2.14  Keep original DB: 2.32  CSV export: 0.00
<input checked="" type="checkbox"/> Content	Archive to XML: 2.38  Keep original DB: 2.51  CSV export: 0.00
<input checked="" type="checkbox"/> appearance	Archive to XML: 1.14  Keep original DB: 1.16  CSV export: 1.06 
<input checked="" type="checkbox"/> context	Archive to XML: 1.23  Keep original DB: 1.17  CSV export: 0.88 
<input type="checkbox"/> behaviour	Archive to XML: 1.00  Keep original DB: 1.12  CSV export: 1.00 
<input checked="" type="checkbox"/> Format characteristics	Archive to XML: 1.36  Keep original DB: 1.20  CSV export: 1.13 
<input checked="" type="checkbox"/> Tool characteristics	Archive to XML: 1.34 

Results: Weighted sum

Result-Tree with all Alternatives, Aggregation method: Weighted sum.

This tree contains only strategies that do not have knock-out evaluation criteria; see above

Node	Results
<input checked="" type="checkbox"/> Requirements	Archive to XML: 4.25  Keep original DB: 4.07 
<input checked="" type="checkbox"/> Object characteristics	Archive to XML: 2.40  Keep original DB: 2.60 
<input checked="" type="checkbox"/> Content	Archive to XML: 2.63  Keep original DB: 2.76 
<input checked="" type="checkbox"/> appearance	Archive to XML: 0.41  Keep original DB: 0.45 
<input checked="" type="checkbox"/> context	Archive to XML: 0.69  Keep original DB: 0.57 
<input type="checkbox"/> behaviour	Archive to XML: 0.07  Keep original DB: 0.35 
<input checked="" type="checkbox"/> Format characteristics	Archive to XML: 0.95  Keep original DB: 0.57 

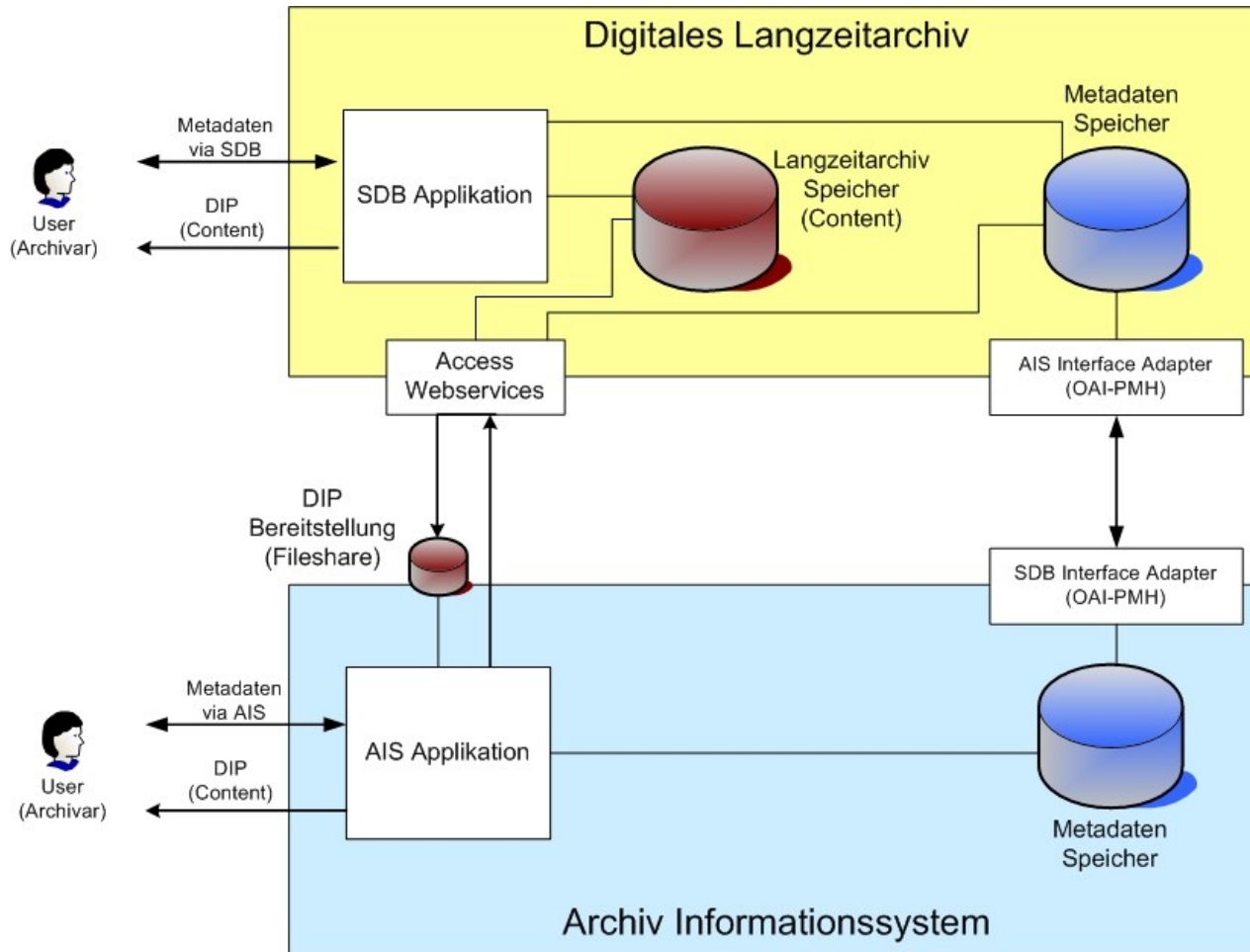
Agenda

- Preservation Planning/Operations
- **Synchronisation mit AIS**
- Showcase

AIS-Schnittstelle

- Vollständige (bidirektionale) Synchronisation von Metadaten zwischen Preservica und einem AIS
- Zugriffsanforderungen vom AIS an das SDB für Dissemination Information Packages (DIPs)
- Protokoll für Synchronisierung basierend auf OAI-PMH
 - OAI-PMH Datenanbieter (Server): Antwortet auf HTTP GET Requests mit Metadaten
 - OAI-PMH Diensteanbieter (Client): Erzeugt HTTP GET Anforderungen an einen bekannten OAI-PMH Datenanbieter und verarbeitet die zurückgesandten XML-Metadaten.

AIS-Schnittstelle (Aufbau)



Agenda

- Preservation Planning/Operations
- Synchronisation mit AIS
- **Showcase**

Vielen Dank für Ihre Aufmerksamkeit!