
Vertrauenswürdige digitale Langzeitarchive

Workshop

„Aus der Praxis der Langzeitarchivierung“
an der Bayerischen Staatsbibliothek München
11.05.2007

Peter Rödiger



Fakultät für Informatik
Institut für Softwaretechnologie
Prof. Dr. Uwe M. Borghoff
Tel.: 089/6004-2274 Fax: 089/6004-4447
Uwe.Borghoff@unibw.de
www.unibw.de/Uwe.Borghoff



Vertrauenswürdigkeit - Einführung

Herausforderungen bei der Langzeitarchivierung digitaler Information

- **Inhärente Komplexität digitaler Objekte:**
 - „Abstand“ physisches Medium und Information
 - Vielzahl möglicher interner (digitaler) Repräsentationsformen
 - Vielzahl möglicher externer Repräsentationsformen („Sichten“) durch (parametrisierbare) Operationen (Interaktion)
- **Langfristigkeit:**
 - Abstand Konsument von Information und Produzent von Information
 - Abstand Konsument von Information und Produzent / Nutzer von Werkzeugen (= technisches Interpretationswissen)
kurz: Technische Überalterung

Vertrauenswürdigkeit - Einführung

Komplexität und Bezug auf künftige Leistungen und Produkte erschweren Aussagen zur Qualität:

Keine direkte Messbarkeit bzw. Bewertbarkeit

Eingeführter Qualitätsbegriff: **Vertrauenswürdigkeit (Trustworthiness)**

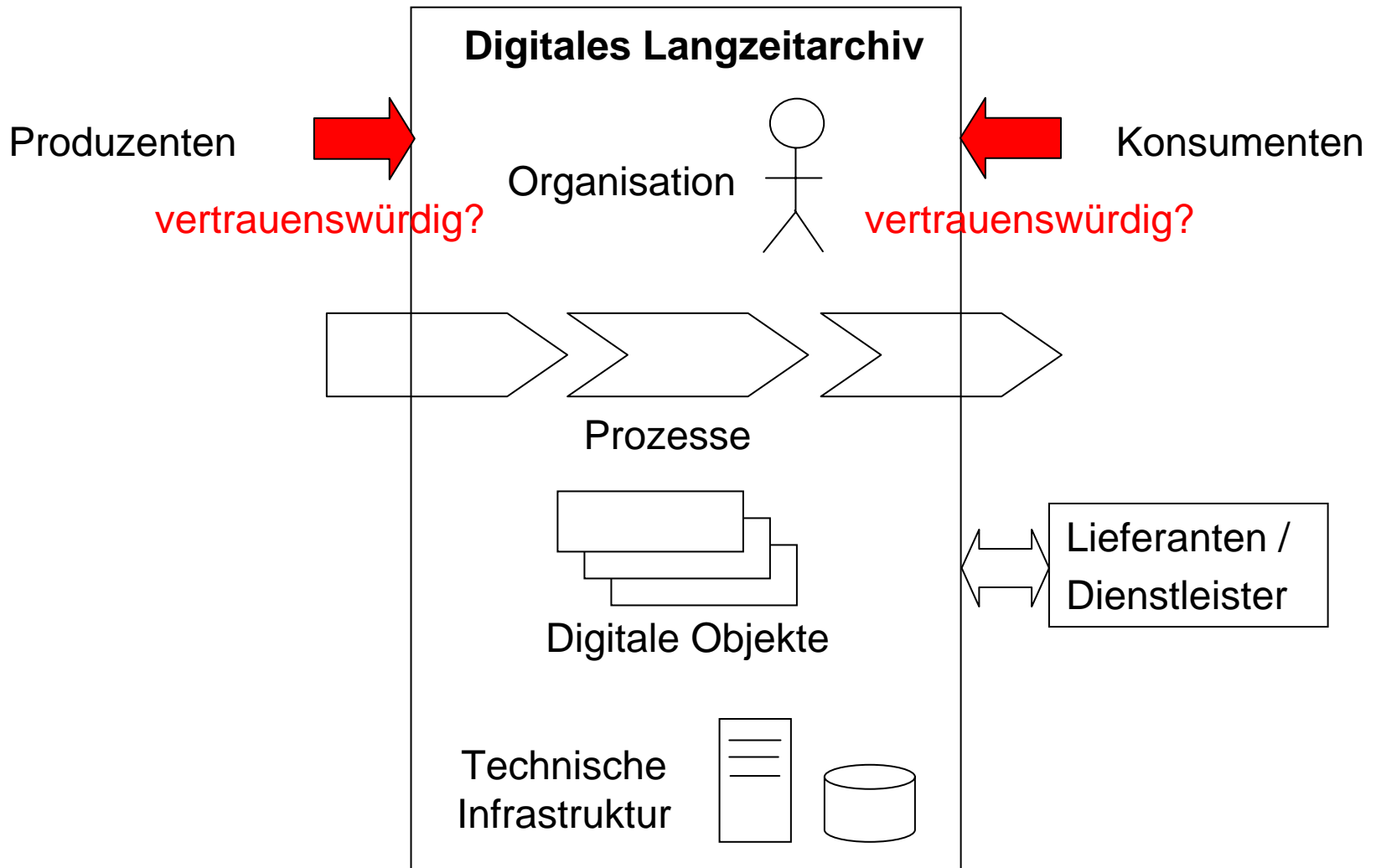
= Eigenschaft eines Systems, gemäß seinen Zielen und Spezifikationen zu operieren

(d.h. es tut genau das, was es zu tun vorgibt)

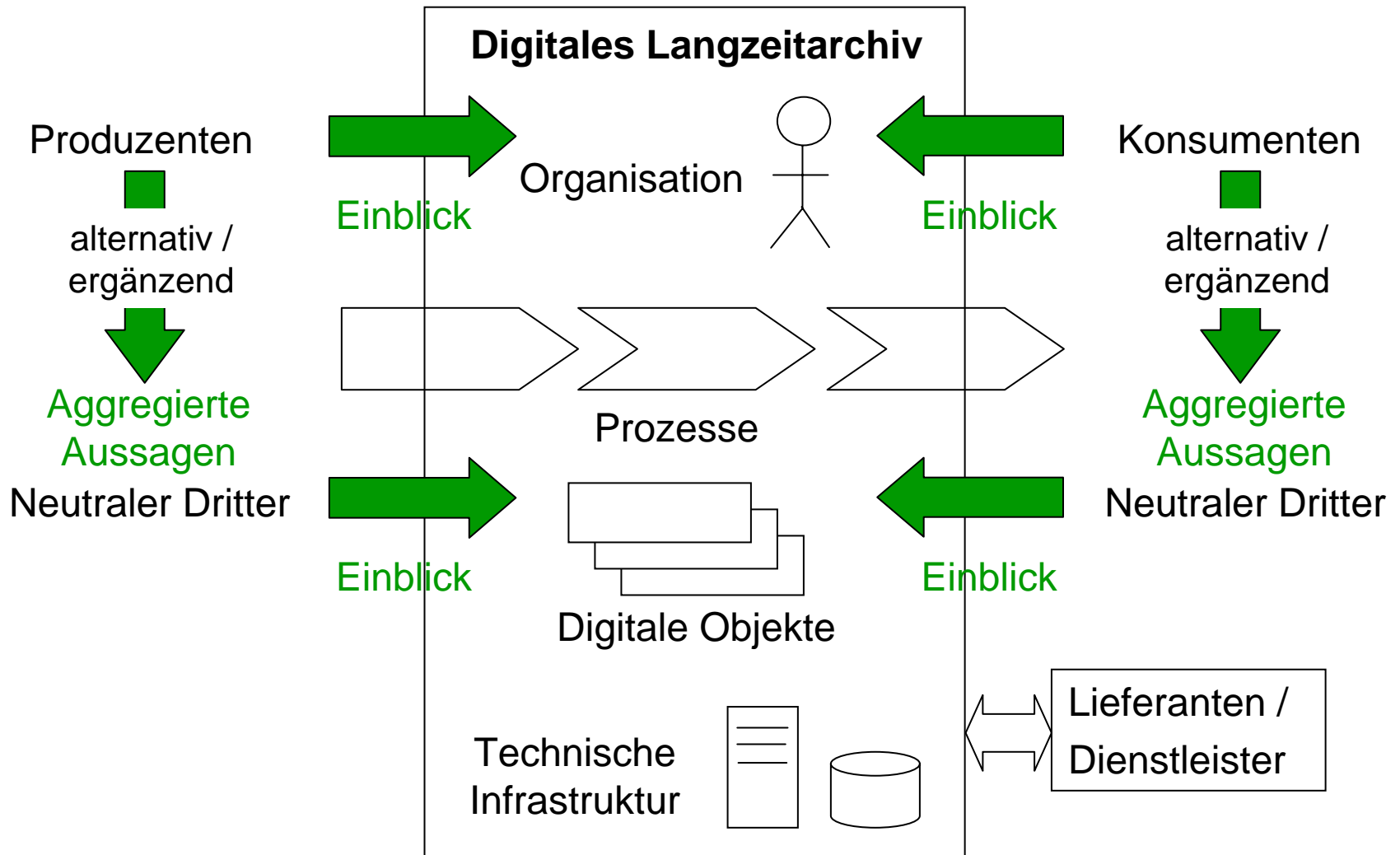
Vgl. Qualitätsstandards:

Qualität ist keine absolute Größe: Zusammenhang mit Anforderungen

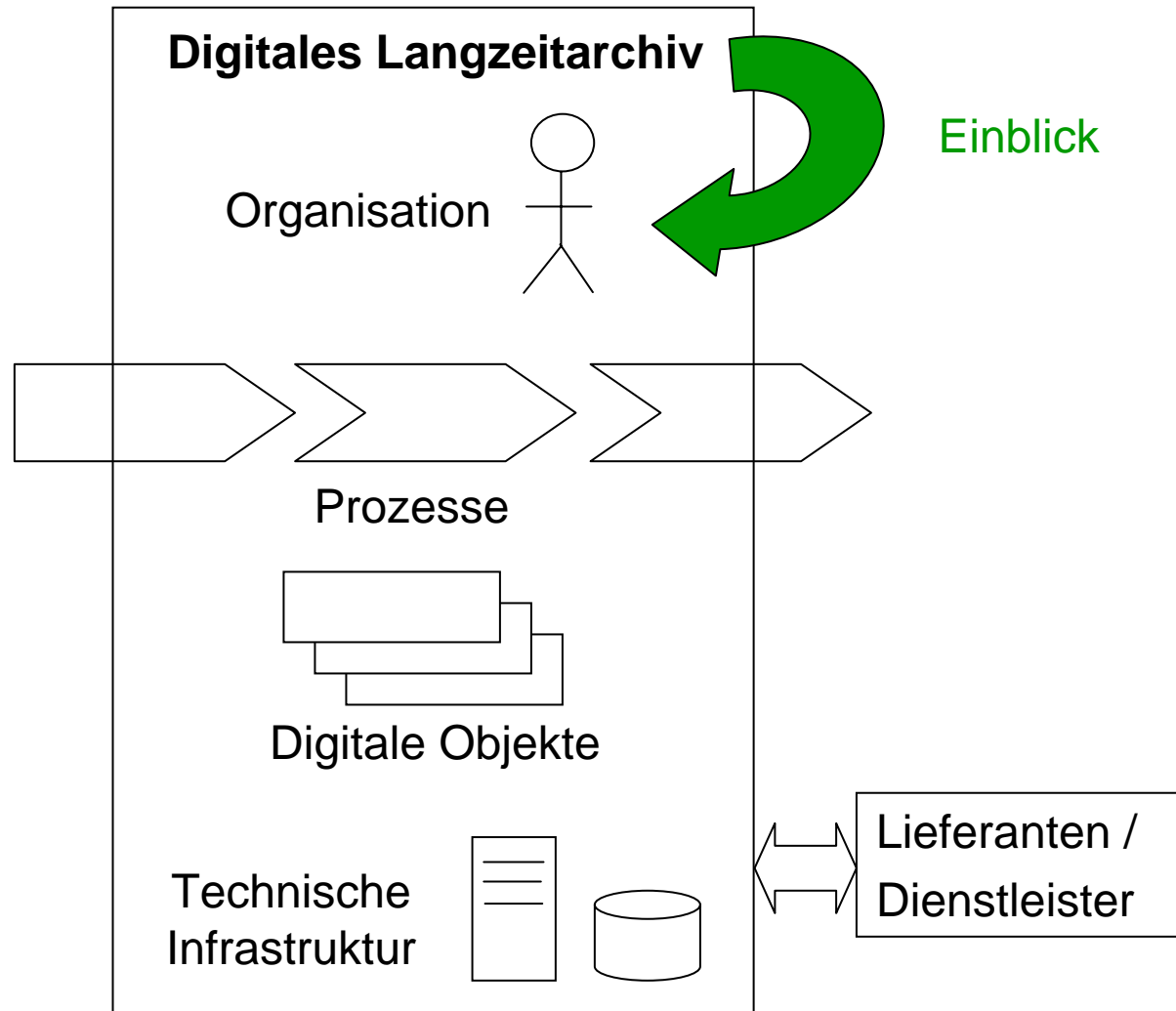
Architektur eines digitalen Langzeitarchivs



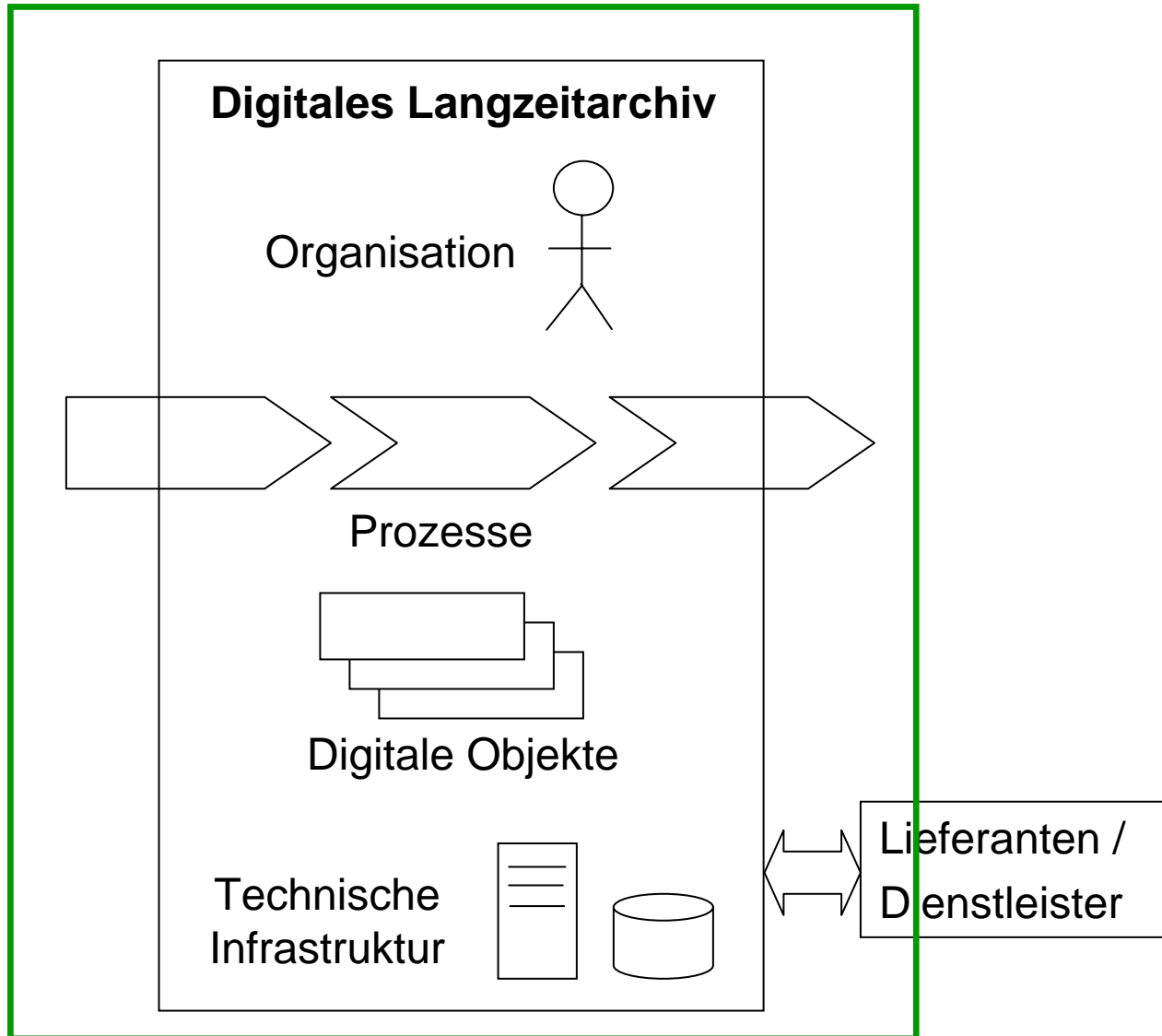
Prinzip der Transparenz - nach Außen



Prinzip der Transparenz - nach Innen



Basis für Transparenz: Dokumentation



Indikatoren für Vertrauenswürdigkeit

Indikatoren für Vertrauenswürdigkeit

Organisatorischer Rahmen

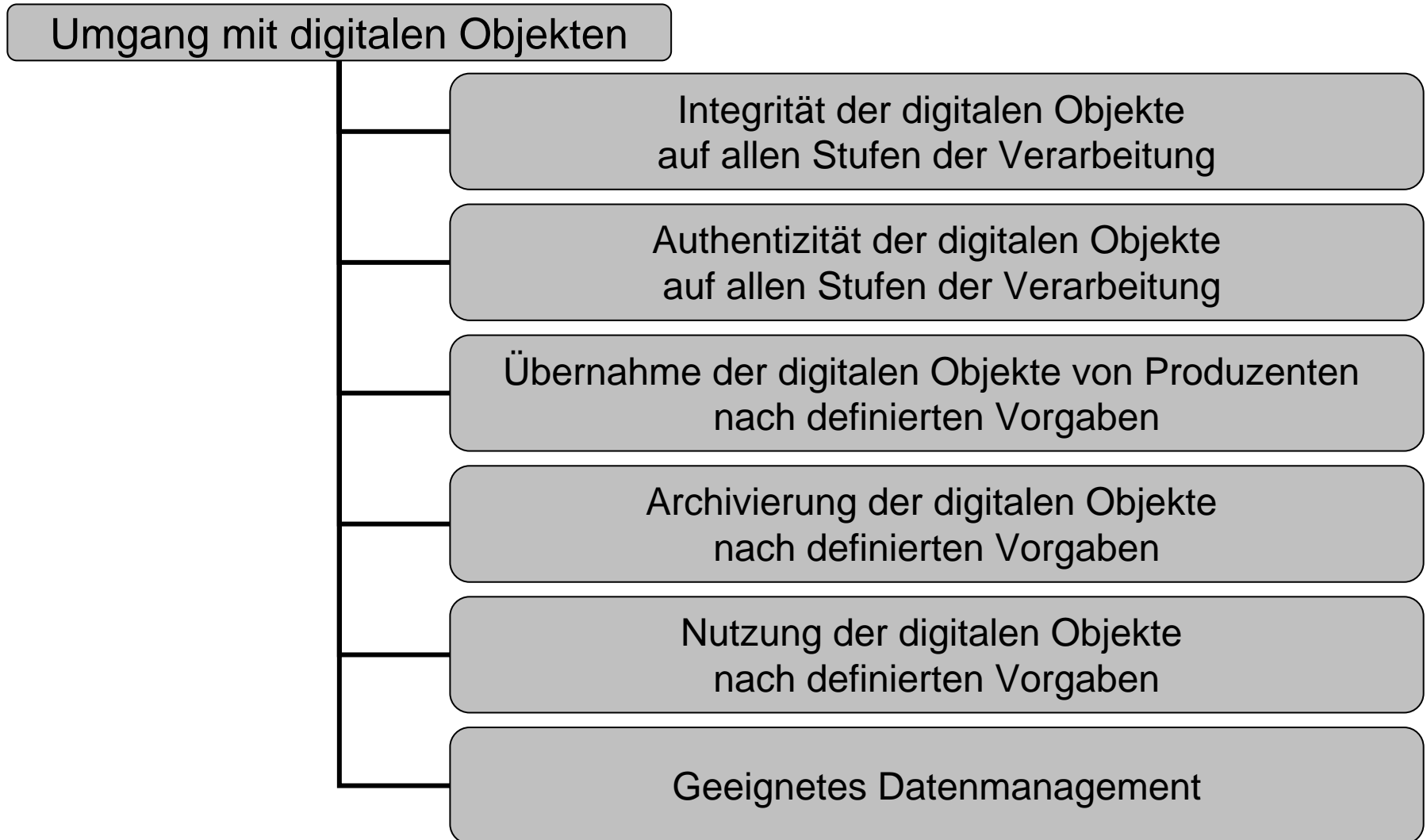
Umgang mit digitalen Objekten

Technische Infrastruktur einschl. Sicherheitsinfrastruktur

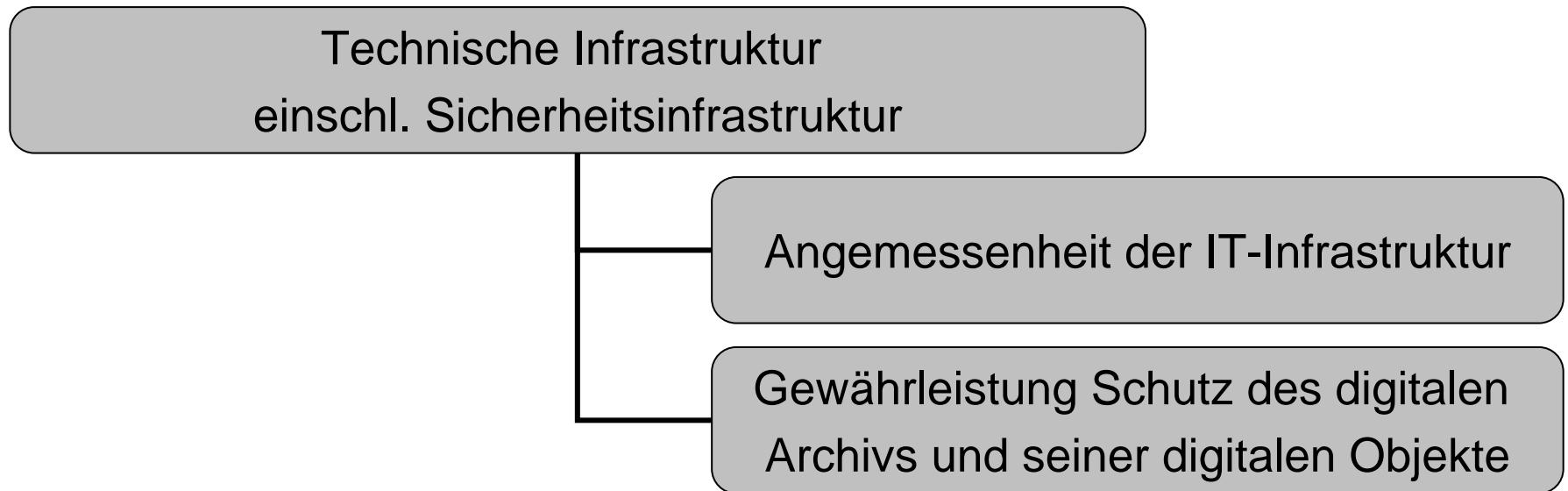
Indikatoren für Vertrauenswürdigkeit: Organisatorischer Rahmen



Indikatoren für Vertrauenswürdigkeit: Umgang mit digitalen Objekten



Indikatoren für Vertrauenswürdigkeit: Technische Infrastruktur



Indikator für Vertrauenswürdigkeit: Integrität

Was ist Integrität:

- Integrität (Vollständigkeit und Unversehrtheit) der digitalen Objekte (Bitsequenz)
Ist das die Problemstellung?
- Integrität der digitalen Information
eigentliches Ziel: Erhalt der Information

Zu klären:

- Was repräsentieren die Teile eines digitalen Objektes?

Dies ermöglicht:

- Auswahl geeigneter Formate
- Bestimmung erlaubter (akzeptabler) Migrationen / Alternativen
- Beseitigung von Redundanz / Irrelevanz
- Verminderung von Komplexität

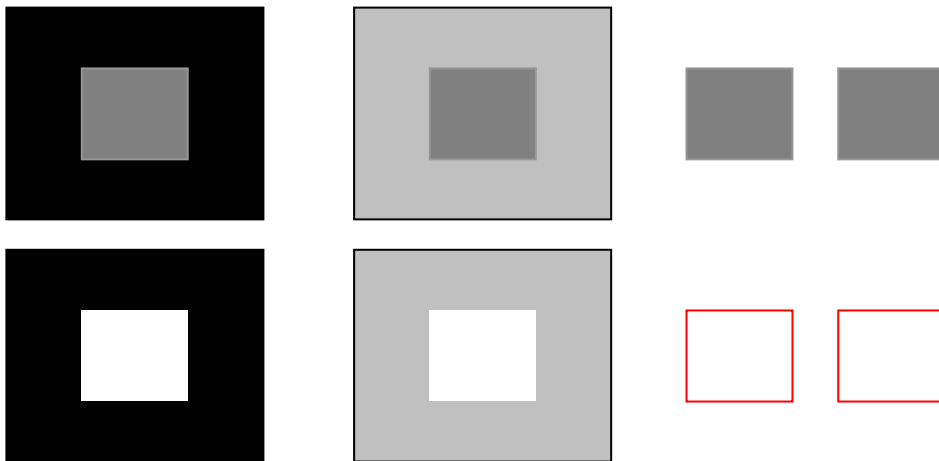
Zum Erhalt von Integrität

rn -> rn -> r n (eingeschobenes Leerzeichen)

1365,89 -> -1365,89

1365,89 -> 1365,89

1365,89 + 1365,89 = 0,00 -> 1365,89 + 1365,89 = 2731,78



Farbraumtransformationen:
Invarianz bei Hintergrundwechsel
Messung vs. Wahrnehmung!

Zum Erhalt von Integrität

Transformation einer Tabulatur (Verlust der Repräsentation von Zeit)

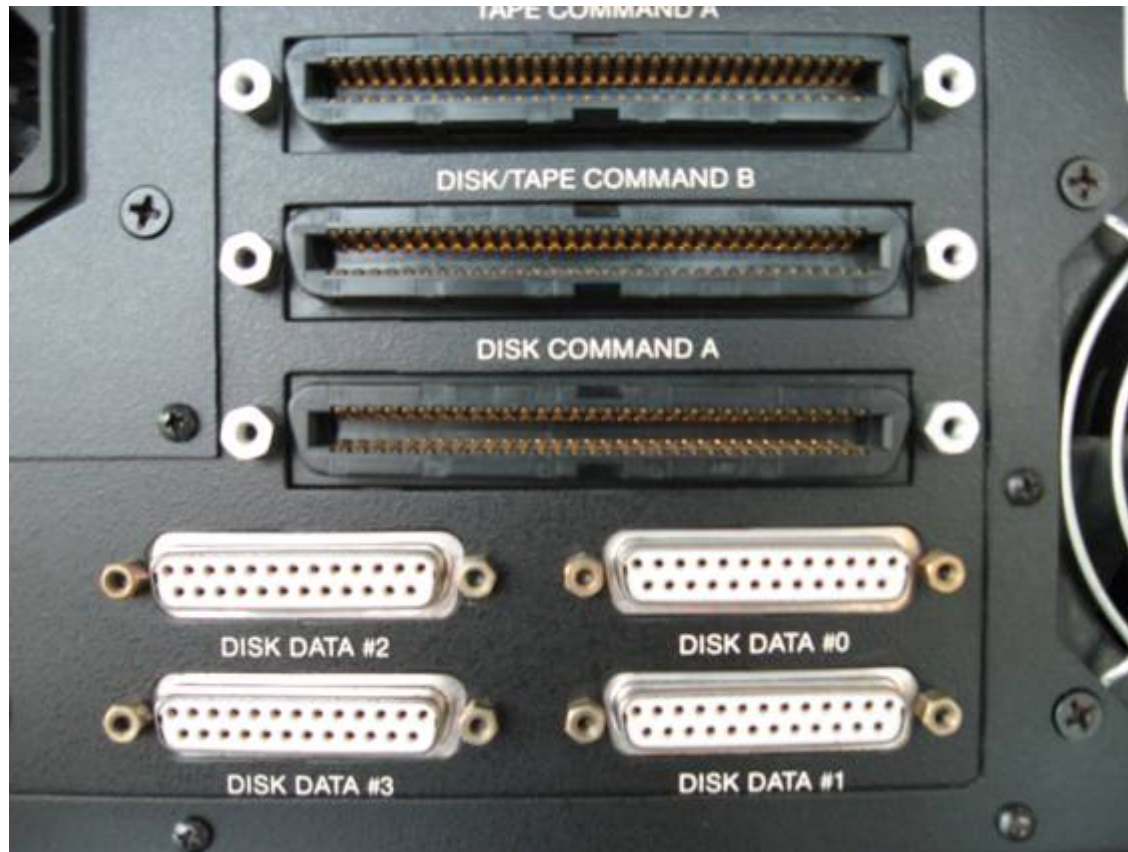
e|-----
h|-----
G|-----
D|-----5-----8-----6-----
A|-----7-----7-----2-----
E|--0-----6--5-----1-----

e|-----
h|-----
G|-----
D|-----5-----8-----6-----
A|-----7-----7-----2-----
E|--0-----6--5-----1-----

Zur Sicherstellung der Integrität in der Praxis

- Gründliche Analyse von Anforderungen an den Informationserhalt im jeweiligen Archivierungskontext
(Allgemeine Ratschläge sind mit Vorsicht zu genießen)
- Rückgriff auf formale Modelle und darauf aufbauende Werkzeuge
- Ausführliche Tests von Methoden und Werkzeugen an realen Beständen (zusätzlich Aussagen zu Beständen und Aufwand)
- Dokumentation aller Erkenntnisse
- Austausch der Erkenntnisse (z.B. nestor als Plattform)

Vielen Dank für Ihre
Aufmerksamkeit !



Anlage

mehr ...

- nestor-Arbeitsgruppe Vertrauenswürdige Archive - Zertifizierung
Kriterienkatalog vertrauenswürdige digitale Langzeitarchive
<http://nbn-resolving.de/urn:nbn:de:0008-2006060710>
- Center for Research Libraries (CRL)
Assessing and Certifying Repositories
<http://www.crl.edu/content.asp?l1=13&l2=58&l3=162>