

Was ist digitale Langzeitarchivierung und warum braucht man sie? Welche technischen und organisatorischen Voraussetzungen für digitale Langzeitarchivierung gibt es?

Pia Rudnik

ZB MED – Informationszentrum Lebenswissenschaften

Leibniz Info-Tag zur digitalen Langzeitarchivierung, Berlin, 30.10.2018



# Was wird angesprochen?

- Begriffsklärung: Was bedeutet LZA?
- Herausforderungen bei der LZA digitaler Ressourcen
- Vorstellung verschiedener Ebenen der LZA
- Vorstellung des OAIS-Referenzmodells
- Organisatorische Voraussetzungen für die LZA
- Technische Voraussetzungen für die LZA



# Früher...

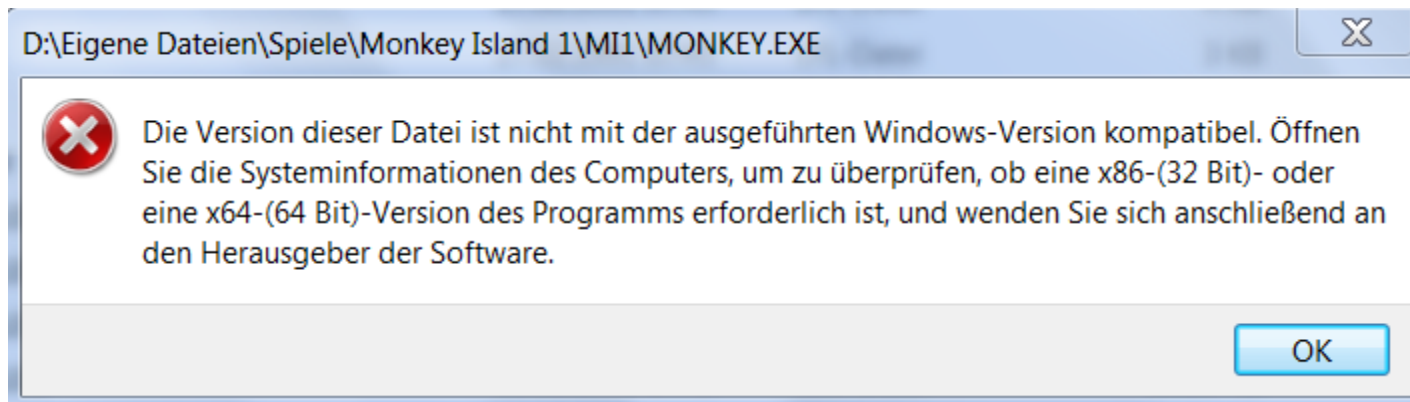


Quelle: <https://pixabay.com/de/schwarz-gesch%C3%A4ft-computer-18320/>, CC0



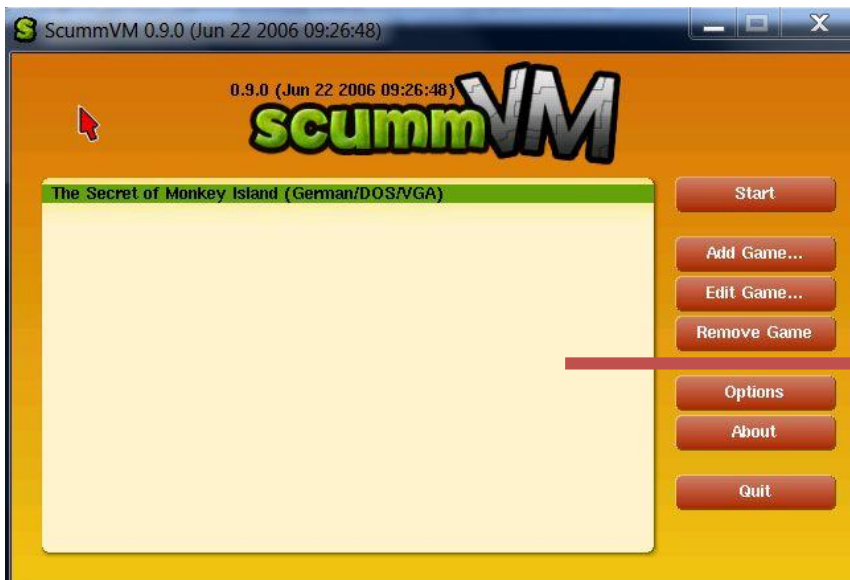
Quelle: Screenshot P. Rudnik, „Monkey Island 1 – The Secret of Monkey Island“, LucasArts 1990

# Heute...



Quelle: Screenshot P. Rudnik

Oder...



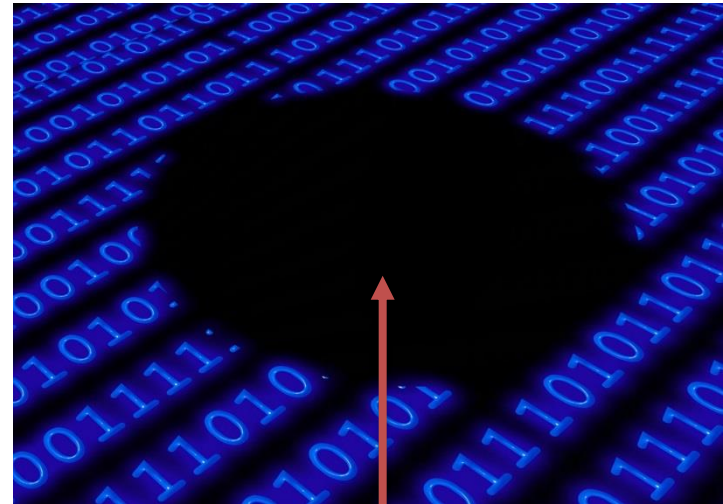
Quelle: Screenshot P. Rudnik, scummVM, scummVM-Team



Quelle: Screenshot P. Rudnik, „Monkey Island 1 – The Secret of Monkey Island“, LucasArts 1990

# Ausgangslage: Informationsflut und drohender Informationsverlust in der digitalen Welt

- Informationen liegen heutzutage zunehmend ausschließlich digital vor
- Große Datenmengen
- Datenheterogenität (unterschiedliche Formate, unterschiedlich strukturiert)
- Flüchtigkeit von Hard- und Software
- Unterschiedliche und sich verändernde Anforderungen an Daten



Quelle: <https://pixabay.com/de/daten-datenverlust-fehlende-daten-2764820/>, CC0

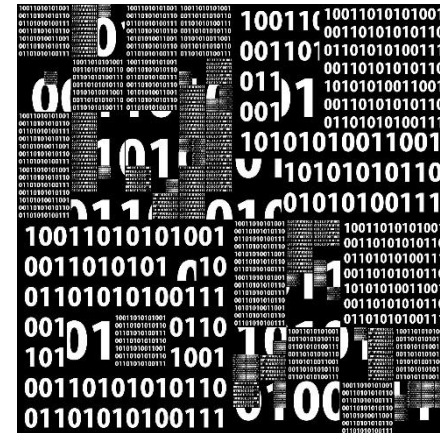
**Risiko: Datenverlust bedeutet häufig unwiederbringlichen Informationsverlust und damit den Verlust des kulturellen und wissenschaftlichen digitalen Erbes!**

# Herausforderungen und Risiken (I)

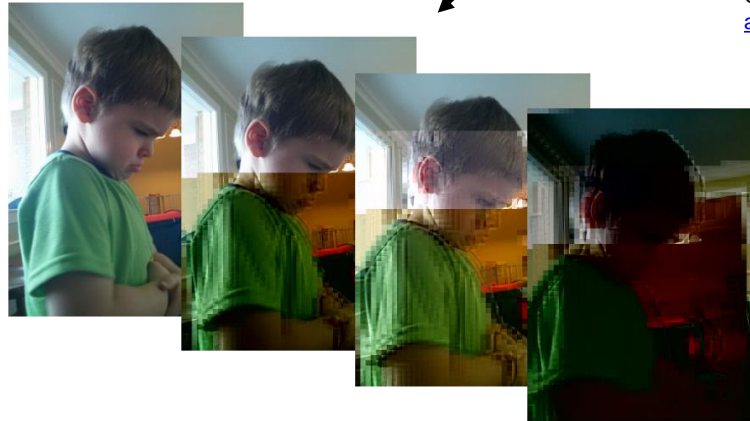
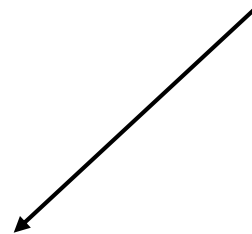
## Veraltete oder beschädigte Speichermedien



Quelle: <https://pixabay.com/de/speicher-daten-diskette-cd-dvd-769053/> , CC0



Quelle: <https://pixabay.com/de/http-www-absturz-administrator-at-216835/> , CC0

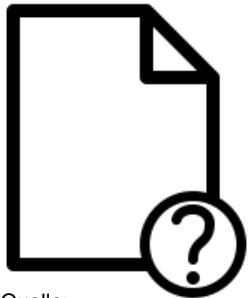


Quelle: Jim Salter: [https://commons.wikimedia.org/wiki/File:Bitrot\\_cascade.png](https://commons.wikimedia.org/wiki/File:Bitrot_cascade.png) , CC BY-SA



# Herausforderungen und Risiken (II)

## Obsolete Dateiformate, veraltete Soft- und Hardware



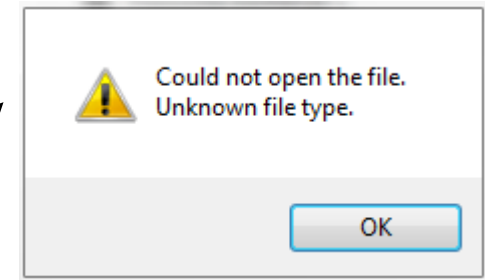
Quelle: <https://thenounproject.com/term/unknown-file/532168/>, iconsphere, CC BY



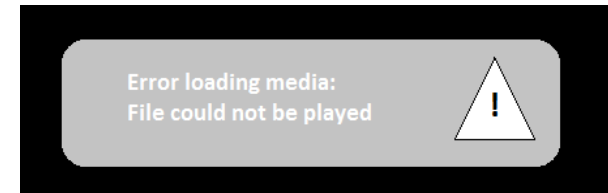
Quelle: <https://pixabay.com/de/computer-pc-rechner-hardware-277775/>, CC0



Quelle: <https://pixabay.com/de/bin%C3%A4r-bin%C3%A4rcode-fragezeichen-3725327/>, CC0



Quelle: Screenshot P. Rudnik



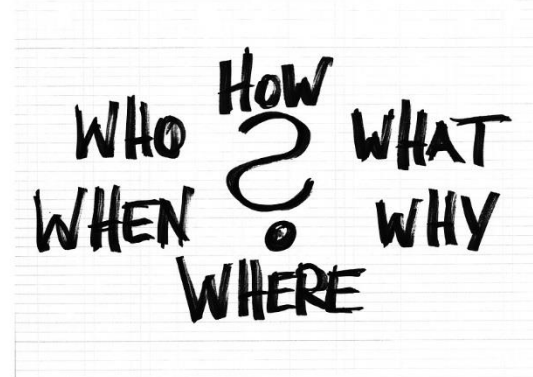
Quelle: P. Rudnik, CC BY SA

# Herausforderungen und Risiken (III)

## Unzureichende Kontextinformationen



Quelle: <https://pixabay.com/de/telefon-alt-holz-w%C3%A4hlscheibe-2663655/>, CC0



Quelle: <https://pixabay.com/de/fragen-wer-was-wie-warum-wo-1328351/>, CC0



Quelle: <https://pixabay.com/de/verwirrt-unsicherh%C3%B6hung-h%C3%A4nde-3082831/>, CC0

# Herausforderungen und Risiken (IV)

## Unterschiedliche Anforderungen der Zielgruppen

Wie können digitale Objekte erhalten werden, damit sie langfristig den Anforderungen (welche sich ggf. im Laufe der Zeit verändern) ihrer Zielgruppen gerecht werden?



Quelle:  
<https://thenounproject.com/search/?q=pdf%20file&i=749513>,  
 Yamini Ahluwalia, GB, CC BY

**Evidenzbasierte Medizin – die Basis einer guten Gesundheitspolitik**

**Zusammenfassung**

Für die neue Legislaturperiode hat die Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF) Forderungen zur Gesundheitspolitik aufgestellt.

**Dennis Makoschey\***

1. AWMF, Berlin

**Text**

In Mittelpunkt einer patientenorientierten Gesundheitspolitik steht die Verbesserung der Qualität der medizinischen Versorgung für alle Menschen in Deutschland. Die Gesundheitspolitik muss sich an empirisch belegten Fakten orientieren, denn nur eine evidenzbasierte Medizin sichert die Qualität der Versorgung und nur eine evidenzbasierte Gesundheitspolitik kann dafür die Rahmenbedingungen setzen. Eine gute medizinische Versorgung setzt eine qualitativ hochwertige Aus-, Weiter- und Fortbildung sowie medizinische Forschung auf international kompetitivem Niveau voraus. Für ein zukunftsorientiertes, nachhaltiges und effektives Gesundheitssystem für Deutschland fordert daher die AWMF:

1. Intensive Kooperation der Gesundheitspolitik mit der wissenschaftlichen Medizin
2. Unabhängige Finanzierung der Entwicklung und Implementierung von Leitlinien
3. Koordination des Infektionsschutzes zwischen Humanmedizin und Tiermedizin
4. Verbesserung der Rahmenbedingungen für die Gesundheitsforschung
5. Personalentwicklung und Nachwuchsförderung in der Medizin

Für die Gesundheitsgesetzgebung sollten ähnlich hohe Maßstäbe gelten wie für die medizinische Versorgung. Bevor neue Wege zur Verbesserung der medizinischen Versorgung eingeschlagen werden, ist anhand überprüfbarer Fakten nachzuweisen, dass die gesetzgebenden Maßnahmen ausreichend, notwendig und zweckmäßig sind. Die AWMF bietet hierzu gern ihre Unterstützung an.

**1. Intensive Kooperation der Gesundheitspolitik mit der wissenschaftlichen Medizin**

Bei gesundheitspolitischen Entscheidungen gilt es darauf zu achten, dass hierbei die vorhandene fachliche Kompetenz angemessen berücksichtigt wird (Bund, Länder, Selbstverwaltung, Medizinische Wissenschaft). Dabei sollten die AWMF und die in ihr organisierten wissenschaftlichen Medizinischen Fachgesellschaften häufiger und zu einem

GMS | AWMF  
 GMS Mitteilungen von der AWMF 2018, Nr. 15, ISSN 1860-4269 L15

Quelle: Makoschey, Dennis, German Medical Science,  
<https://www.egms.de/static/pdf/journals/awmf/2018-15/awmf000320.pdf>;  
<https://www.egms.de/static/xml/journals/awmf/2018-15/awmf000320.xml>, CC BY

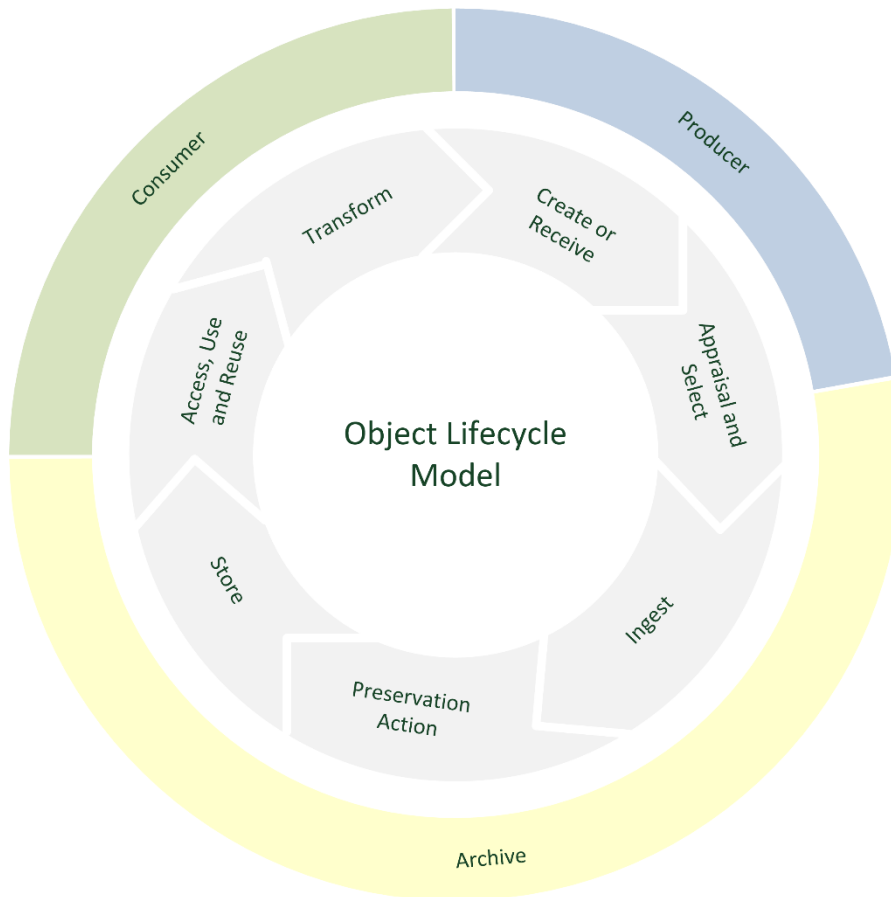
```
<?xml version="1.0" encoding="UTF-8" standalone="no" ?>
<GmsArticle xmlns:xlink="http://www.w3.org/1999/xlink">
  <Metadata>
    <Identifier>awmf000320</Identifier>
    <IdentifierDoi>10.2305/awmf000320</IdentifierDoi>
    <IdentifierUrn>urn:nbn:de:0183-awmf000320</IdentifierUrn>
    <ArticleType>Mittelteil</ArticleType>
  </Metadata>
  <Title language="de">
    Evidenzbasierte Medizin – die Basis einer guten Gesundheitspolitik
  </Title>
  </TitleGroup>
  <CreatorList>
    <Creator>
      <PersonNames>
        <LastName>Makoschey</LastName>
        <LastNameHeading>Makoschey</LastNameHeading>
        <FirstName>Dennis</FirstName>
        <Initials>D</Initials>
      </PersonNames>
      <Address>
        AWMF-Geschäftsstelle, Birnenstraße 67, 10559 Berlin
        <Affiliation>AWMF, Berlin</Affiliation>
        <WebPage>http://awmf.org/webPage</WebPage>
        <Address>
          <Email>office@awmf.org</Email>
          <CreatorRole corresponding="yes" presenting="no">author</CreatorRole>
        </Creator>
      </CreatorList>
      <PublisherList>
        <Publisher>
          <Corporation>
            <CorporationName>German Medical Science</CorporationName>
            <Corporation>
              <Address>Berlin</Address>
            </Corporation>
          </Publisher>
        </PublisherList>
      </SubjectGroup>
      <SubjectHeadingDOI>610</SubjectHeadingDOI>
      </SubjectGroup>
      <DateReceived>20180207</DateReceived>
      <DatePublishedList>
        <DatePublished>20180209</DatePublished>
      </DatePublishedList>
      <Language>germ</Language>
      <License license-type="open-access" xlink:href="http://creativecommons.org/licenses/by/4.0/>
        <AltText language="en">
          This is an Open Access article distributed under the terms of the Creative Commons Attribution License.
        </AltText>
        <AltText language="de">
          Dieser Artikel ist ein Open-Access-Artikel und steht unter den Lizenzbedingungen der Creative Commons Attribution License.
        </AltText>
      </License>
    </SourceGroup>
    <Journal>
      <ISSN>1860-4269</ISSN>
      <Volume>15</Volume>
      <JournalTitle>GMS Mitteilungen von der AWMF</JournalTitle>
    </Journal>
  </GmsArticle>
```



Quelle:  
<https://thenounproject.com/search/?q=xml%20file&i=1658765>,  
 David, US, CC BY



# Lebenszyklus von Daten



Daten langfristig verfügbar zu halten, ist eine **andauernde** Aufgabe, die niemals abgeschlossen ist.

LZA bedeutet also eine stetige Anpassung von Daten im Hinblick auf:

- technische Veränderungen
- Entwicklung neuer Zugriffsformen
- sich ändernde Anforderungen der Nutzenden

Quelle: Michelle Lindlar, CC BY SA

# Langzeitarchivierung

## Was bedeutet das – was nicht?

### Digitales Langzeitarchiv gemäß DIN 31644:

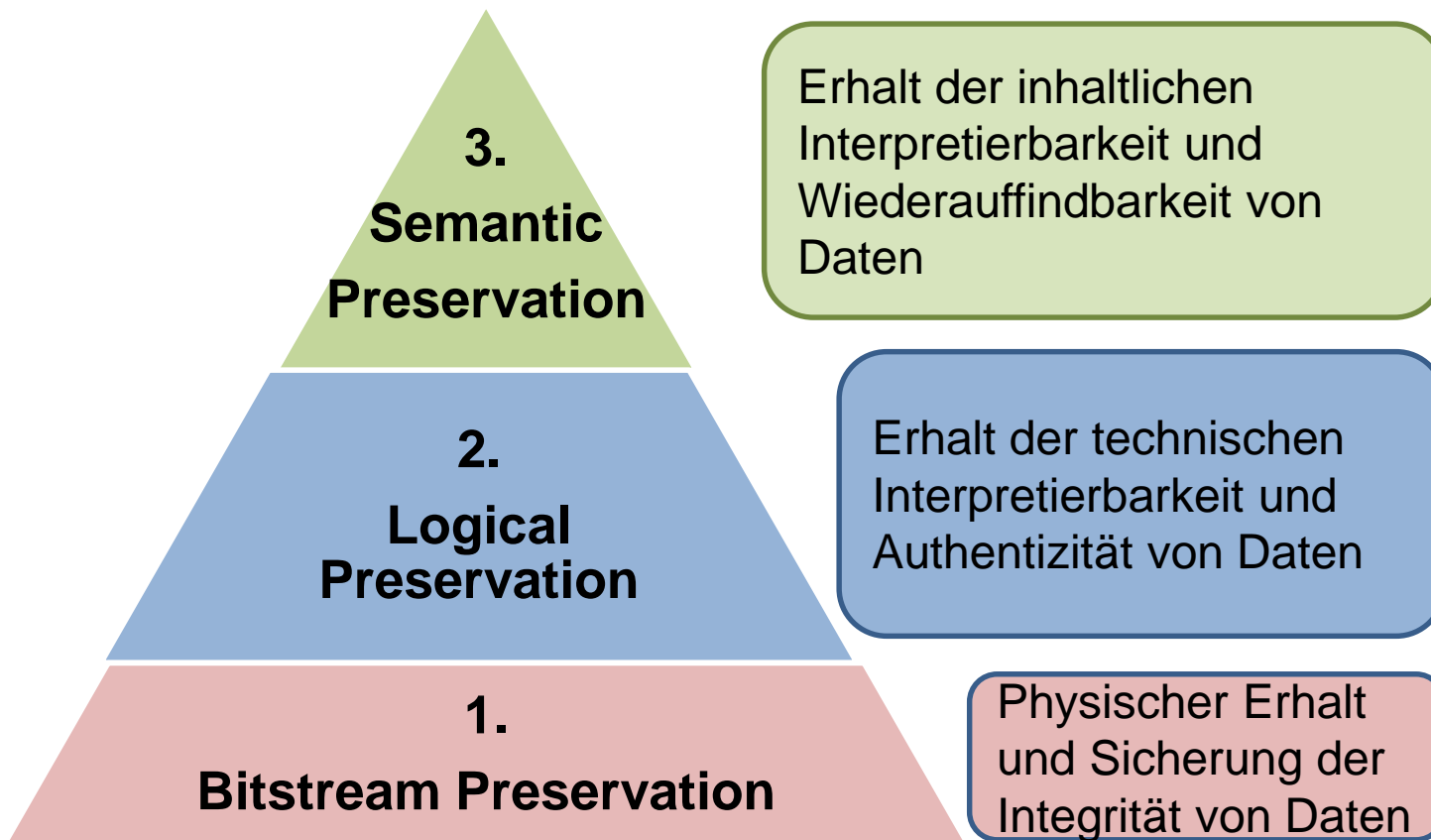
„Organisation (bestehend aus **Personen und technischen Systemen**), die die Verantwortung für den **Langzeiterhalt und die Langzeitverfügbarkeit** von **Information in digitaler Form** sowie die Bereitstellung für eine bestimmte **Zielgruppe** übernommen hat.“

#### → LZA ist:

- keine rein technische Angelegenheit
- zeitlich nicht beschränkt
- Informationserhaltung, nicht nur Datenerhaltung
- abhängig von den Anforderungen der Zielgruppen
- ein niemals abgeschlossener Prozess

**Ziel: Erhalt digitaler Information über die Veränderung von Soft- und Hardware hinaus**

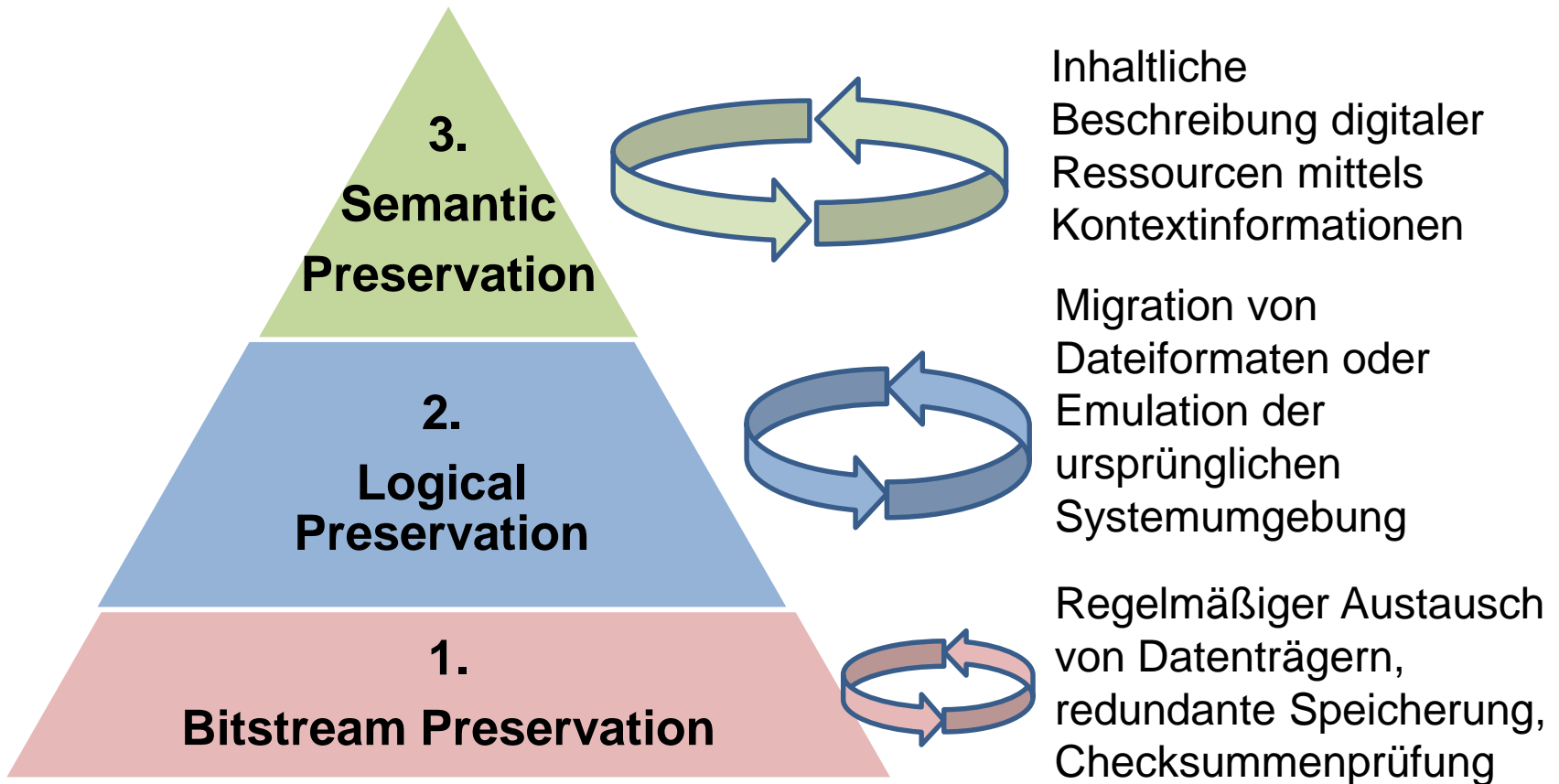
# Die drei Ebenen der Langzeitarchivierung (I)



Quelle: P. Rudnik, CC BY SA



# Die drei Ebenen der Langzeitarchivierung (II)



Quelle: P. Rudnik, CC BY SA

# Open Archival Information System (OAIS)

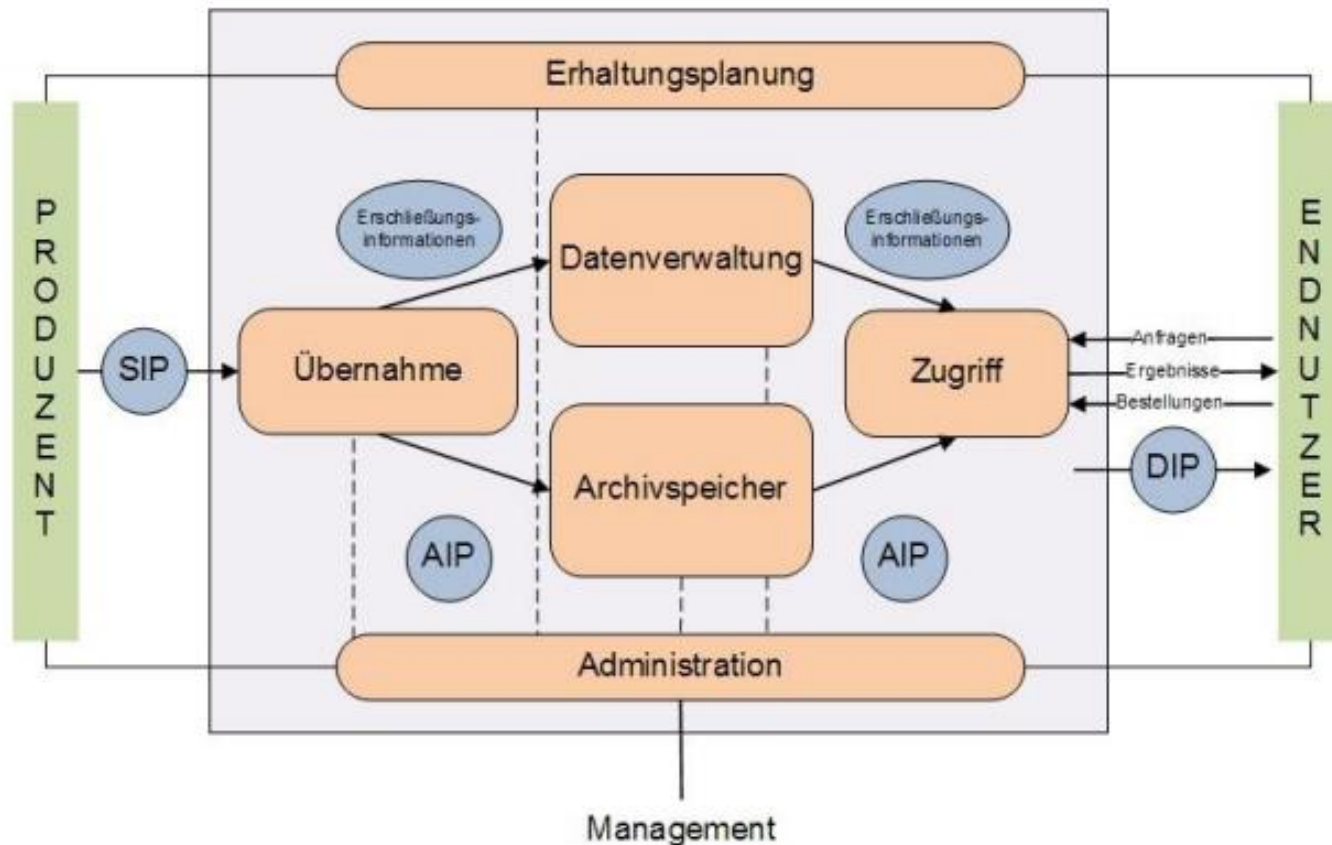
## Referenzmodell für die LZA (I)

- Generisches Modell, entwickelt 1997 unter Federführung der NASA
- Seit 2003 ISO-Standard (ISO 14721)
- Als Kommunikationsgrundlage in LZA-Community anerkannt
- Modell beschreibt offenes, modular aufgebautes LZA, in dem Menschen und Systeme zusammen verschiedene Aufgabe erfüllen
- Keine Beschränkung auf bestimmte Institutionstypen, Formate oder Systemarchitekturen → OAIS nennt keine konkreten Lösungen!
- OAIS-Funktionseinheiten *Erhaltungsplanung, Administration, Übernahme, Datenverwaltung, Archivspeicher* und *Zugriff* beschreiben organisatorische und technische Aufgabenbereiche eines LZA von der Übernahme von DatenproduzentInnen bis hin zur Übergabe an die jeweiligen EndnutzerInnen
- Nennt drei Arten von Informationspaketen, die in einem LZA weiterverarbeitet werden (SIP, AIP, DIP)



# Open Archival Information System (OAIS)

## Referenzmodell für die LZA (II)



Quelle: nestor - Kompetenznetzwerk Langzeitarchivierung: Referenzmodell für ein Offenes Archiv-Informationssystem – Deutsche Übersetzung 2.0, nestor-Materialien 16, <https://d-nb.info/104761314X/34>

# Organisatorische Voraussetzungen für die LZA

- Verantwortung für den Erhalt digitaler Informationen übernehmen
- Langfristig Ressourcen (Personal, Finanzen für technische Infrastruktur oder für entsprechende Dienstleistung) für die LZA zur Verfügung stellen
- Verbindliche Regelung der Verantwortlichkeiten innerhalb des LZA, z.B. anhand von Policies
- Klärung rechtlicher Rahmenbedingungen
- Bestimmung der Zielgruppen und Abstimmung der LZA auf die Anforderungen der Zielgruppen (Community Watch)
- Dokumentation des LZA und der Prozesse, damit das LZA und seine digitalen Ressourcen langfristig sowohl für die Betreiber als auch für die EndnutzerInnen nachvollziehbar bleibt

# Technische Voraussetzungen für die LZA

- Skalierbare technische Infrastruktur: Archivspeicher, Datenverwaltung, Schnittstellen zu Repositories, Tools zur Durchführung von Formaterkennung, Formatvalidierung, Migration oder Emulation
- Monitoring von Hard- und Software (Technology Watch), um frühzeitig auf obsoletere Hard- und Software und Formate reagieren zu können
- Sicherheitskonzept: Redundante Speicherung, Tools zur Sicherstellung der Datenintegrität, regelmäßiger Austausch von Speichermedien
- Automatisierte Workflows (insbesondere bei großen Datenmengen), für Übernahme, Weiterverarbeitung und Ausgabe der Informationsobjekte und für das Rechte-Management
- Implementierung von Metadatenstandards (z.B. PREMIS)
- Planung für die Überführung der Informationsobjekte in eine andere technische Infrastruktur bei einem Exit-Szenario

# Fazit

- LZA hat viele Namen
- LZA findet auf verschiedenen Ebenen statt
- LZA sichert unser digitales Wissen
- LZA ist ein niemals endender Prozess

**Herzlichen Dank für Ihr Interesse!  
Haben Sie Fragen?**

Thomas Bähr, [Thomas.Baehr@tib.eu](mailto:Thomas.Baehr@tib.eu), Tel. 0511 / 762-17281

Pia Rudnik, [rudnik@zbmed.de](mailto:rudnik@zbmed.de), Tel. 0221 / 478-32176

Dr. Monika Zarnitz, [m.zarnitz@zbw.eu](mailto:m.zarnitz@zbw.eu), Tel. 0431 / 8814 - 431

