

Digitale Langzeitverfügbarkeit für Hochschulen in NRW

Kolloquium UB Bielefeld, 16.06.2023

Dr. Philip Gross | hbz

Paul Kirschner | hbz

LZV und LZA in NRW

- [LZA.nrw](#)
 - Projekt der Universitätsarchive (Federführend UDE)
 - Finanziert durch Kanzler:innenkonferenz
 - Verwaltungsdokumente und e-Akte
- [DA NRW](#)
 - Land, Kommunen, Landesbibliotheken
 - Kunst- und Kulturgüter
- [Datensicherung.nrw](#)
 - RWTH Aachen
 - Bitstream-Preservation für Forschungsdaten
- [LZV.nrw](#)
 - DH.NRW Projekt aus hbz, WWU und UzK
 - LZV für Hochschulen
 - Wissenschaftliche Publikationen und Forschungsdaten



Landesinitiative Langzeitverfügbarkeit

LZV Pilotphase

- RWTH Aachen
- RRZK Köln
- USB Köln

Vorprojekt-Phase

- Bedarfsermittlung
- Konzeptionierung
- Vernetzung



März, 2017

Januar, 2019

Januar, 2020

September, 2021

Konzept Workshop

- Landeslösung
- Kooperativer Ansatz

Landesinitiative LZV.nrw

- Start der Initiative
- Konsortiumsbildung



Landesinitiative Langzeitverfügbarkeit

Arbeitspakete

LZV-Services



- LZV.nrw Web-Präsenz
- LZV-Servicekatalog
- Service-Entwicklung

Kooperation



- Wissensmanagement
- DH.NRW Schnittstellen
- Governance

Entwicklung



- Schnittstellenentwicklung
- Digital Curation Management
- Cloud-Services

Zielsetzung

- 🎯 Awareness
- 🎯 Konkrete Services für Publikationen und FD für NRW Hochschulen
- 🎯 Kooperativer ganzheitlicher Ansatz: Beratung, Services, Wissensaustausch



Serviceangebot

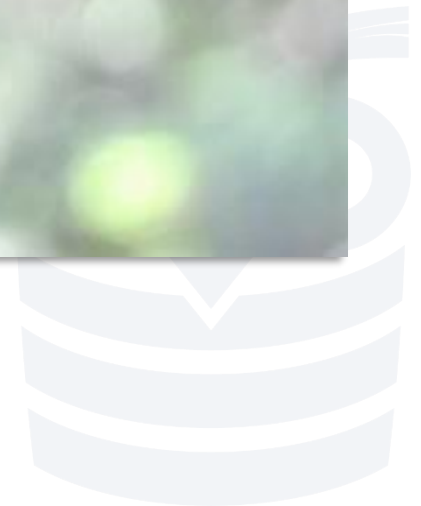


- Beratung und Begleitung
 - LZV Strategieentwicklung
 - Onboarding
- LZV-Infrastruktur und Services
 - Betrieb und Speicherung
 - Services für Publikationen
 - Services für Forschungsdaten
- Beratung und Support
 - Umsetzung der LZV an der Hochschule
 - Nutzung der Infrastruktur
- Technische Weiterentwicklung
 - Schnittstellenentwicklung
 - Tool-Entwicklung



Kooperation

- Wissensmanagement
 - NRW-Wissenscommunity
 - Wissensplattformen
 - Informationsangebote
- DH.NRW-Schnittstellen
 - [fdm.nrw](#)
 - [openaccess.nrw](#)
 - [Digi-Kunst.nrw](#)
- Governance
 - LZV-Beirat
 - LZV-Expert*innengruppe



Informationsangebote

- LZV.nrw Website www.lzv.nrw
- Service Wiki [Langzeitverfügbarkeit für Hochschulen in NRW](#)
- Hochschulbibliothekszentrum des Landes NRW [hbz](#)
- Digitale Hochschule NRW [DH.NRW](#)

Kontakt

Dr. Philip Gross

Leitung Landesinitiative LZV.nrw

philip.gross@hbz-nrw.de

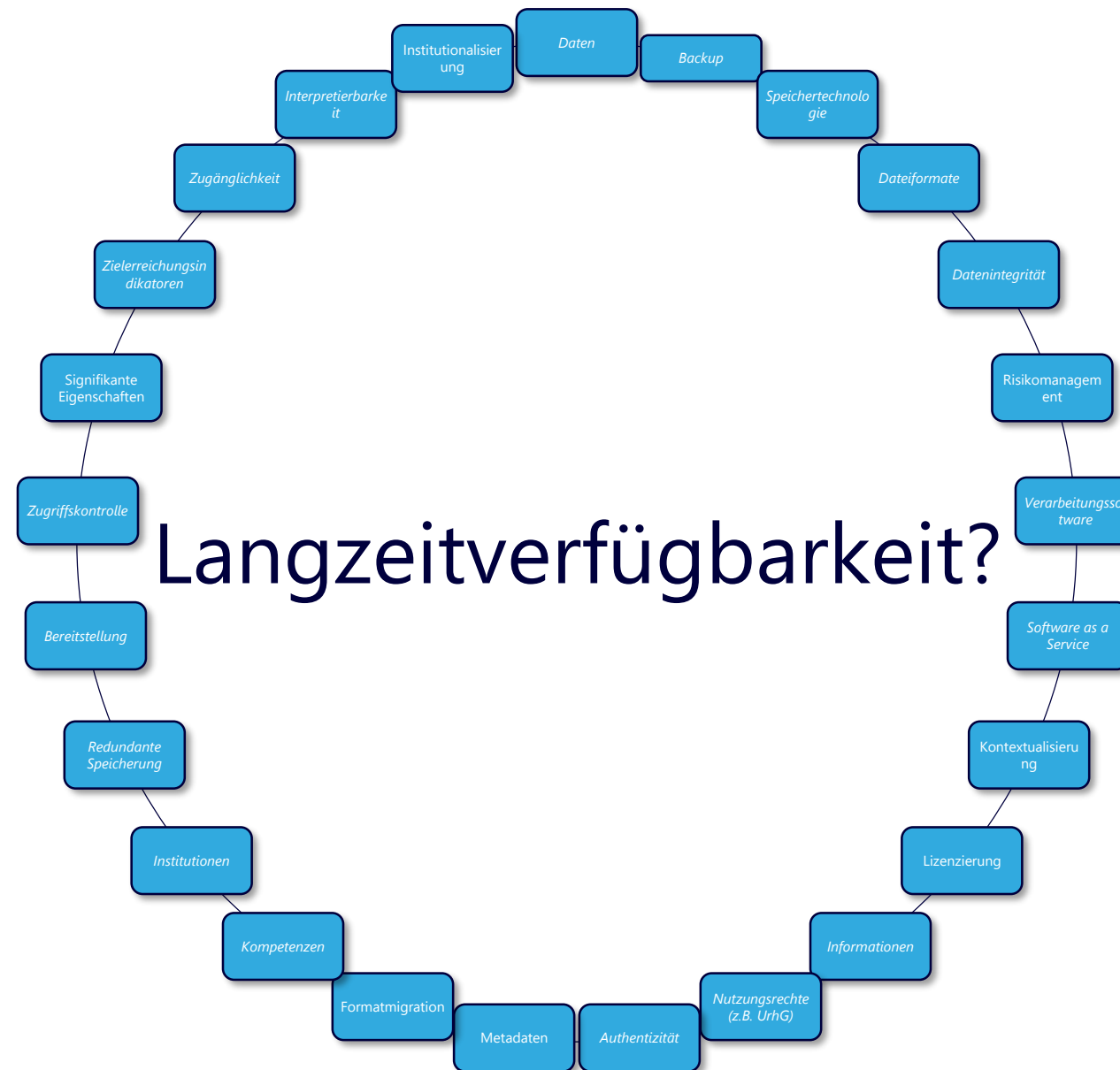
+49 221 – 40075 245



Langzeitverfügbarkeit

Technische Umsetzung mit Rosetta

Paul Kirschner, Koordinator LZV-Betrieb am hbz



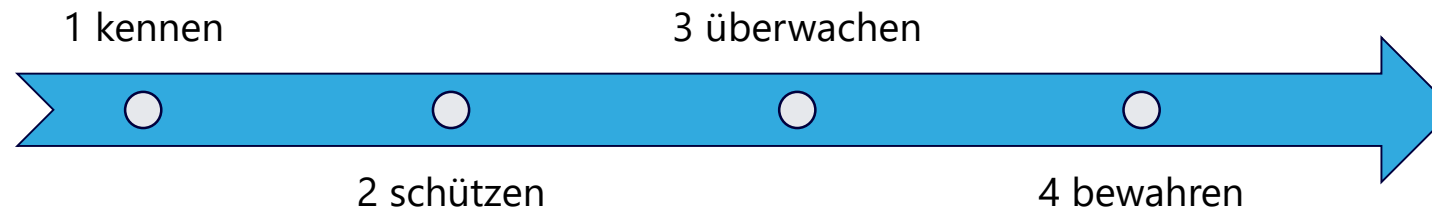
Erhaltungsdimensionen

Organisatorisch

Intellektuell

Technisch

Levels of Preservation



Level 4 - Bewahren

Speicher	Integrität	(Zugriffs)Kontrolle	Metadaten	Inhalt (essentielle Eigenschaften)
<ul style="list-style-type: none"> • Digitalen Inhalt in mindestens drei Kopien vorhalten und jede Kopie an einem geographisch getrennten Ort mit einem anderen Gefahrenpotenzial speichern. • Speicherdiversität durch die Nutzung verschiedener Speichertechnologien erhöhen, um ein mögliches Versagen einzelner Technologien ausgleichen zu können. • Es existiert eine Strategie, um dem Risiko der Obsoleszenz von Speicher-Hardware, -Software und -Medien zu begegnen. Entsprechende Maßnahmen durchführen. 	<ul style="list-style-type: none"> • Im Anschluss an bestimmte Ereignisse oder Maßnahmen Integritätsprüfungen durchführen. • Korrupten digitalen Inhalt je nach Anforderung ersetzen oder reparieren. 	<ul style="list-style-type: none"> • Regelmäßige Überprüfung der Maßnahmen-/Zugriffsprotokolle. 	<ul style="list-style-type: none"> • Protokollieren, welche Langzeitarchivierungsmaßnahmen am digitalen Inhalt durchgeführt worden sind und wann diese Maßnahmen stattgefunden haben. • Ausgewählte Metadatenstandards implementieren. 	<ul style="list-style-type: none"> • Dateiformatmigrationen, -Normalisierungen, Emulation und ähnliche Maßnahmen durchführen, welche die zukünftige Zugänglichkeit des digitalen Inhalts gewährleisten.

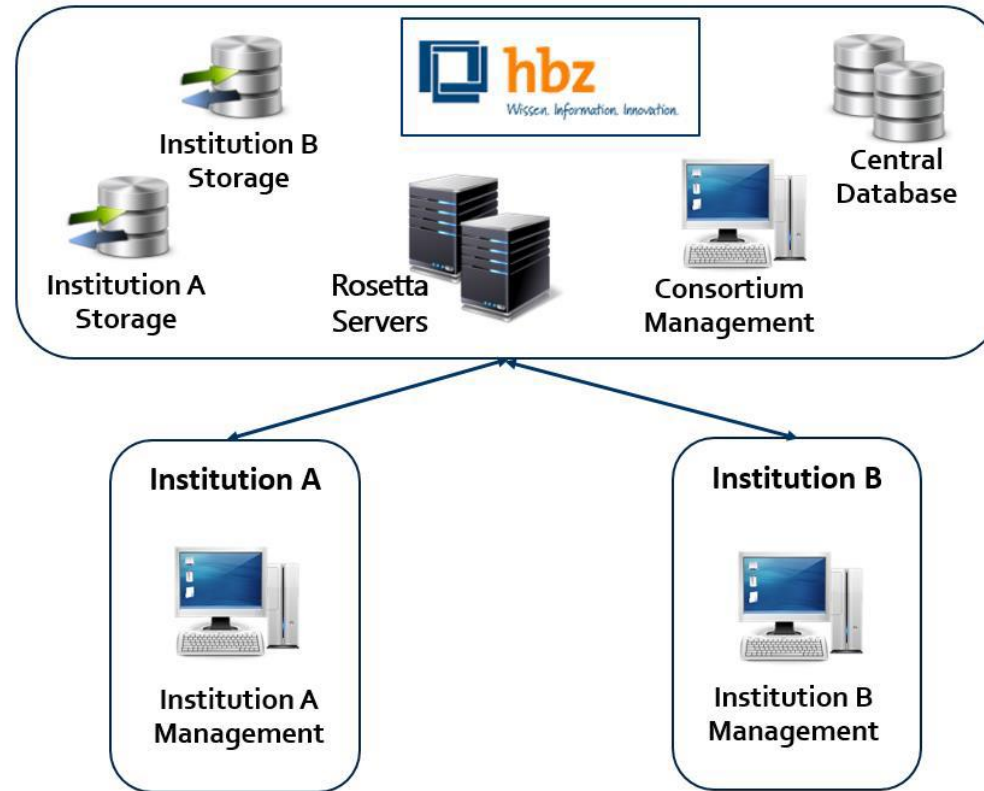
Quelle: Lindlar, Micky (Orcid: 0000-0003-3709-5608); Rudnik, Pia (Orcid: 0000-0003-4081-9646) (2021). NDSA 2019 Digital Preservation Levels Translation: German Translation of Version 2.0. <https://osf.io/2mkwx/> and <https://doi.org/10.5281/zenodo.4718661>. Original: <https://ndsa.org/publications/levels-of-digital-preservation/> / <https://web.archive.org/web/20221116015202/https://ndsa.org/publications/levels-of-digital-preservation/>

Merkmale der digitalen Langzeitverfügbarkeit

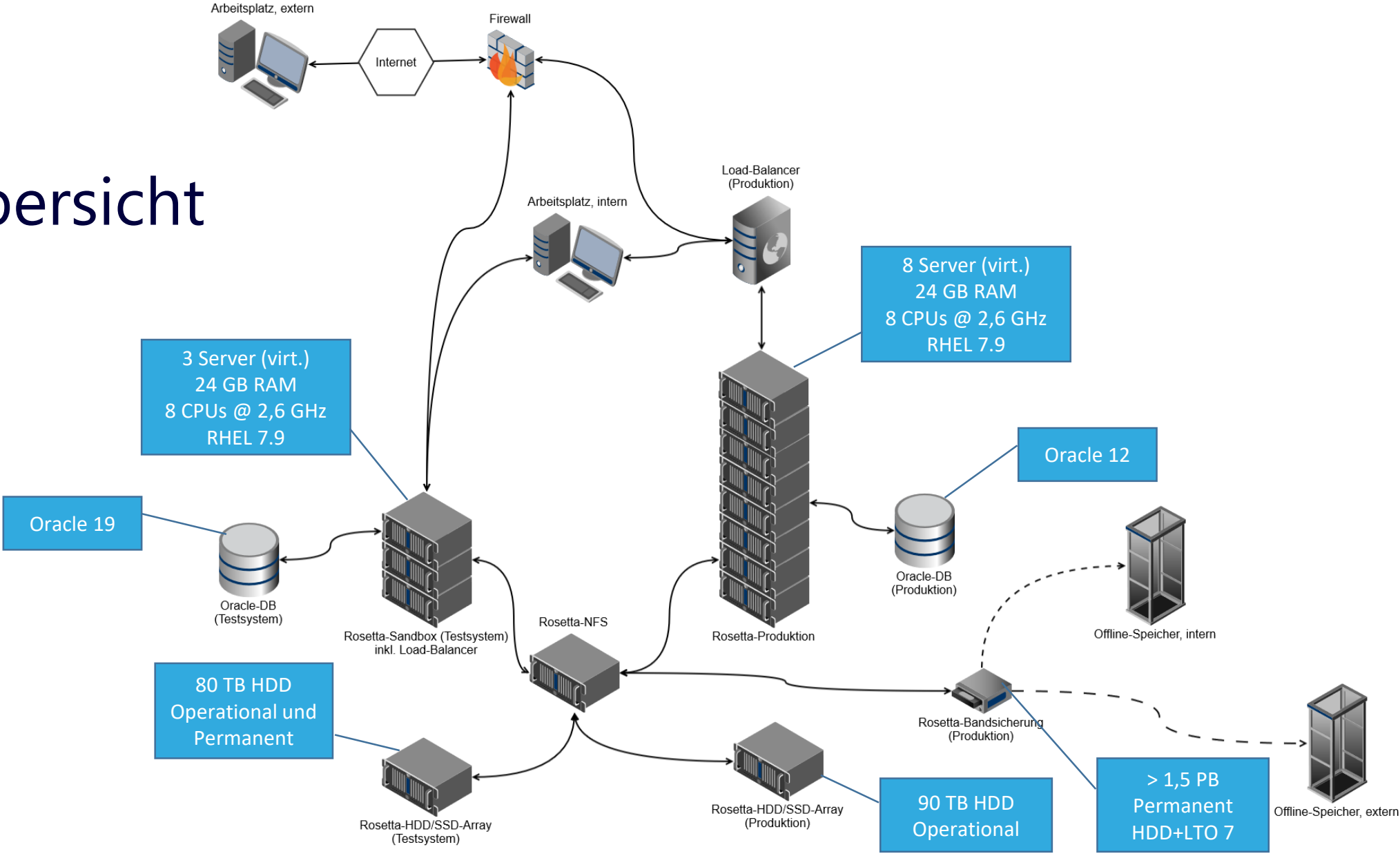
- Zweck: Sicherstellung der zukünftigen
 - Zugänglichkeit, d.h. Anzeige/Verarbeitung durch jeweils aktuelle Technologie
 - Nachnutzbarkeit, d.h. inhaltliche Verständlichkeit für Nutzer:innen
- Anforderungen
 - Bereitstellung hinreichender Ressourcen (organisatorisch, personell, finanziell, technologisch)
 - Entwicklung geeigneter Maßnahmen
 - Definition von Verantwortlichkeiten
 - Klärung rechtlicher Fragen
 - Sicherstellung von Datenintegrität und -authentizität
 - Festlegung institutioneller Regeln/Policies
 - Verstetigung und Organisationsentwicklung



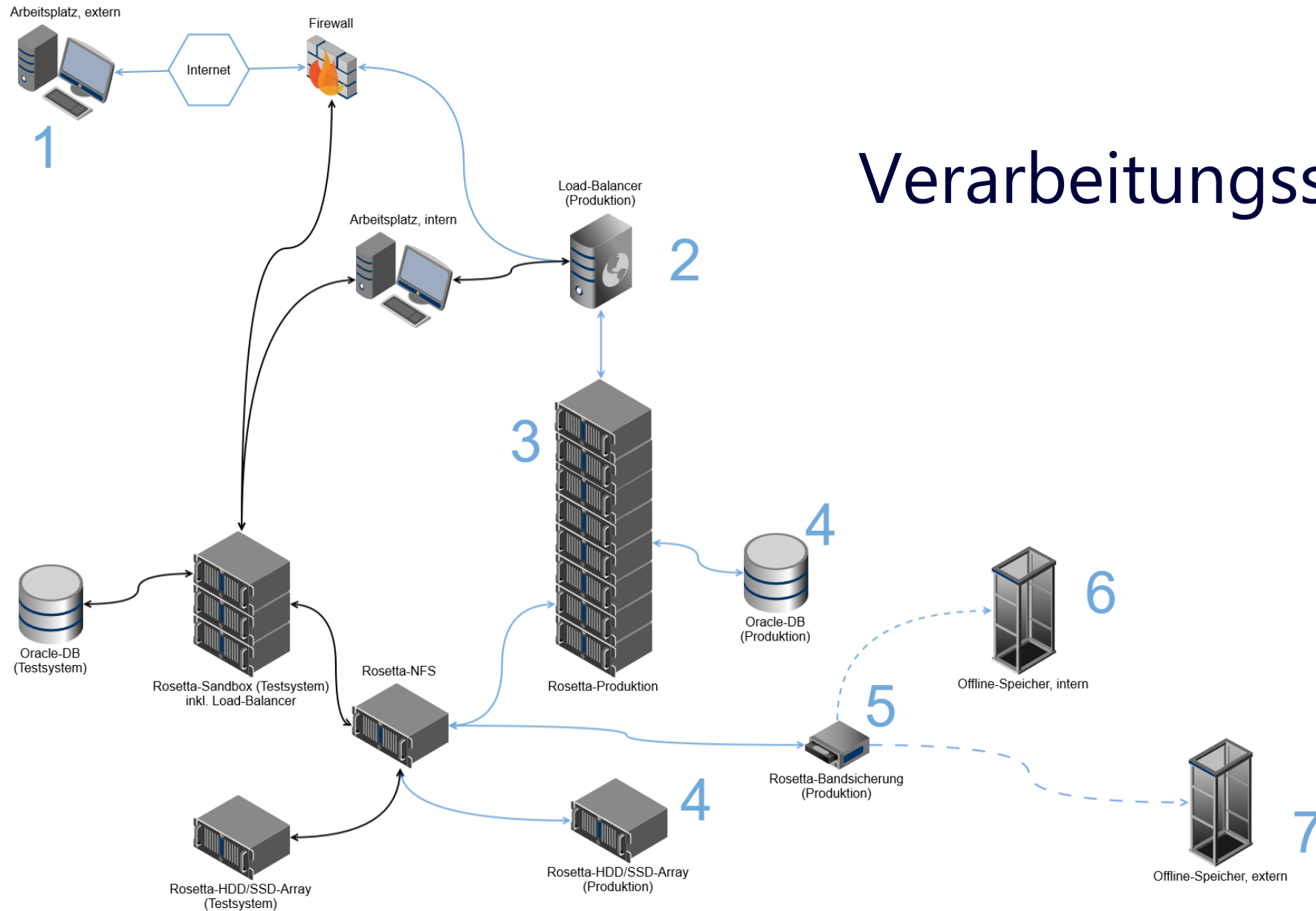
Infrastruktur – Rosetta im hbz



Systemübersicht



Verarbeitungsstationen



Interaktive Tafel gängiger Dateiformate

fmt/476 **.pdf** Acrobat
001

Für Details zum Dateiformat auf die PRONOM-ID klicken z.B. *fmt/476* (Weiterleitung auf *The National Archives UK*)

Bedeutung des Zahlencodes

LZV-Kriterien	Format	Lizenz	Nutzung			
Wert	bin	text	proprietär	frei	selten	häufig
Code	0	1	0	1	0	1

Filteroptionen

- Textformate
- Dokumentenformat
- Grafikformate
- Multimediaformate
- Daten- und 3D-Formate

Empfohlene Dateiformate zur Langzeitverfügbarkeit

Filter zurücksetzen

ASCII												XML													
x-fmt/111 .txt Editor 111												fmt/903 .blend Blender 011													
fmt/1149 .md Markdown 111												fmt/410 .arc WayBack 010													
x-fmt/276 .doc MS-Word 001	fmt/280 .tex LaTeX 111	fmt/160 .dvi LaTeX 010	fmt/126 .ppt PowerPoint 001	fmt/119 .bmp MS-Paint 001	fmt/4 .gif 001	fmt/92 .svg Inkscape 111	fmt/1 .wav 001	x-fmt/386 .mpg 001	fmt/569 .mkv vlc 011	fmt/506 .swf Flashplayer 001	x-fmt/18 .csv 111	fmt/244 .kml 111	fmt/1210 .obj 101	fmt/1281 .warc WayBack 101											
x-fmt/195 .sgml Docbook 111	fmt/412 .docx MS-Word 101	fmt/552 .indd InDesign 001	fmt/124 .eps Ghostscript 111	fmt/215 .pptx PowerPoint 101	fmt/730 .dng 001	fmt/645 .jpg 001	fmt/559 .ai Illustrator 001	fmt/279 .flac 011	fmt/640 .m2v kaffeine 001	fmt/199 .mp4 vlc 001	not listed .pez Prezi 001	fmt/57 .xls MS-Excel 001	fmt/1134 .gpx 001	fmt/1209 .dae collada 111	x-fmt/263 .zip 001										
fmt/471 .html Chrome 111	fmt/291 .odt LibreOffice 011	fmt/1091 .sla Scribus 110	fmt/476 .pdf Acrobat 001	fmt/293 .odp LibreOffice 011	fmt/13 .png 011	x-fmt/392 .jp2 001	fmt/297 .odp LibreOffice 011	fmt/134 .mp3 001	fmt/573 .webm chrome 011	x-fmt/384 .mov Quicktime 001	fmt/214 .xlsx MS-Excel 101	fmt/582 .xthreed 110	x-fmt/265 .tar 011												
fmt/103 .xhtml 110													x-fmt/387 .tif 001	fmt/427 .cdr CorelDraw 000	fmt/596 .m4a 001	fmt/5 .avi vlc 011	fmt/337 .mj2 001	fmt/295 .ods LibreOffice 011	fmt/831 .ply 001	fmt/287 .hdf5 C# 010					
												fmt/532 .dxf Autocad 101													fmt/203 .ogg 011



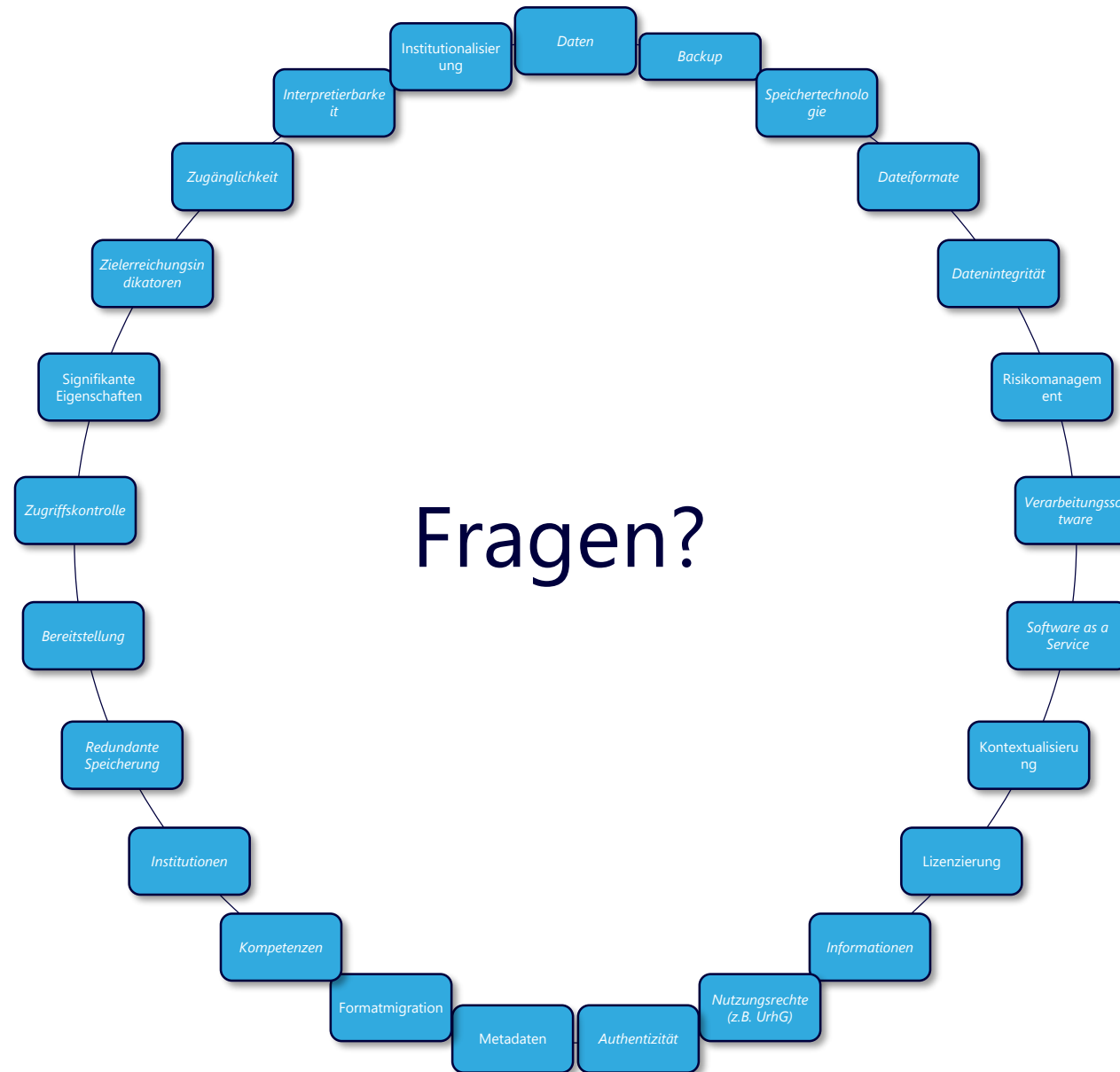
Einlieferungswege und Schnittstellen

- REST-API
- OAI-PMH
- BagIt
- METS
- Rsync

The collage consists of five overlapping panels, each representing a different submission method:

- REST-API:** A diagram showing a GET request to `https://lvz-test.hbz-nrw.de/rest/v0/deposits+?_wadt`. It breaks down the URL into Method (Abfrageart), Endpoint (Serveradresse), Ressource (Aktionen/Datensatz), and Parameter.
- OAI-PMH:** A screenshot of the OAI-PMH interface for 'Deposits: OAI-PMH: Harvest Job'. It shows fields for Name, Base URL, and various options like 'Contract and Edit', 'Set', 'User Name', 'Ignore Last Run Time', 'Match', 'XSL File', and 'Metadata Flow'.
- BagIt:** A screenshot of a BagIt manifest file showing a directory structure with files like `bagit.txt`, `manifest-md5.txt`, and `bag-info.txt`.
- METS:** A screenshot of an XML METS file structure, showing elements like `file` and `fileID` with various attributes.
- Rsync:** A screenshot of a terminal window showing the output of an `rsync` command, including details about the source and destination paths and options.







lzv.nrw

**Landesinitiative
Langzeitverfügbarkeit**

