



Universität
Zürich^{UZH}

The Zurich Open Repository and Archive ZORA

Transitioning from EPrints to DSpace CRIS

Dr. Martin Brändle

Zentrale Informatik/University Library, Open Science Services

Tessa Gerber, Margit Dellatorre, Olivier Vogel, Laetitia Kaiser, Andrea Malits

University Library, Open Science Services

Jens Witzel, Jann Forrer, Roberto Mazzoni, Stefan Vogt

Zentrale Informatik

Outline

- ZORA and its characteristics
- Rationale for change to DSpace-CRIS
- How we try to manage the transition
- Challenges, risks and mitigation

ZORA (Zurich Open Repository and Archive)

- Open and worldwide access to the research and scholarly output of the University of Zurich
- Focus is on qualified scientific publications
- Currently: >200K publications (69% with full texts, 51% OA)
- Operated by the University Library together with the Central IT of UZH
- Currently based on EPrints

ZORA on EPrints



Welcome to ZORA

ZORA (Zurich Open Repository and Archive) provides open and worldwide access to the research and scholarly output of the University of Zurich, Switzerland. A focus is on qualified scientific publications. ZORA is operated by the [University Library](#) together with the [Central IT](#) of the University of Zurich.



Latest Additions

Formal Theory of Mediation in Secession Conflicts: Analysis of Possible Cooperation Based on a Game Theoretical Model.

Meier, Nora .
2025, University of Zurich, Faculty of Arts. Item availability may be restricted.

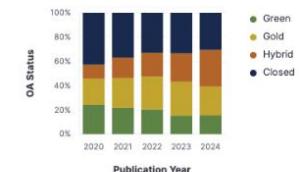
[→ Link](#)

Beyond Removal: Visibility Reduction in Content Moderation.



Open Access Distribution

Open Access Status by Year



Open Access by Faculty



Open Access News

23 Apr 2025

Highlights from the “PLATO Diamond Open Access Workshop Day”

Around 70 participants took part in 10 workshops on Diamond Open Access – a valuable platform for the exchange of knowledge, experiences and ideas.

[→ Link](#)

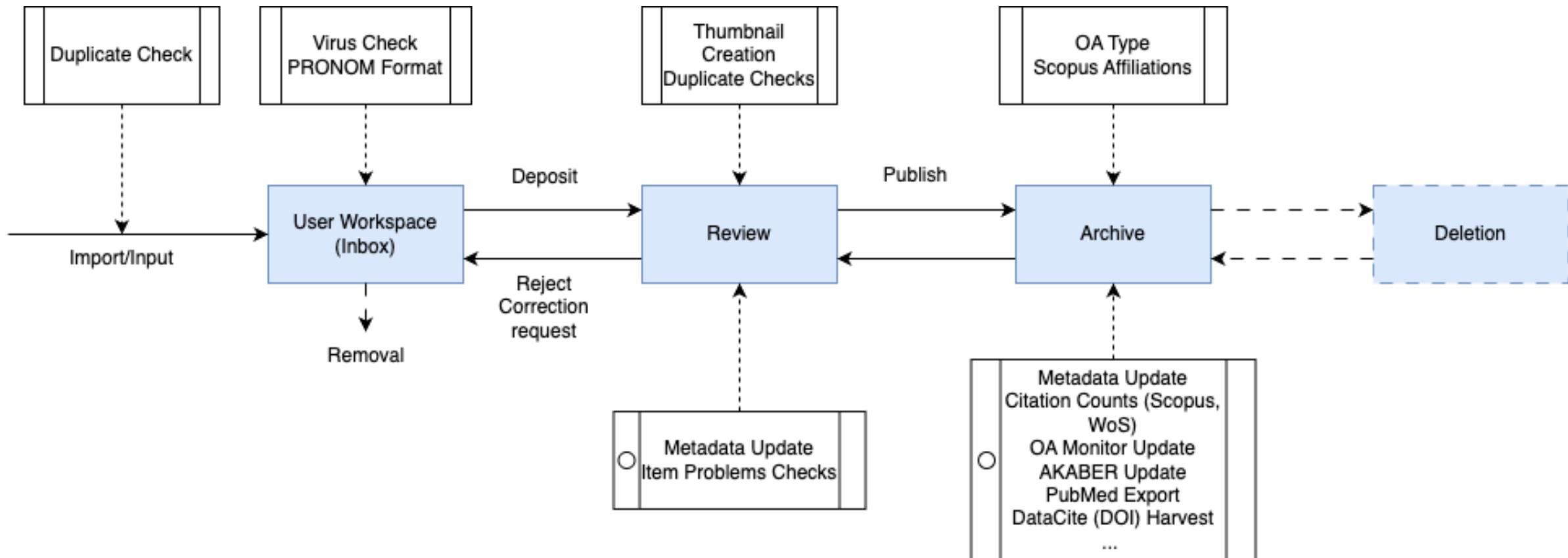
03 Apr 2025

ZORA 2.0 is launching – with 200 000 entries

ZORA Characteristics: Serving the UZH community

- Repository serves as publication database for UZH's annual Academic Report
=> high submission quota (10'000 publications yearly)
- Large active user base: 2000 submitters, 30 editors => Trainings
- Use of internal and external sources for metadata enrichment
- In-house ZORA-IT, serving requirements of University Library

ZORA characteristics: High automation in publication workflow



ZORA Characteristics: UZH extensions (1)

■ Journal Database (JDB)

- Journals and Publishers authority tables
- Copyright regulations, legal information e.g. from Open Policy Finder
- APC data
- R&P contracts

Lookup components in ZORA workflow

What can be self-archived in ZORA - Copyright regulations			
Accepted manuscript:	<input checked="" type="radio"/>	☒ 12 Months	 Publisher source must be acknowledged with citation Must link to publisher version with set statement (see policy) and DOI Policy
Published PDF:	<input type="radio"/>		Policy
Published PDF:	<input type="radio"/>		Policy
Published PDF:	 Hybrid OA option	 Published source must be acknowledged  Policy	
Published PDF:	 Hybrid OA option	 Published source must be acknowledged  Policy	
Published PDF:	 Hybrid OA option	 Published source must be acknowledged  Policy	
Read&Publish Agreement:	The publication is possibly published as part of a Read&Publish agreement and therefore could have an open licence. Please check the licence of the publication on the publisher's website or in the PDF.		
Additional Information:	Hybrid OA option: Please check the document status on the publisher's website. Keine Angaben online		

PDF (Published Version) available	
A PDF (Published Version) is available Open Access on the publisher's website at the following URL: https://onlinelibrary.wiley.com/doi/pdfdirect/10.1002/vrc2.478 . The publication is available under a CC BY-NC-ND license.	Download

ZORA Characteristics: UZH extensions (2)

■ OA Type Calculation

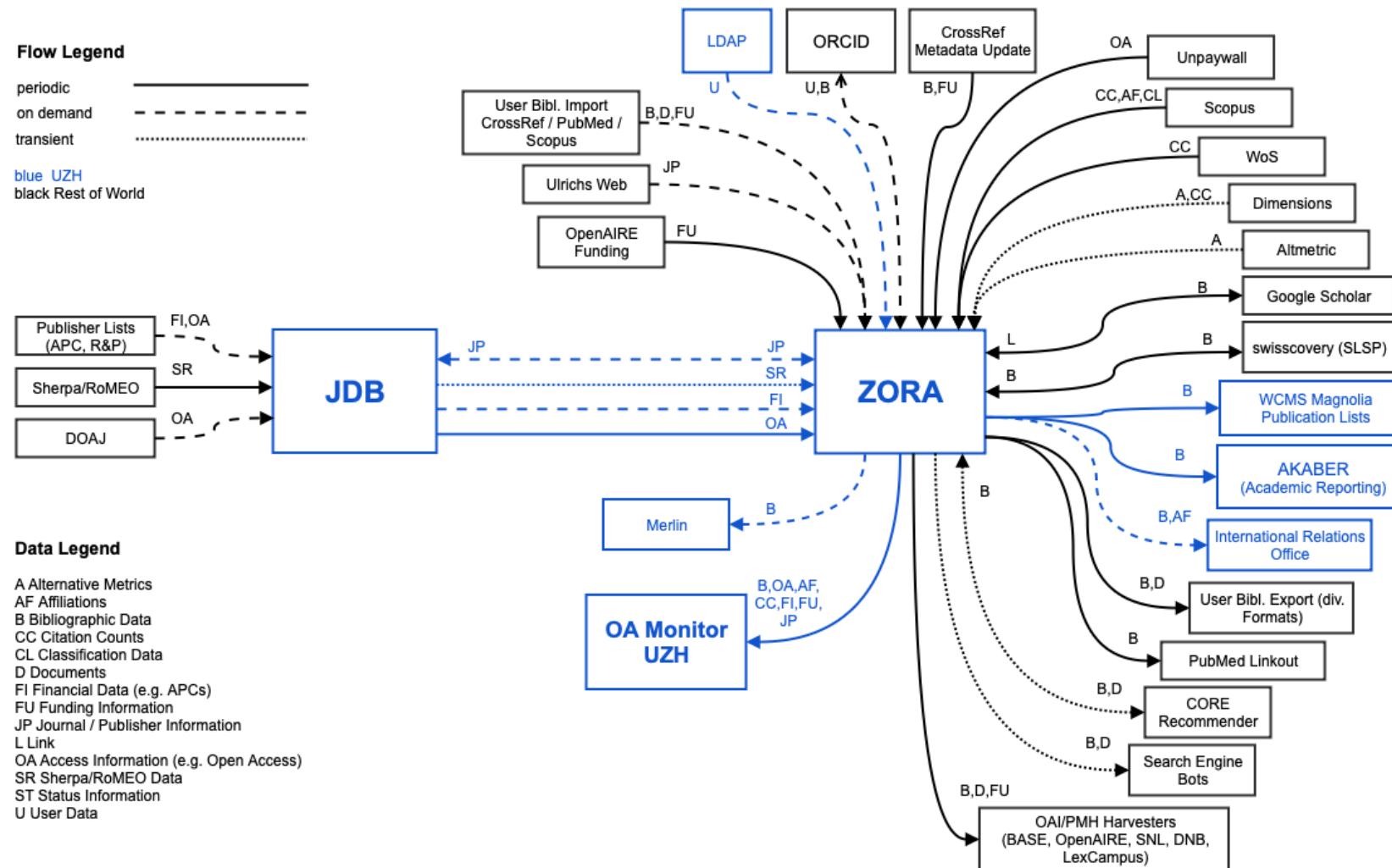
- Closed, Green, Hybrid, Gold
- Dynamic quantity
- Set of weighted (prioritized) sources:
ZORA, JDB, Unpaywall, DOAB
- Internal triggers for recalculation: Item change
- External triggers: Yearly recalculation



ZORA Characteristics: UZH Extensions (3)

- Further Extensions
 - Publication lists (component in WCMS)
 - ORCID Integration
 - Search Engine (Elasticsearch): Facets, Hit Highlighting, More-Like-This
 - Coversheets
 - UZH Layout
 - Affiliations and Correspondence Author
 - Author Collaborations (Visualization)

ZORA Data Flows: Integration of Data Sources and Data Consumers



Rationale for transition to DSpace: EPrints' weaknesses and strengths

- Weaknesses
 - Old programming language: Perl
 - No separation of data, control, presentation
 - Inbuilt search engine insufficient
 - No significant development in last few years
 - Community organisation and support
 - Increased internal maintenance
- Strengths
 - Flexible data model (relations, hierarchies, compound fields, multi-language)
 - Simple extensibility with plug-in architecture
 - Sophisticated system of roles and rights
 - Task queue and triggered tasks
- => Most Swiss universities switch to DSpace(-CRIS)
- => Decentralised, interoperable Swiss repository landscape

Transition Project

- Usual project life cycle
 - ✓ Current/Target Analysis
 - ✓ Requirements Engineering
 - ✓ Tendering
 - ✓ Awarding
 - ✓ Implementation
 - Testing
 - Trainings
 - Data migration
 - Switch, Production
 - Maintenance
- In parallel
 - Interviews with Swiss university DSpace operators
 - Get acquainted with DSpace: Setup of two small test servers
 - DSpace User Group Meetings (CH, DE)
 - Operation of current platform

Requirements

- 890 individual requirements in 36 groups
 - General
 - Architecture, Security
 - Functional, Interfaces
 - Metadata Model (e.g. “Berlin” model)
 - User/Roles/Authentication
 - Data Life Cycle (“Workflow”)
 - Migration
 - Presentation/Search/Discovery
 - Documentation/Training
- In scope/out-of-scope decisions
 - e.g. re-implement interface to OA monitor later, freeze data
- Requirements as base for test protocols
- In addition: Test scenarios

Challenges, Risks and Mitigation

- Challenges and Risks
 - Performance
 - Data model – EPrints strongly entity-based
 - Data curation and migration
Metadata mapping correct?
 - Mapping of old IDs and URLs (items, OUs)
 - Interfaces to data consumers
 - Triggered processes
- Mitigation
 - Load balanced server architecture, data caches
Performance tests on test environment
Transfer of bad bot rules
 - Consulting by PCG Academia, compromise model
 - Curation scripts (UZH)
Logging of migration process (export/import)
Checksums for documents, but metadata?
 - Solution by PCG Academia
 - Inform and test as early as possible
 - Solution by PCG Academia

Organization

- Preanalysis phase – helped to find common understanding
- Frequent meetings
- MS Teams chats
- UZH JIRA for issue tracking
- Code sharing

Opportunities

- Standard, state-of-the art software with active, lively community
- Modern UI/UX
- More CRIS functions for UZH?
- Interoperable repositories – exchange metadata and documents
- Open Source development: Give and Take