



From Fragmented Systems to One Repository:

Migrating Multiple Research
Data to DSpace-CRIS

EMPOWERING EXPERTISE.

Béatrice Keller, Monika Kriemler

12 May, 2026

Agenda

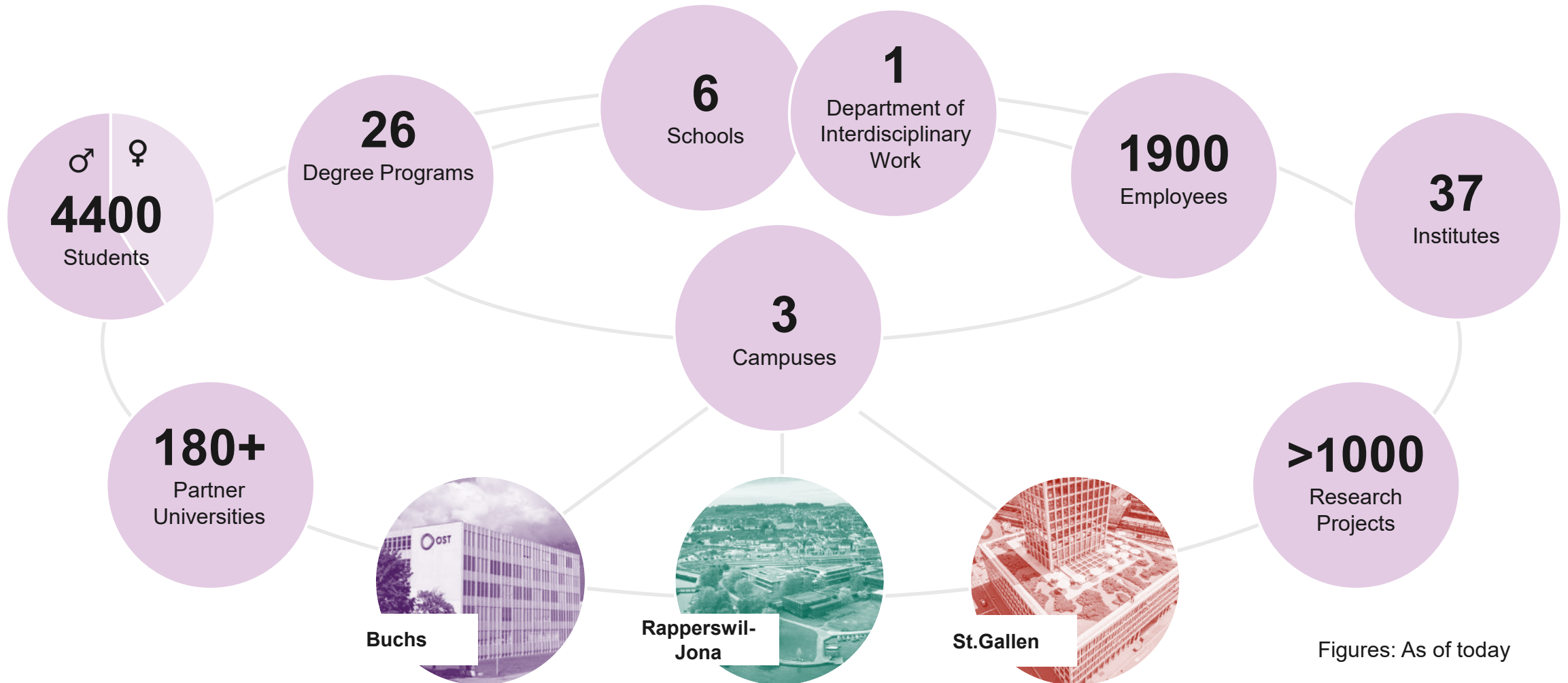
- Introduction
- The Data Landscape
- Migration Strategy
- Challenges and Lessons Learned
- Outcomes and Next Steps

Learning Objectives

1. Understand the **key challenges of migrating research data from multiple systems into a single repository.**
2. Learn strategies for **metadata harmonization and data cleaning during migration.**
3. Identify **technical and organizational considerations when implementing next-generation DSpace.**
4. Gain practical insights into **planning and executing complex migrations.**

Introduction

Eastern Switzerland University of Applied Sciences (OST)



Introduction

Why Migrate to Next-Generation DSpace?

Limitations

- Fragmented systems and workflows
- No integration between data sources
- Inefficient research reporting
- Limited visibility

Goal

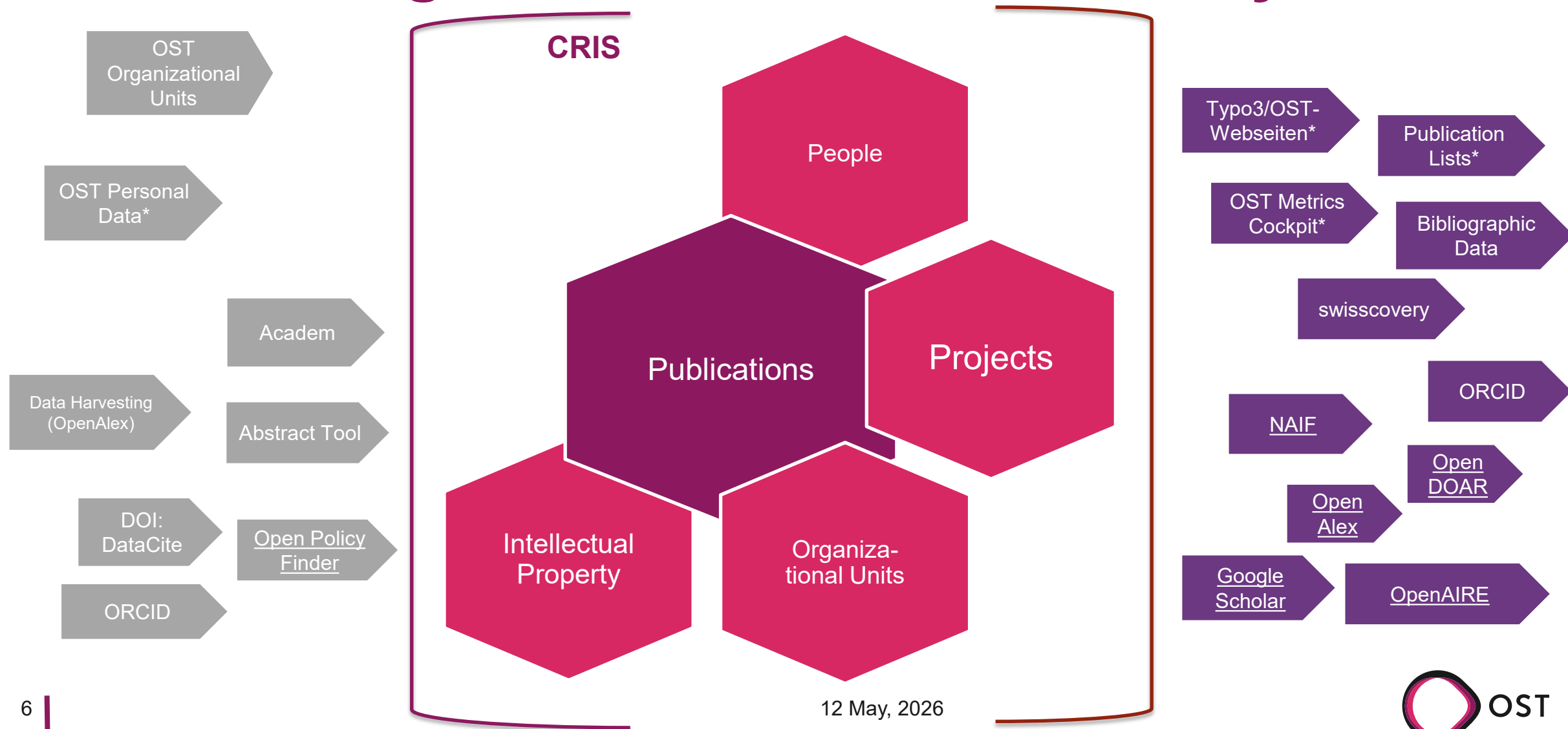
- Unified repository and CRIS
- Improved data quality and consistency
- Support for Open Access and reporting
- Better visibility of research outputs



OST Research Information Exchange

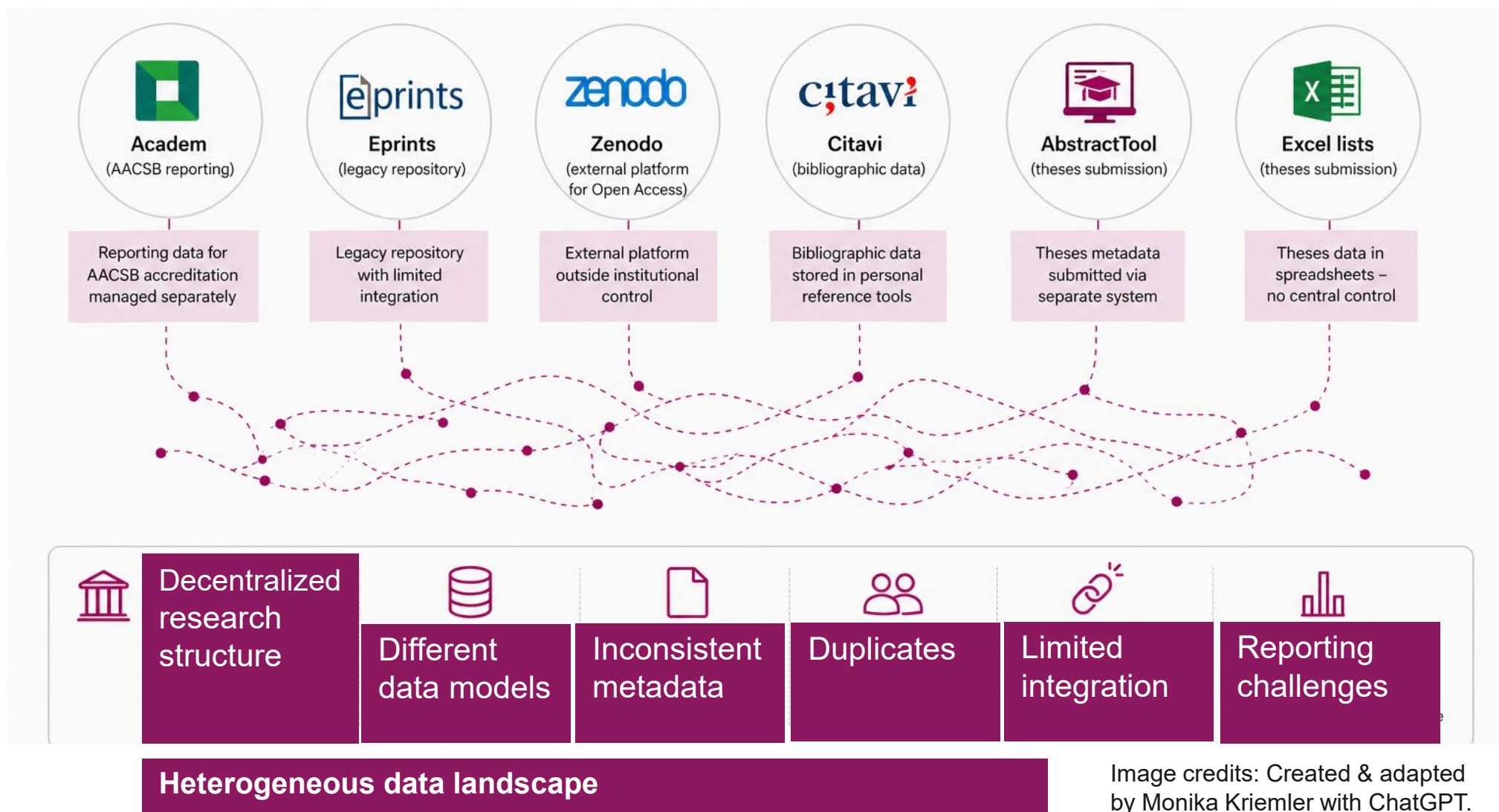
Introduction

Establishing a Research Information System



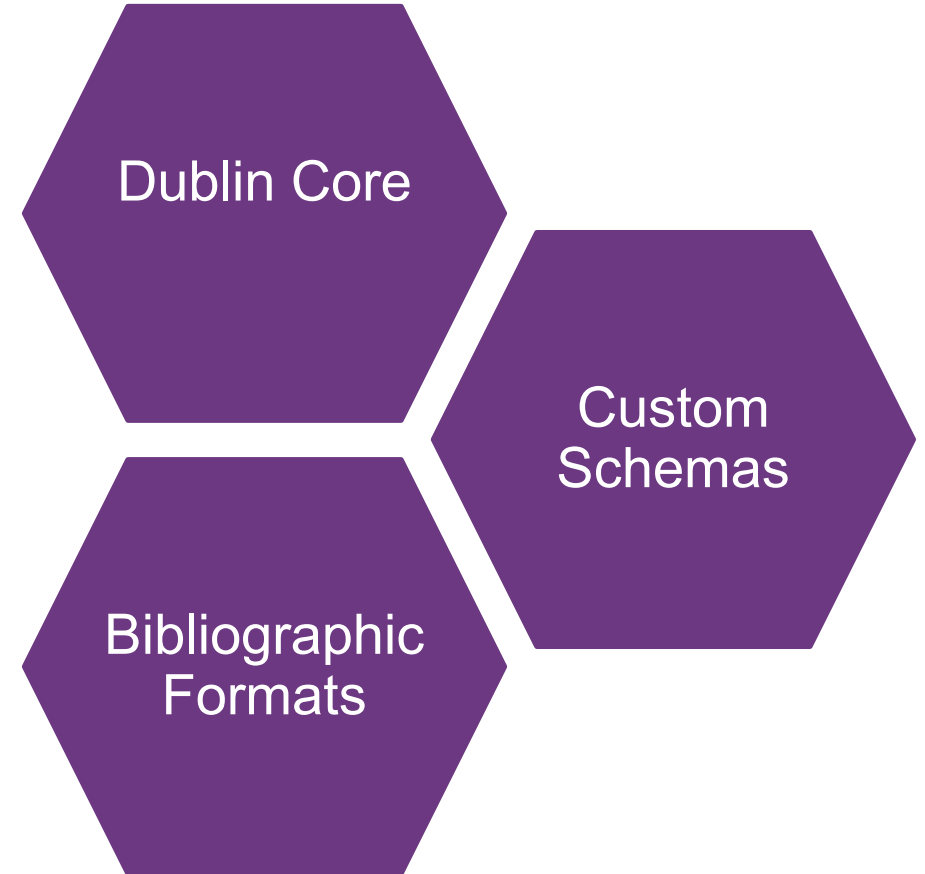
The Data Landscape

A Fragmented Research Data Ecosystem for Publications



Metadata Diversity

- Different field names and structures
- Inconsistent author or affiliation representation
- Missing or partial identifiers
- Metadata must be mapped, normalized, and enriched before integration in DSpace-CRIS



OST Subsystems for People & Organizational Units

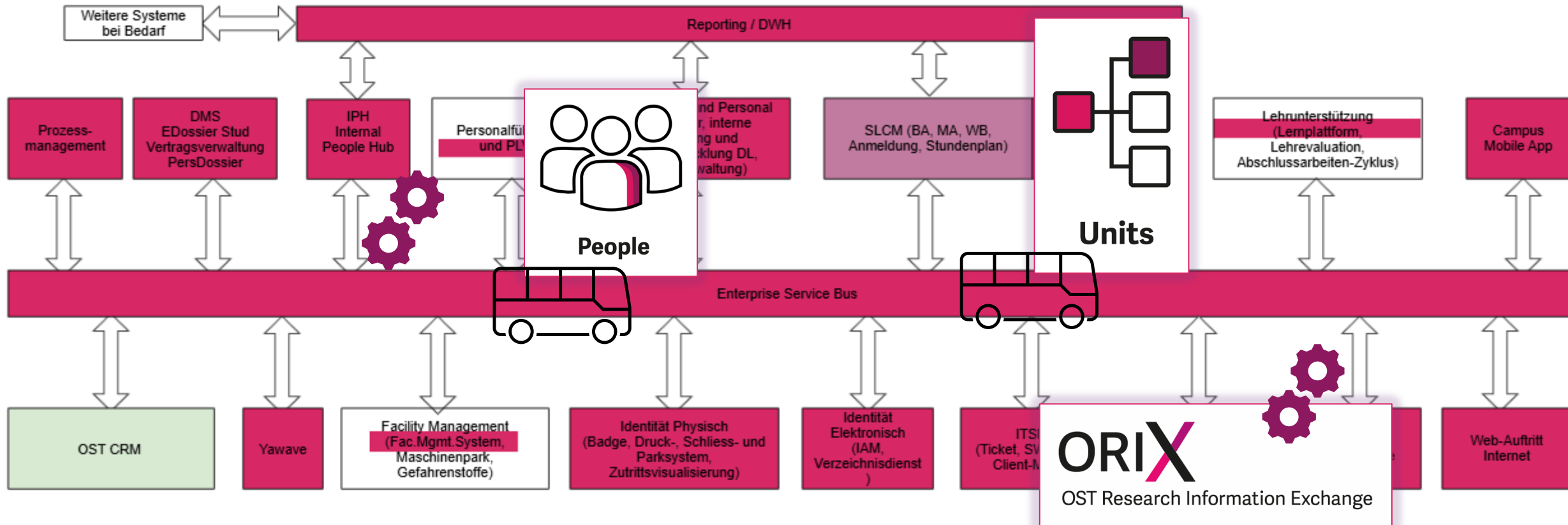
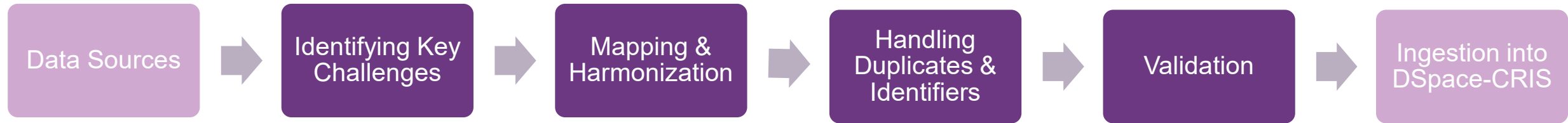


Image Source:
ICT OST

Migration Strategy

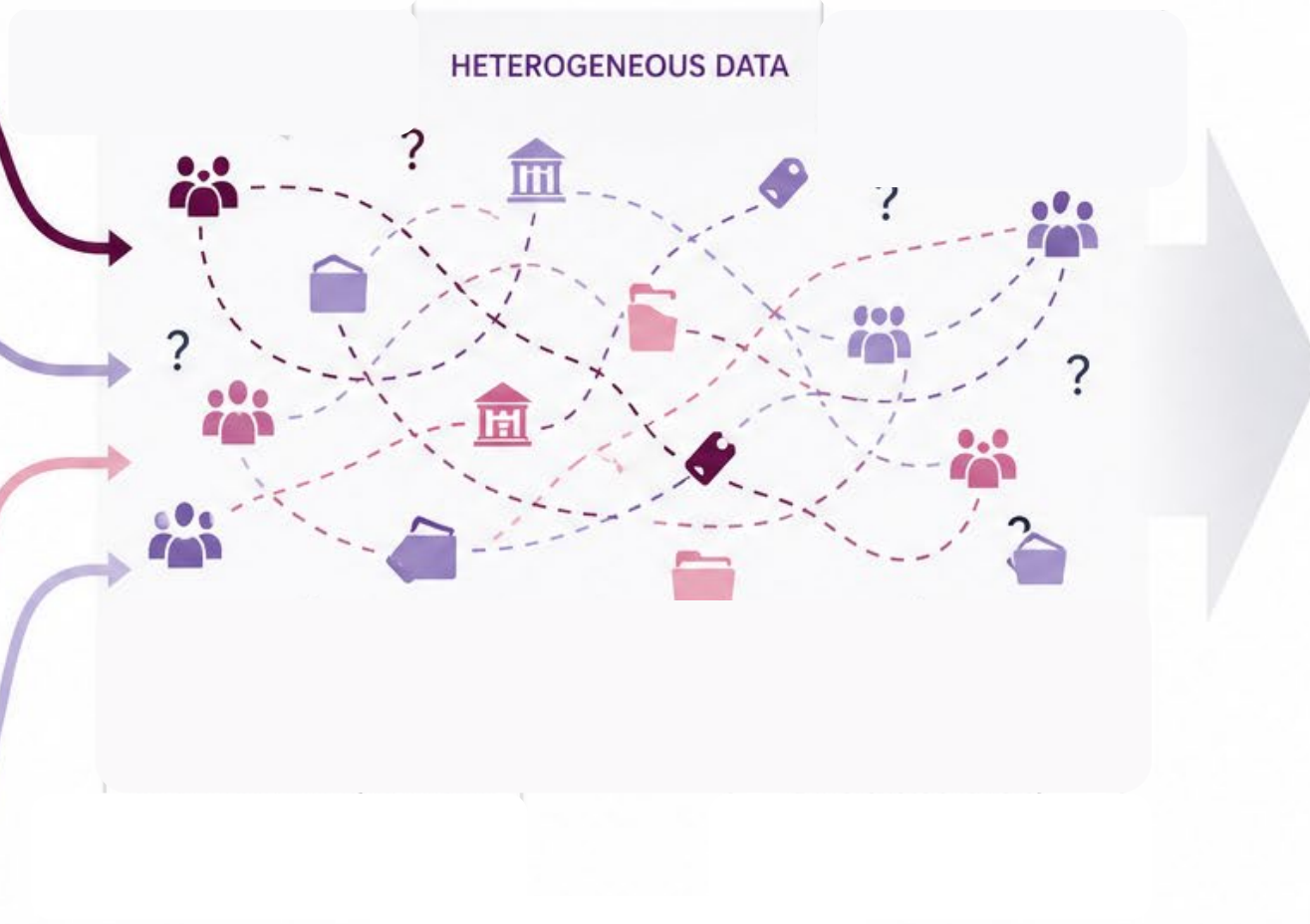
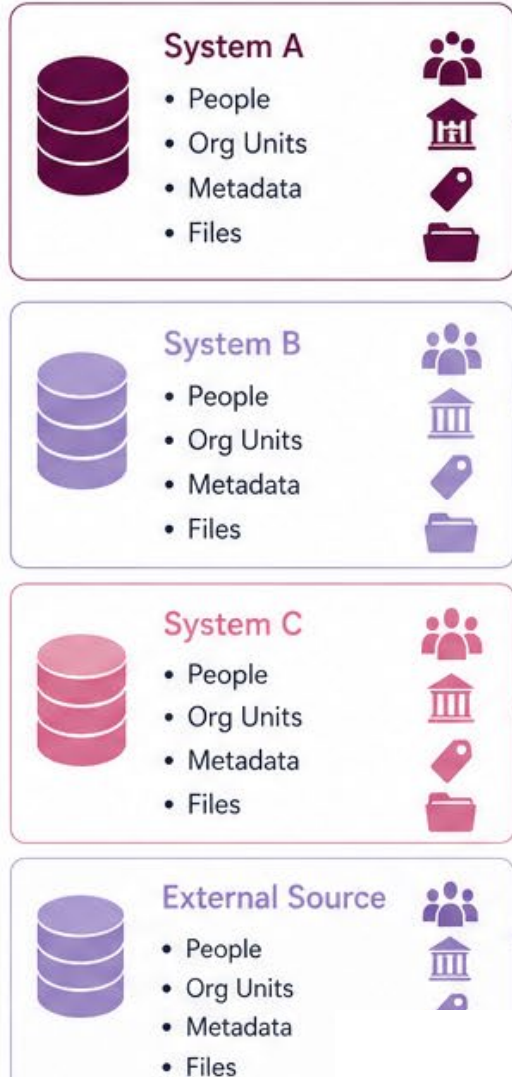
Migration Approach



- Migration is a multi-step transformation process — not a simple transfer

MULTIPLE SOURCE SYSTEMS

Migration Strategy



TARGET SYSTEM



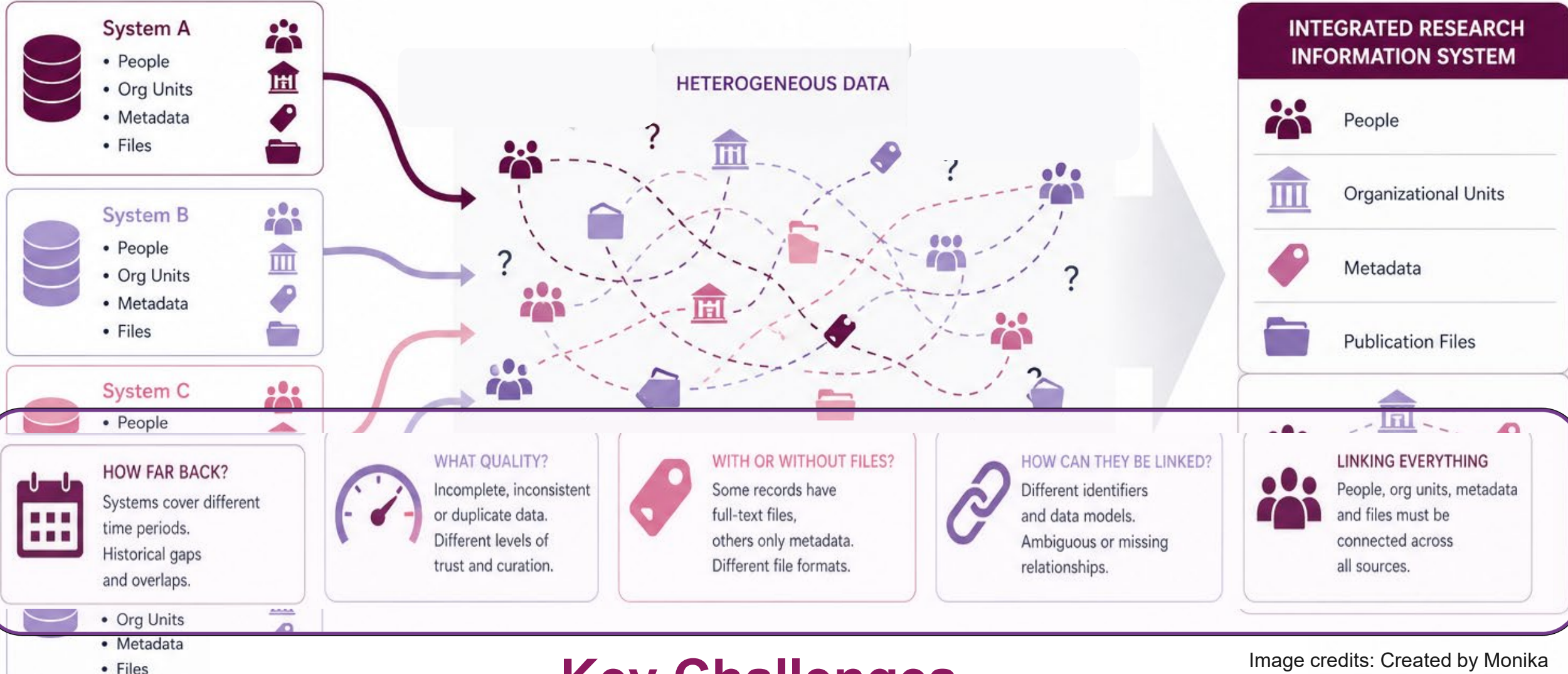
Key Challenges

Image credits: Created by Monika Kriemler with ChatGPT.

MULTIPLE SOURCE SYSTEMS

Migration Strategy

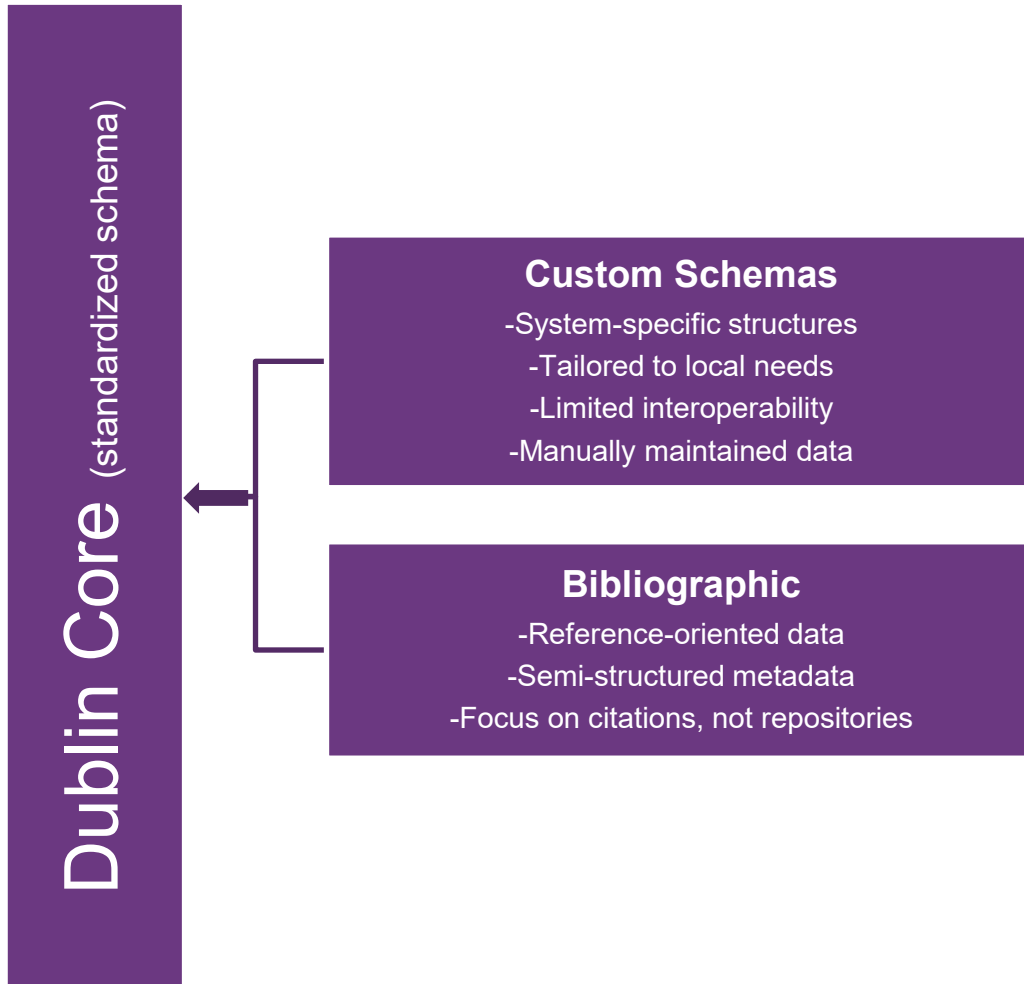
TARGET SYSTEM



Key Challenges

Image credits: Created by Monika Kriemler with ChatGPT.

Metadata: Mapping & Harmonization



- Mapping schemas to CRIS model
- Converting formats, normalizing structures, standardize fields
- Aligning publication types and metadata

JournalArticle				
CITAVI-Label (de)	Citavi xml	DublinCore	CRIS-Label (en)	Hint
Dokumententyp	DocumentType	dc.type	Type of Publication	
	FALSE	dc.affiliation.institution	Institution	default = Eastern
	FALSE	dc.affiliation.school	School / Division	
	FALSE	dc.affiliation.department	Department	only Technik
	FALSE	dc.affiliation.institute	Institute / Unit	
	Title	dc.title	Title	
	Subtitle	dc.title.alternative	AlternativeTitle	
	Language	dc.language.iso	Language	
	Persons.Person.Role	FALSE	FALSE	
	Persons.Person.FirstName	dc.contributor.author	Author	
	Persons.Person.LastName	dc.contributor.author	Author	
	Persons.Person.MiddleName	dc.contributor.author	Author	
	Periodical.Name	dc.relation.ispartof	Journal	
	Periodical.StandardAbbreviation	FALSE	FALSE	
Verlag	Publishers.Publisher	dc.publisher	Publisher / Ed. Institution	
	Volume	dc.description.volume	Volume	
	Number	dc.description.issue	Issue	
475-486	PageRange	dc.description.pagestart	Page Start	476
	PageRange	dc.description.pageend	Page End	-486

Mapping & harmonization for publication type JournalArticle, Citavi into Dublin Core

Handling Duplicates & Identifiers

- Detecting and merging duplicate records
- Using DOIs, ORCID, and internal IDs
- Pre-enrichment of author identifiers (PIDs)
- Consistency and identifiers are the foundation for linking data

The screenshot shows a user interface for a contact record. At the top, it says 'KONTAKT : CONTACT' and 'Monika Kriemler'. Below this is a section titled 'Zusammenfassung' (Summary). It contains two tables. The first table, 'KONTAKT INFORMATIONEN', lists: PId (100222), Vollständiger Name (Monika Kriemler), and Geschlecht (Weiblich). The second table lists: Abteilung (Bibliothek), Position (Projektleitung Publikationsdatenbank), Telefon Geschäft (+41 58 257 46 79), and Telefon OST (+41 58 257 46 79). Below these tables is a snippet of XML code. A large purple arrow points from the 'PId' field in the first table to the XML snippet. The XML snippet shows a root element with an 'Abstract' child and a 'Persons' child. The 'Persons' child contains a 'Person' element with 'Role' (Autor), 'LastName' (Kriemler, Monika:100222 | Brossard, Martin), and 'Sex' (Unknown). At the bottom of the XML snippet, there is a line of text: 'Keller, Béatrice::2001|Meier, Max|Huber, Felix::3845'. Below this line, it says 'xml for export out of CITAVI'.

KONTAKT INFORMATIONEN	
PId	100222
Vollständiger Name	Monika Kriemler
Geschlecht	Weiblich

Abteilung	Bibliothek
Position	Projektleitung Publikationsdatenbank
Telefon Geschäft	+41 58 257 46 79
Telefon OST	+41 58 257 46 79

```
<Persons>
  <Person>
    <Role>Autor</Role>
    <LastName>Kriemler, Monika:100222 | Brossard, Martin</LastName>
    <Sex>Unknown</Sex>
  </Person>
</Persons>
```

Keller, Béatrice::2001|Meier, Max|Huber, Felix::3845

xml for export out of CITAVI

Workflow and Tools

- Workflow design
 - Iterative process (test -> **validate** -> refine)
 - Data quality checks before / after ingestion
 - Sampling and manual review
 - Continuous feedback
 - Source-by-source ingestion
- Tools used
 - Scripting for transformation
 - Data processing pipelines
 - DSpace import tools

Challenges and Lessons Learned

Organizational Challenges

- Multiple stakeholders with different expectations and priorities
- Governance gaps & new requirements
- Limited understanding of the overall OST system landscape
- Limited availability of sufficiently skilled resources
- Careful change management

Challenges and Lessons Learned

Unexpected Challenges

- Legal and copyright issues
- System-specific edge cases and technical limitations
- Underestimated manual effort for data cleanup

Challenges and Lessons Learned

Key Lessons Learned

- Migration revealed data issues more than it created them
- Start with clear data models & standards
- Involve stakeholders early
- Don't underestimate availability of experts and university bodies
- Linking strategy is critical
- **Exchange ideas with colleagues from other repositories and learn from them**

Outcomes

Outcomes of ORIX

- Data quality & consistency
 - Improved data reliability
 - Reduced duplicates and inconsistencies
- Researcher benefits
 - Easier publication management
 - Increased visibility and citation impact
- Integration & Open Access
 - Improved system integration and transparency
 - Strengthened Open Access compliance

Outcomes

Next Steps

- Integration to OST web platforms
- Graduated visibility model for people
- Workflow optimization for submissions while adapting / evolving OST subsystems
- AI-powered search and discovery
- Communication at OST for ensuring a welcome acceptance of ORIX

Questions?

ORIX
OST Research Information Exchange

[Browse](#) [Communities](#)

[People](#) [Publications](#) [Projects](#) [Units](#) [Thesis](#) [Open Access](#)

Search in the repository

				
1534	13	0	0	1
Publications	Persons	Projects	Intellectual property	Organizational units

Kontakt

Béatrice Keller
Leitung Bibliothek

T direkt +41 58 257 45 30
T zentral +41 58 257 41 11
beatrice.keller@ost.ch



Bibliothek
OST – Ostschweizer Fachhochschule | Oberseestrasse 10 | 8640
Rapperswil | Schweiz | [ost.ch](https://www.ost.ch)

WO WISSEN WIRKT.

Monika Kriemler, Dipl. Natw. ETH
Projektleitung Publikationsdatenbank

T direkt +41 58 257 46 79
T zentral +41 58 257 41 11
monika.kriemler@ost.ch



Bibliothek
OST – Ostschweizer Fachhochschule | Oberseestrasse 10 | 8640
Rapperswil | Schweiz | [ost.ch](https://www.ost.ch)

WO WISSEN WIRKT.



Alle Logos und Symbole von Organisationen sind urheberrechtlich geschützt. Sofern ansonsten nicht anders angegeben, ist dieser Foliensatz lizenziert mit einer [CC-BY-4.0 International](https://creativecommons.org/licenses/by/4.0/) Lizenz.

Zitieren als: Béatrice Keller, Monika Kriemler. (2026. 12. Mai). From Fragmented Systems to One Repository: Migrating Multiple Research Data to DSpace-CRIS. [Konferenzbeitrag] The Big Migration, Warsaw. [DOI]