3-3-2021

Planning Batch Cataloging Projects

Rebecca B. French
James Madison University, frenchrb@jmu.edu

Planning Batch Cataloging Projects

Rebecca B. French, James Madison University

About Me

- Head of Metadata Analysis & Operations, JMU Libraries
- Background
  - Special formats cataloging
  - E-resources
  - Metadata for special and digital collections
  - Batch processing and automation

Overview

- Assess your starting point
- Plan a batch cataloging workflow
- Select appropriate technologies and tools

"books" icon by Fasobrun Jamil from The Noun Project

Batch Cataloging

- "Obtaining (or creating), transferring, manipulating, and editing groups of MARC bibliographic records" (Young, 2012)
- "Editing and adding large batches of MARC records to a catalog at once"; differentiated from "individually cataloging each title" (Turner, 2015)

Turner, A. (2015). Implementation of batch cataloging: A case study. In D. M. Mueller (Ed.), *Creating Sustainable Community: The Proceedings of the ACRL 2015 Conference, March 25-28, 2015* (pp. 303–314). Association of College and Research Libraries. http://www.ala.org/acrl/sites/ala.org.acrl/files/content/conferences/confsandpreconfs/2015/Turner.pdf
Young, P. (2012). A survey of batch cataloging practices and problems. *Technical Services Quarterly*, *29*(1), 22–41. https://doi.org/10.1080/07317131.2012.624450

Batch Cataloging

- Any of the following metadata actions performed in bulk
  - Collecting
  - Searching
  - Transforming
  - Matching
  - Editing
  - Loading

"books" icon by Fasobrun Jamil
from The Noun Project

Good Candidates for Batch Cataloging

- Similar items
  - Format
  - Other characteristics
- Need
  - Use
  - Discoverability
- Might not get cataloged if handling one-by-one

"Stack of CDs" icon by Juan Pablo Bravo from The Noun Project

Do you have materials that would benefit from batch cataloging?

Evaluate Resources

- Who will be working on this project?
- What skills do they already have?
  - Tools
  - Level of expertise
- What skills could be learned?
  - Professional goals
  - Availability of learning resources
  - Time needed to get to level of competency required for the project
  - Opportunity cost

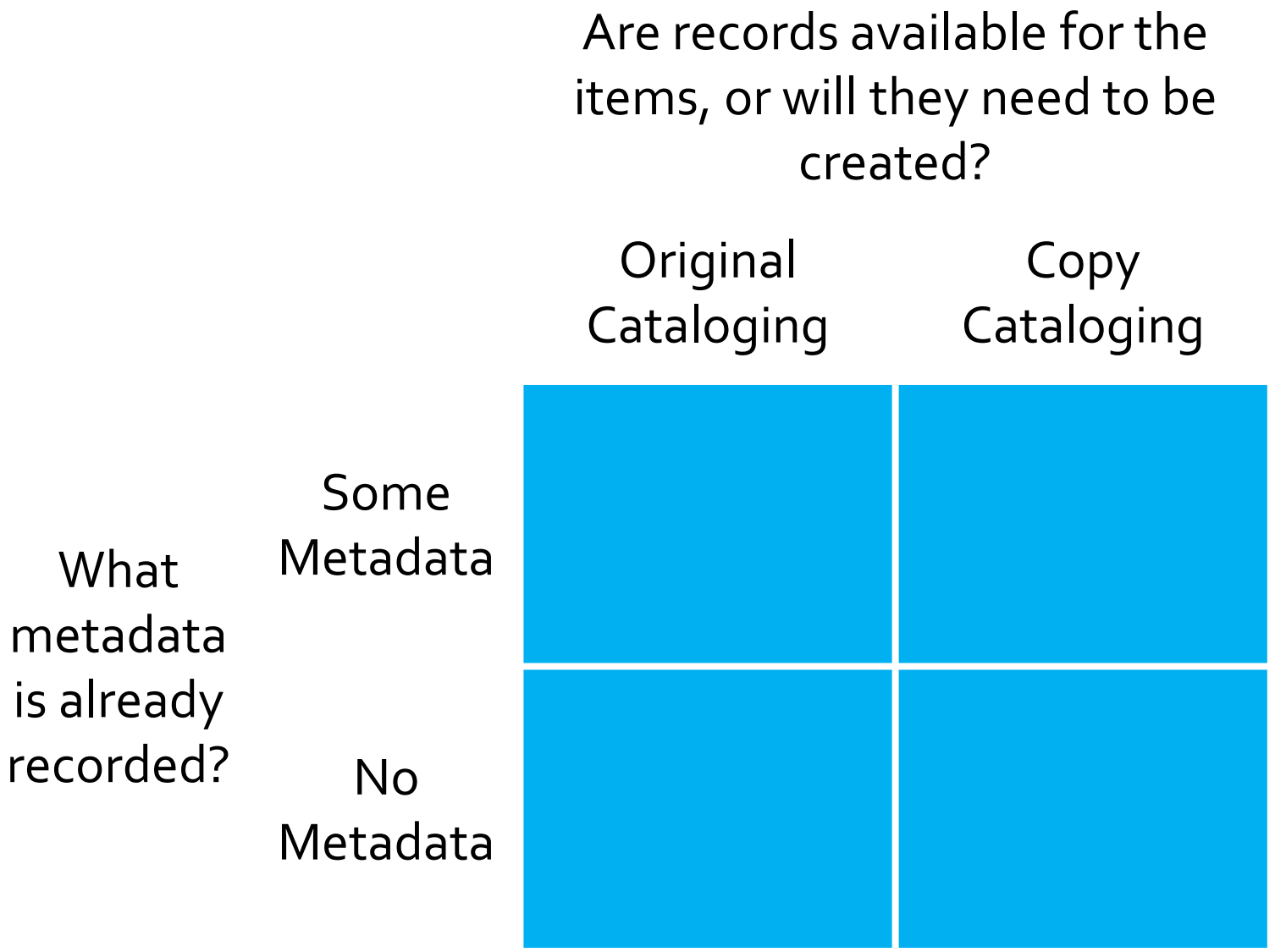
"hard skills" icon by lastspark
from The Noun Project

Two Key Questions

- Are records available for the items, or will they need to be created?
  - Records need to be created → original cataloging
  - Records are available → copy cataloging
- What metadata is already recorded about the items?

"Question" icon by Alice Design from The Noun Project

Batch Cataloging Matrix



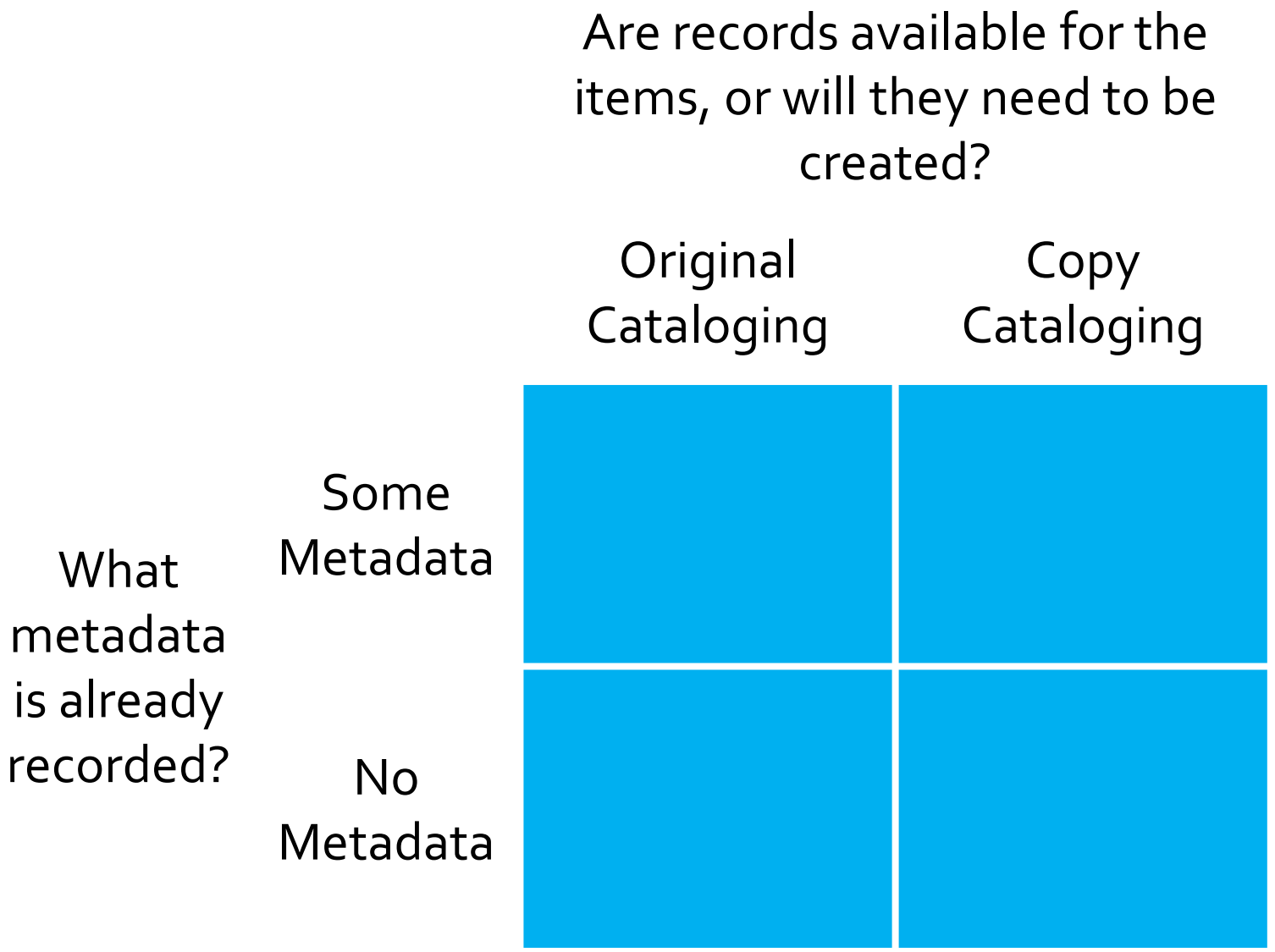
Two Key Questions

- Are records available for the items, or will they need to be created?
  - If there are existing records, what metadata is needed to search for those records?
- What metadata is already recorded about the items?
  - Is it unique enough to search on?
  - If not, what additional metadata would need to be recorded?

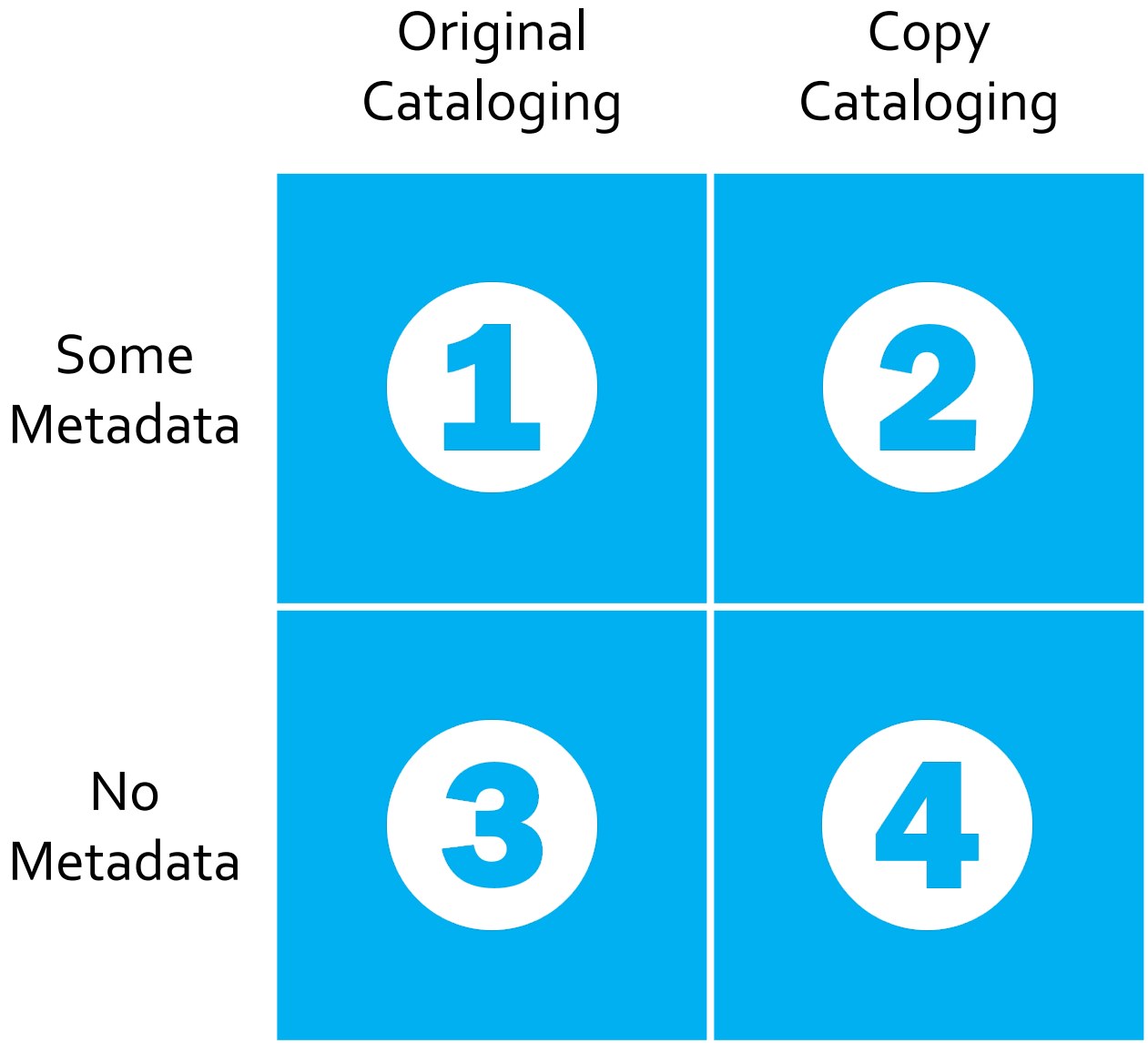
Additional Questions

- What editing needs to be done?
  - To existing metadata
  - To copy cataloged records
- How will these records be added to the ILS or other system?
  - If there are existing records that will need to be overlaid, what fields could be used as a match point?
- Is this a one-time or ongoing project?
- What checks do you need to perform during or after the project for quality control?

Batch Cataloging Matrix



Batch Cataloging Matrix



Which category does your batch cataloging project fall into?

Some Existing Metadata, Original Cataloging

1. Transform
2. Edit
3. Load

Example 1: ETDs

1. Transform
   - Student-submitted metadata (Qualified Dublin Core) into MARCXML
   - MARCXML into MARC binary
2. Edit
   - Add subject headings and classification
   - Adjust capitalization, spacing, etc.
3. Load

French, R. B. (2018). JMU ETD scripts [XSLT scripts]. https://github.com/frenchrb/JMU_ETDs

Some Existing Metadata, Copy Cataloging

1. Search
2. Match
3. Edit
4. Load

Example 2: Jazz LPs

1. Search
   - Batch search OCLC using publisher name and issue number
2. Match
   - OCLC records to local bib numbers
3. Edit
   - Clean up fields per local guidelines
4. Load

French, R. B. (2017). Direct database access to OCLC Connexion's local save file. *Code4Lib Journal, 38*. http://journal.code4lib.org/articles/12821
French, R. B. (2020, February 25). *Secrets of the save file: Using Connexion and Microsoft Access for efficient batch cataloging projects*. Presentation at the Music OCLC Users Group Meeting, Norfolk, VA. https://commons.lib.jmu.edu/letfspubs/185/

No Existing Metadata, Original Cataloging

1. Collect
2. Transform
3. Edit
4. Load

Example 3: Comic Books

1. Collect
   - Staff and students record metadata in spreadsheet
2. Transform
   - Spreadsheet into MARC binary
3. Edit
   - Add boilerplate fields
   - Reformat data from spreadsheet (e.g., names in direct/indirect order)
4. Load

French, R. B. (2021). comics2marc [Python script]. https://github.com/frenchrb/comics2marc

No Existing Metadata, Copy Cataloging

1. Collect
2. Search
3. Match
4. Edit
5. Load

Example 4: CD Backlog

1. Collect
   - Add UPC barcodes to brief records in ILS
2. Search
   - WorldCat Search API lookup in OpenRefine
3. Match
   - OCLC records to local bib numbers
4. Edit
   - Generate call numbers
   - Clean up fields per local guidelines
5. Load

Batch Cataloging Matrix with Project Stages

	Original Cataloging	Copy Cataloging
Some Metadata	Transform Edit Load	Search Match Edit Load
No Metadata	Collect Transform Edit Load	Collect Search Match Edit Load

Do you have experience editing records in batch?

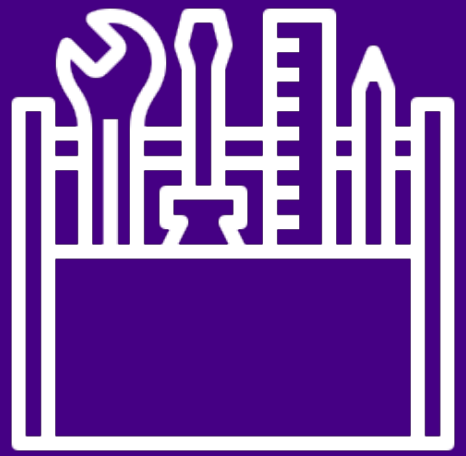
Do you have experience searching for records in batch?

Do you have experience transforming metadata in batch?

Do you have experience matching or merging metadata records in batch?

Project Stages

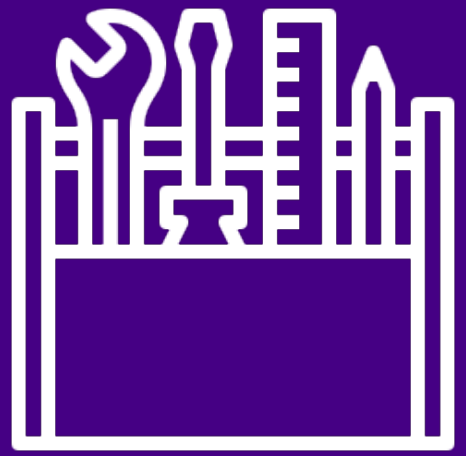
- Collect
- Transform
- Search
- Match
- Edit
- Load



Tools for Collecting

- Spreadsheets
  - Microsoft Excel
  - Google Sheets
- Forms or surveys
  - Google Forms
  - Qualtrics
  - Airtable
- Custom apps

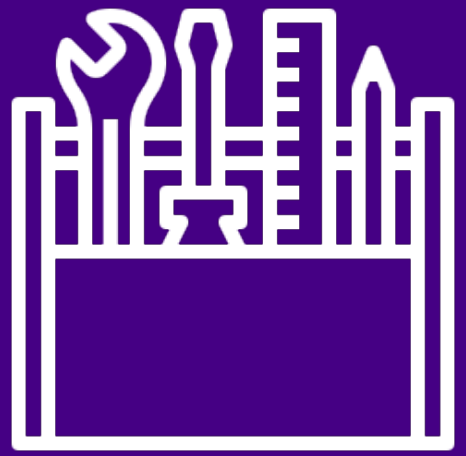
"tools" icon by monkik
from The Noun Project



Tools for Transforming

- MarcEdit's Delimited Text Translator
- OpenRefine export templating
- Scripting
  - XSLT
  - Python

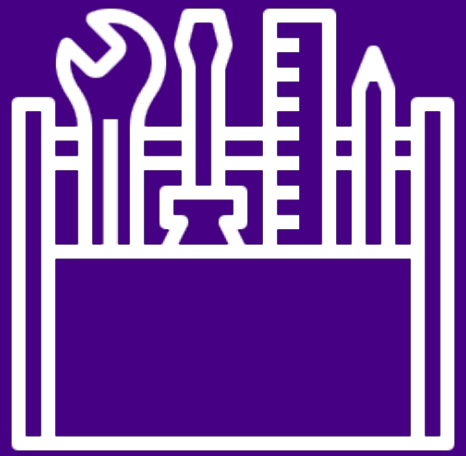
"tools" icon by monkik
from The Noun Project



Tools for Searching

- OCLC Connexion client batch searching
- Z39.50
- APIs
  - OpenRefine URL lookups
  - Scripts
- SQL

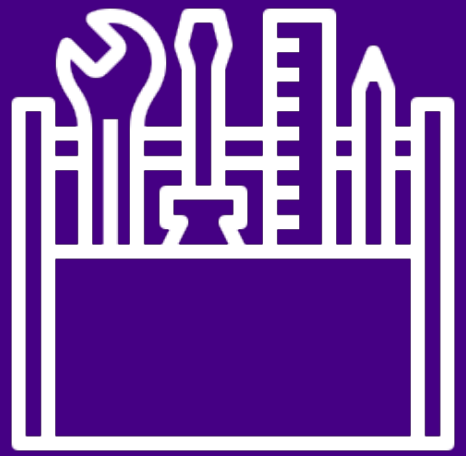
"tools" icon by monkik
from The Noun Project



Tools for Matching

- MarcEdit's Merge Records function
- Excel lookups
- Databases
  - Microsoft Access
  - SQLite
- ILS import profiles/load tables

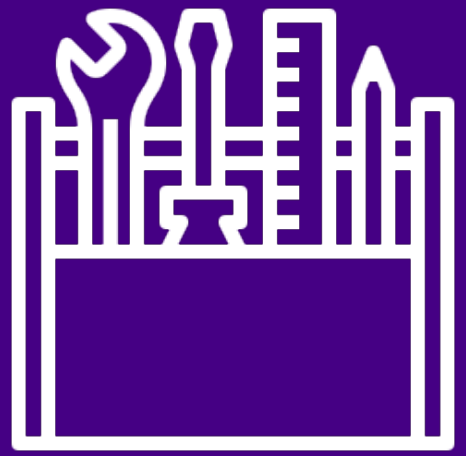
"tools" icon by monkik
from The Noun Project



Tools for Editing

- MarcEdit
- ILS batch update function
  - Sierra Global Update
  - Alma normalization rules
- Scripting
  - XSLT
  - Python

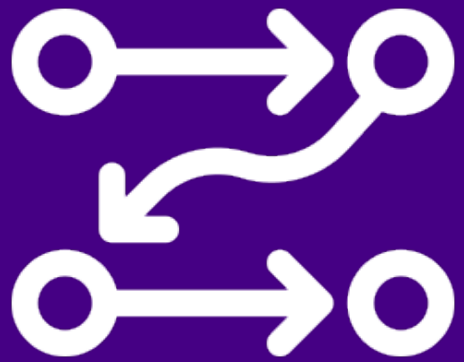
"tools" icon by monkik
from The Noun Project



Tools for Loading

- Function of ILS, IR, or bibliographic utility
- ILS import profiles/load tables
  - Customization
  - If overlaying existing records, consider match point and whether any fields need to be protected
- Manual or automated upload
  - FTP
  - API

"tools" icon by monkik
from The Noun Project



Mapping Pathways between Tools

- Select potential tools for each stage, based on available systems and skills
- Output of one stage becomes the input for the following stage

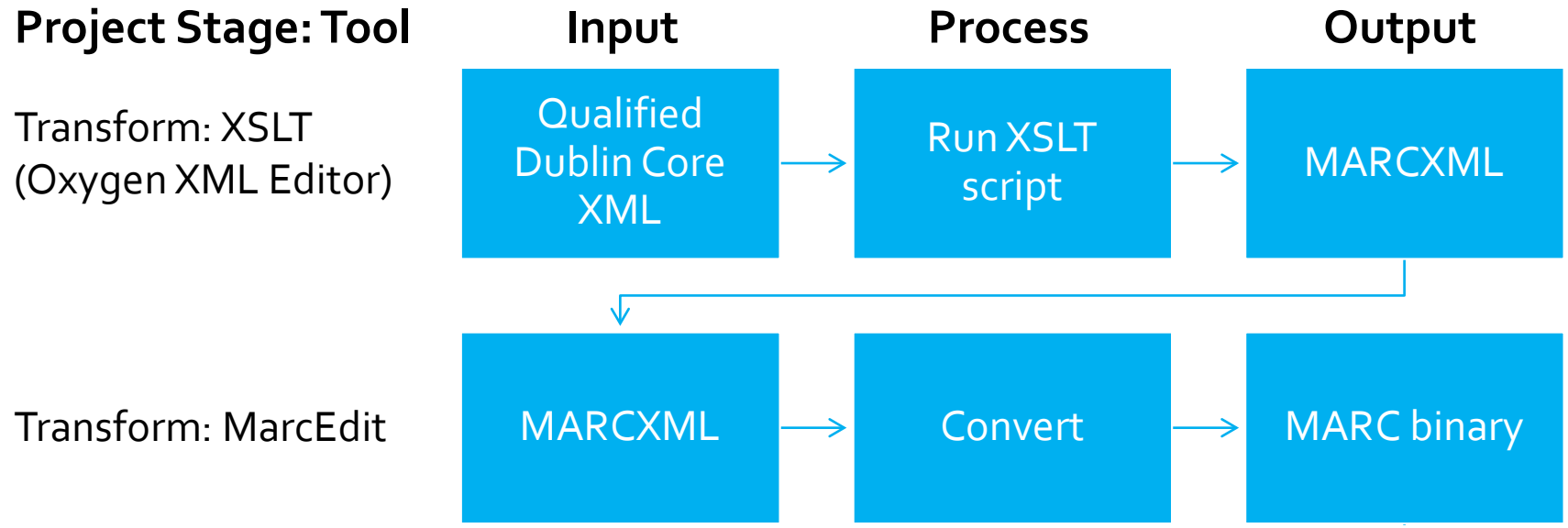
Mapping Pathways between Tools

Example 1: ETDs

Project Stage: Tool	Input	Process	Output
Transform: XSLT (Oxygen XML Editor)	Qualified Dublin Core XML	Run XSLT script	MARCXML

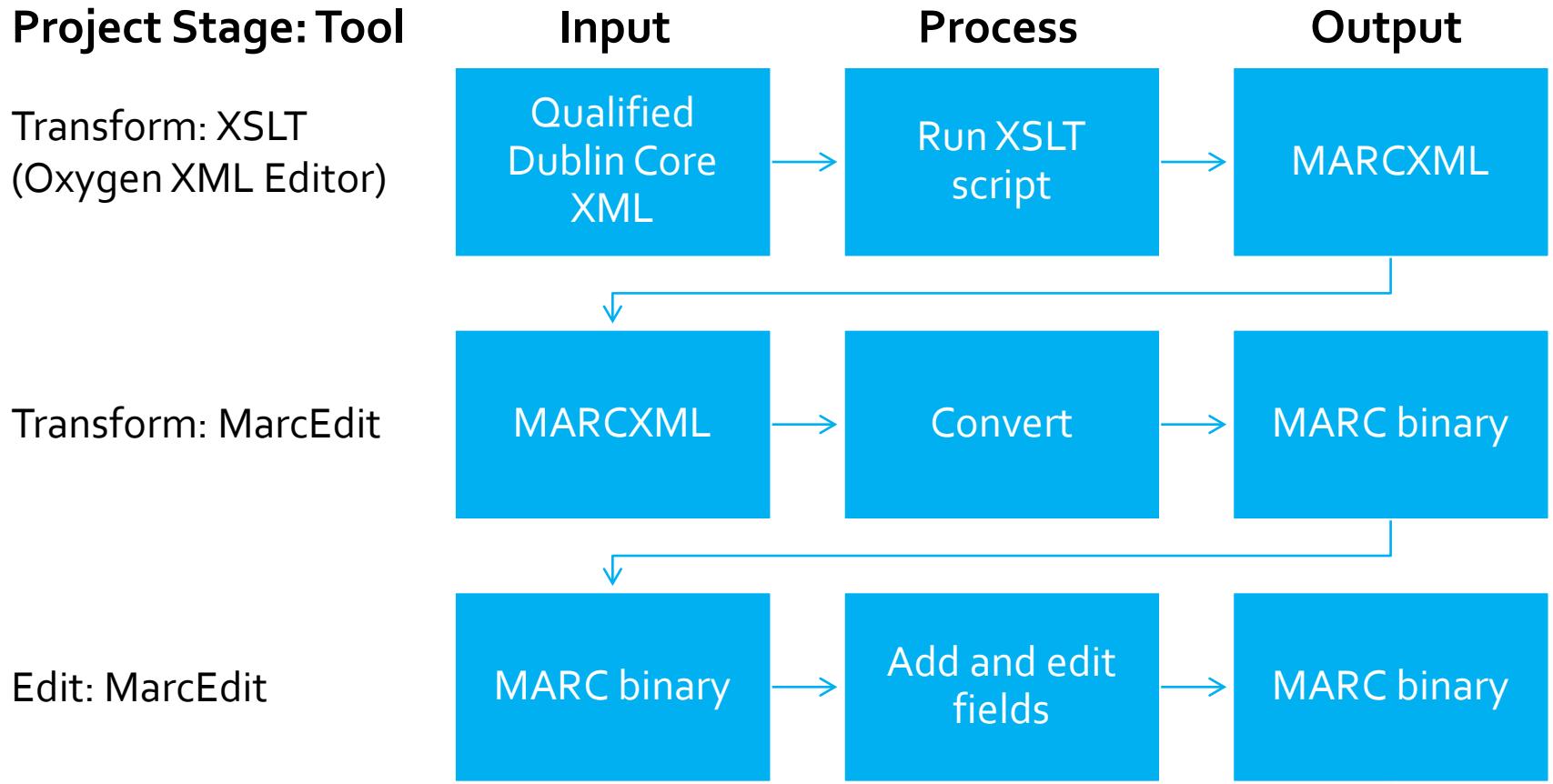
Mapping Pathways between Tools

Example 1: ETDs



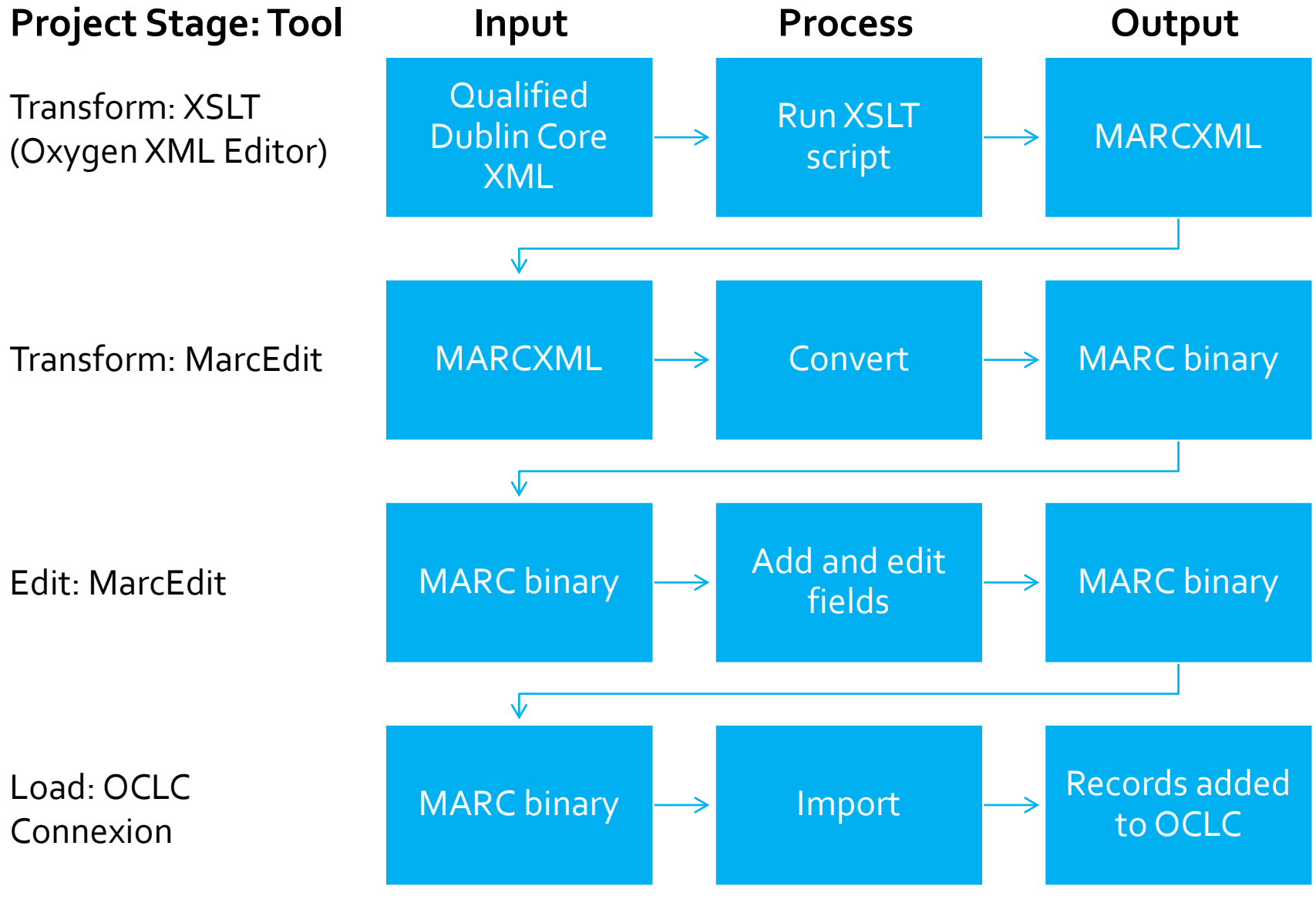
Mapping Pathways between Tools

Example 1: ETDs



Mapping Pathways between Tools

Example 1: ETDs

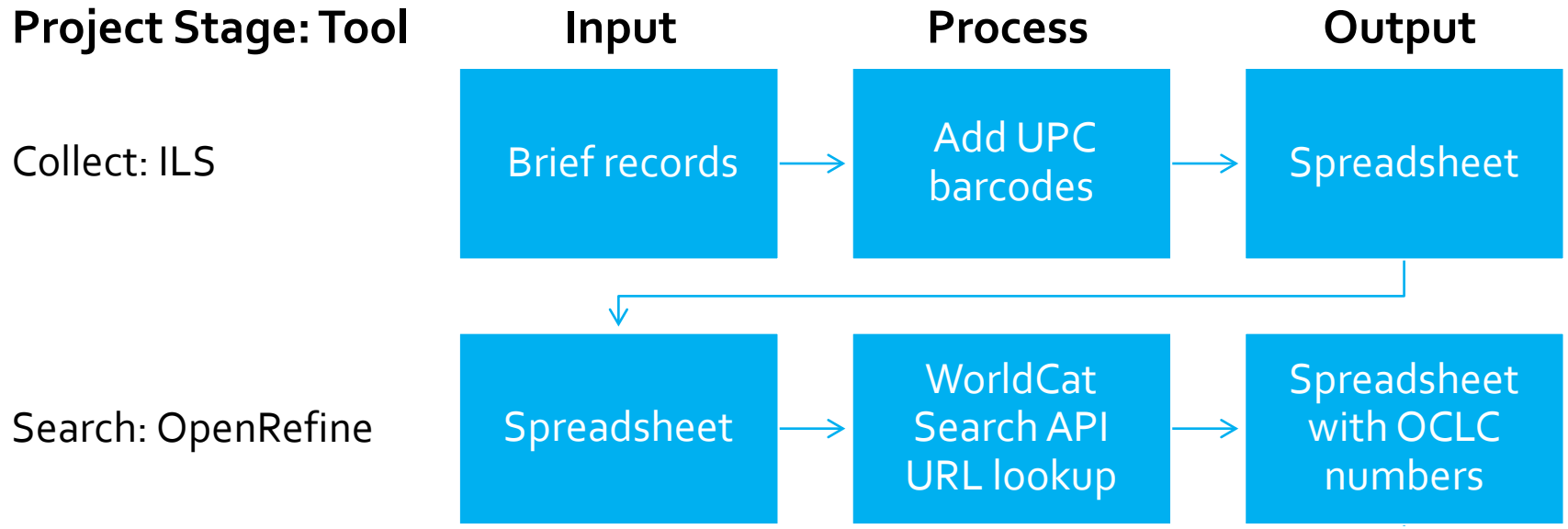


Mapping Pathways between Tools

Example 4: CD Backlog

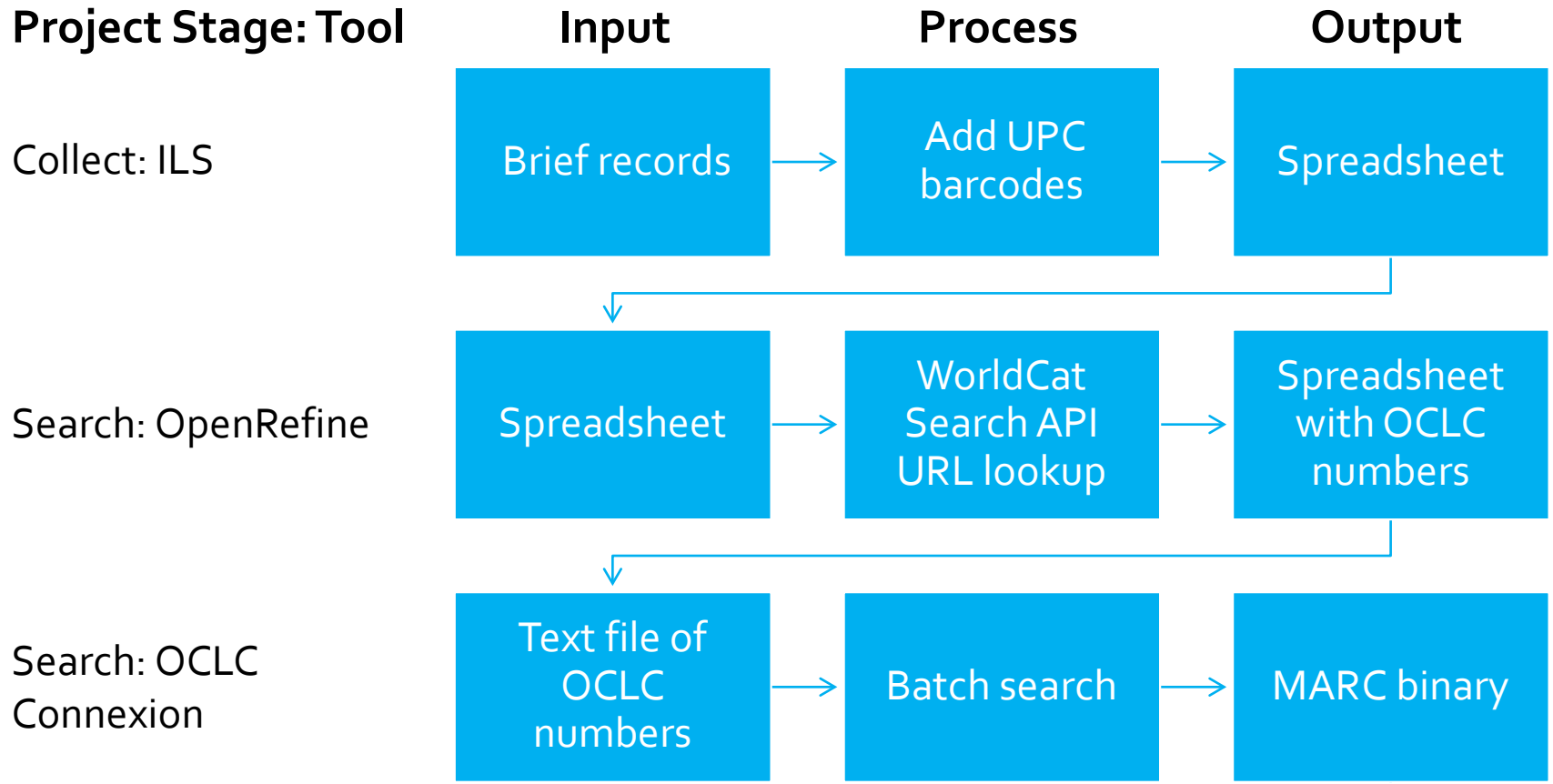
Mapping Pathways between Tools

Example 4: CD Backlog



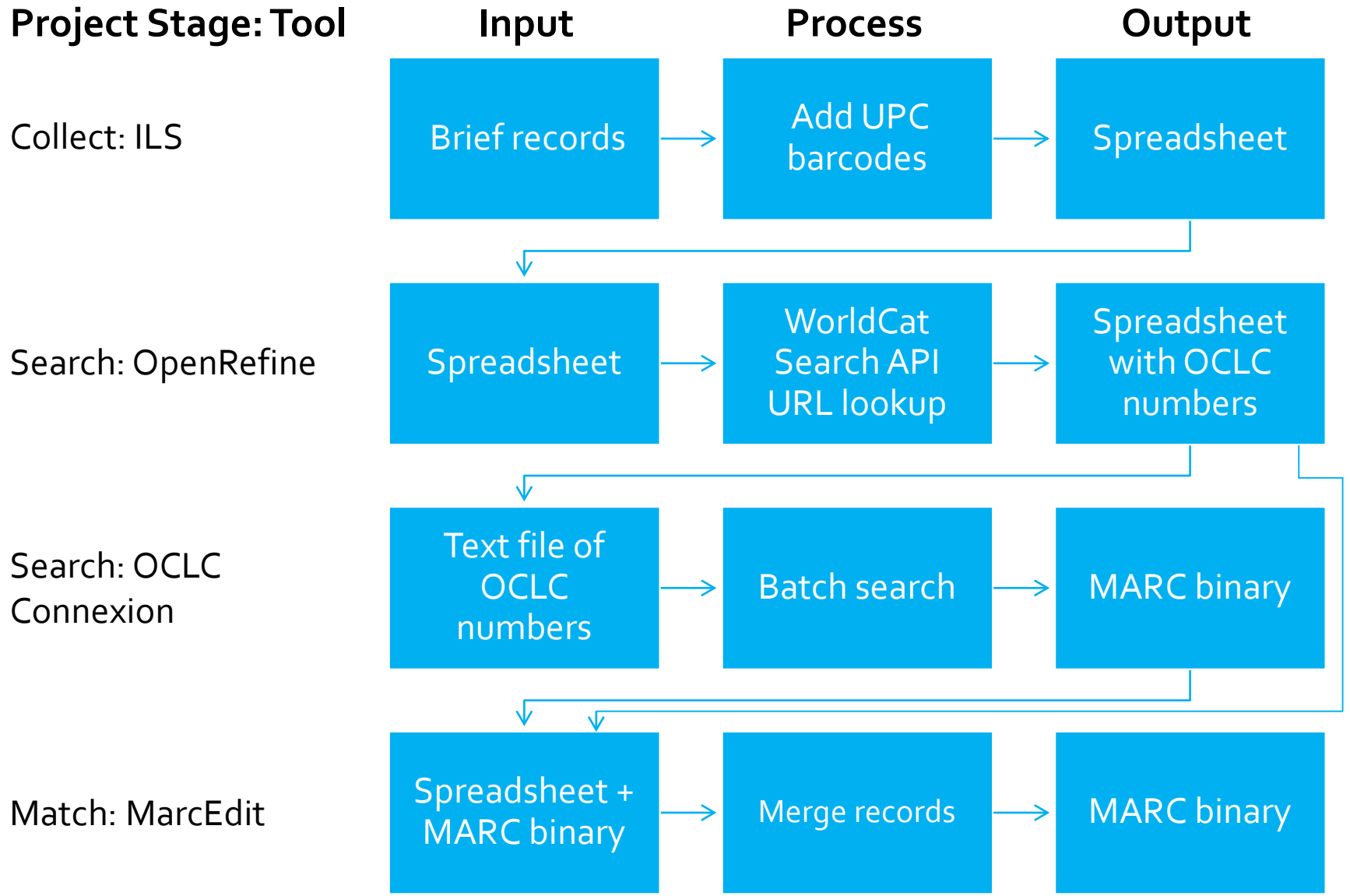
Mapping Pathways between Tools

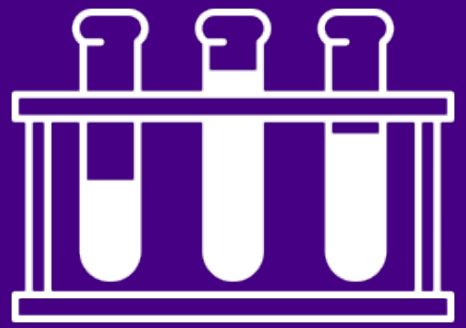
Example 4: CD Backlog



Mapping Pathways between Tools

Example 4: CD Backlog





Testing

- Walk through entire process
  - Single record
  - Small group of records
- Sandbox or testing environment
- Identify points for quality control checks

Documentation and Training

- Write down all the steps, even the ones you think you'll remember later
- If people who were not involved in planning the project or workflow will be performing any of the work, train them on the process
- Also useful if you want to share about your project later

"write documentation" icon by Juicy Fish from The Noun Project

Questions?

Planning Batch Cataloging Projects

	Original Cataloging	Copy Cataloging
Some Metadata	Transform Edit Load	Search Match Edit Load
No Metadata	Collect Transform Edit Load	Collect Search Match Edit Load