Full title: Using OpenRefine's Reconciliation to Validate Local Authority Headings

Authors: Scott Carlson and Amber Seely

Abstract**:** In 2015, the Cataloging and Metadata Services department of Rice University's Fondren Library developed a process to reconcile four years of authority headings against an internally developed thesaurus. With a goal of immediate cleanup as well as an ongoing maintenance procedure, staff developed a "hack" of OpenRefine's normal Reconciliation function that ultimately yielded 99.6% authority reconciliation and a stable process for monthly data verification.

Keywords**:** data reconciliation, authority control, thesauri, OpenRefine, digital archives, linked data

Background

In 2011, Rice University's Fondren Library began the Historical Images (HI) project, a program to digitize Rice-related photographs, maps, letters and other documents of historical interest and make them available digitally to the Rice community. The Woodson Research Center, Fondren's special collections and archives department, began the project with roughly 50 GBs of image files from their internal reference work -- in particular, items digitized in preparation for the university's centennial celebration on October 12, 2012. Three key Fondren departments -- Woodson, Digital Scholarship Services, and Technical Services -- collaborated on the project's guidelines and workflows. By April of 2013, the initial 50 GBs of images were processed and made available on the Rice Digital Scholarship Archive (RDSA), with more materials filling up the workflow queue.

Technical Services' role in the project centered on enriching image metadata beyond basic descriptions, leading to a highly unique local thesaurus. Historical lists of building names, university departments, and schools were cross-referenced for name changes during the school's history and checked against the Library of Congress Name Authority File (LCNAF); naturally, a large number of local, university-centric headings needed to be constructed, along with headings related to the city of Houston. The next focus was a large number of personal name headings (faculty, staff, students, members of the Houston business community, city and state politicians), also checked against LCNAF. The use of local headings gave Technical Services the latitude to delve deeply into the metadata of each image, and every attempt was

made to model local headings upon valid forms already in LC. Robert Estep, Fondren's Copy Cataloging Unit Team Lead and coordinator of the HI thesaurus, would later write about the project:

> "Building a progressive thesaurus, which is both a melding of valid LC headings and local headings, requires the flexibility of being able to return from time to time, sometimes to tweak, other times to undo earlier work and start from scratch. But the pattern we have found is that each return is both easier and shorter, as we digest the lessons of embarking on a project which extends both into the past via the images themselves, and into the future as the life of the University, its teachers and its students, continues to be documented."(1)

By mid-2016, this thesaurus had grown to almost 4,500 unique headings in the **dc.subject.lcsh** field of over 3,000 digital objects. Because of the thesaurus size, processing and validating headings added to new images was time-consuming, but not impossible. Quality control issues were a more pressing concern, given that authorized headings -- especially names -- are subject to change over time (along with the rules and standards governing their creation). As both the number of digital objects and their related headings increase over time, so does the likelihood of human input error and inconsistencies in application. To add fuel to an accelerating fire, the HI thesaurus was originally maintained as an Excel spreadsheet, a practice not recommended(2) largely due to its inability to scale for large projects. Fondren needed a way to batch-reconcile the growing collection data against a local thesaurus (or thesauri) to verify correctly-used headings, clean up incorrect headings, and find headings that had fallen through the cracks during thesaurus updates.

Much has been written in the past decade about the state of thesauri and authority control in digital collections; Dragon neatly summed it up in 2009 as "vitally important but frequently overlooked."(3) Salo cites batch ingests of large amounts of data as one of the primary sources of this problem, implicating the creators of this uncontrolled access data as well as the librarians and digital repository managers who make it available.(4) Salo ends her assessment of name authority control in institutional repositories noting that such librarians and digital repository managers "can only plan to plow large amounts of staff time into managing names, and petition their software developers for more efficient ways to bring order to chaos."(5) With regard to the need for such automatic maintenance, Dragon writes about a postcard digitization project initiated by East Carolina University's Joyner Library: "Without the

ability to effect a global change by changing one authority record, the cataloger must manually update each affected bibliographic record if it is determined that a heading needs to change. There is no workaround for this. The change process can be made as easy as possible with the inclusion of a good search function that identifies all the records that need to be changed and with the ability to cut and paste the updated headings into each record, but nevertheless the process is inefficient and prone to human error."(6)

Unfortunately, DSpace -- the underlying software behind the RDSA -- is not capable of handling the thesaurus verification and automated batch processing we hoped to implement. The repository software includes provisions for developing local, hierarchical controlled vocabularies and/or links to external, persistent authority identifiers(7), but despite the much-appreciated efforts of the greater DSpace community -- such as the effort by Skourlas *et al.* to convert external ontology files into DSpace's Controlled Vocabulary file format(8) -- the use of hierarchical controlled vocabularies is largely aimed at self-deposited items, or other instances where digital objects are uploaded one-at-a-time. Similarly, editing metadata within the DSpace software is performed on individual records; while the software does include a batch metadata editing tool(9), the tool only exports a comma-separated value (CSV) files, which need to be edited by the user outside of DSpace and reuploaded to implement global or small-scale batch changes.

Phillips *et al.* from the University of North Texas Libraries developed a collection of command line tools(10) that were used to introduce name authorities into their institution's electronic theses and dissertations.(11) We considered adding some of these tools to the HI workflow, especially dc_breaker(12), a Python script designed to convert metadata harvested from OAI-PMH repositories into formats for use with other command line tools. However, we ultimately decided to use OpenRefine (formerly Google Refine, hereafter referred to simply as Refine). With its easy-to-learn graphic user interface, Refine's introduction into general usage for Fondren's RDSA projects not only provided a swath of possibilities in cleaning and normalizing data, but also enjoyed a warm reception from departments outside of Technical Services. Because of this, we decided to investigate Refine as a way to automatically match known headings in the HI collection, verify potential matches for incorrect headings, and alert us to headings that were accidentally left out of past HI thesaurus updates.

Refine, Reconciliation, and SKOS

Refine's Reconciliation service is defined in its Github wiki as "a web service that, given some text which is a name or label for something [...] returns a ranked list of potential entities

matching the criteria. The candidate text does not have to match each entity's official name perfectly, and that's the whole point of reconciliation -- to get from ambiguous text name to precisely identified entities."(13) When Reconciliation is run, Refine displays the results as a collection of "judgement" scores. Results scoring above a certain percentage threshold are automatically labeled as matches, while judgements below that threshold are returned as a short list of Best Guess matches (and an accuracy estimate for each) which the user may choose as a correct match, or create a "new" match for data not found within the returned precisely identified entities. In the world of linked data, those "potential entities" often have Uniform Resource Identifiers (URIs) behind them; with a few commands, Refine can extract these URIs from the match results and save them along with the text, the first step to creating linked data.

Standard reconciliation services are generally RESTful APIs -- that is, programs that facilitate the exchange, retrieval and deletion of data over regular internet protocols. However, with the addition of an optional extension(14), Refine can be adapted to reconcile data against locally-uploaded Resource Description Framework (RDF) files. As mentioned previously, the HI thesaurus began -- and to facilitate work in the Technical Services department, still exists -- as a spreadsheet on a departmental server. To use our local thesaurus in Reconciliation, our spreadsheet needed to be converted to RDF. There are several reasons to move data to an RDF-encoded environment for reconciliation; the most obvious is that our HI thesaurus does not exist with a RESTful API. On top of this, moving our thesaurus to a semantic data model pushes our data closer to linked open data, and also provides experience with the Semantic Web that was previously unavailable to us.

After researching several ontology languages, we decided to convert our thesaurus into a SKOS-formatted RDF-XML file. The Simple Knowledge Organization System (SKOS) is a W3C recommendation for semantically expressing thesauri, subject headings, and other lists of controlled vocabulary. SKOS can be used on its own, or in combination with more-formal (and more complicated) hierarchical languages, such as the Web Ontology Language (OWL). Several real-world examples of SKOS being used to facilitate library authority data exist, most notably the Library of Congress's Linked Data Vocabularies (http://id.loc.gov/) which deploys LC Authority data as SKOS. Other examples include UNESCO's multilingual thesaurus (http://skos.um.es/unescothes/) and the NASA Authorized Thesaurus (http://www.sti.nasa.gov/sti-tools/), which holds almost 20,000 subject terms as RDF/SKOS.

The fundamentals of SKOS are too long to be covered in a meaningful way and can be found in greater detail elsewhere, such as the W3C's SKOS System Primer(15), but it should be

noted that the fundamental element of a SKOS vocabulary is the *concept*. Concepts are the units of thought -- ideas, meanings, objects, events -- which form the bedrock of our knowledge organization systems. SKOS introduces the **skos:Concept** class to assert that a given resource (such as a subject heading) is a concept. This is done by

1) creating (or re-using) a URI to uniquely identify the concept, and
2) asserting in RDF that the resource identified by this URI is a **skos:Concept**.

Concepts then include labels -- expressions (such as a name) that are used to refer to the concept. The **skos:prefLabel** property assigns a “preferred” name to a resource, while **skos:altLabel** provides a secondary or alternative label. Below is a portion of LC’s SKOS RDF-XML authority record for John Lennon, reproducing pertinent SKOS entities in a simplified form:

<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:skos="http://www.w3.org/2004/02/skos/core#">
<rdf:Description rdf:about="http://id.loc.gov/authorities/names/n80017868">
<**skos:prefLabel** xml:lang="en">Lennon, John, 1940-1980</**skos:prefLabel**>
<**skos:altLabel** xml:lang="en">Lennon, J. (John), 1940-1980</**skos:altLabel**>
<**skos:altLabel** xml:lang="en">Lennon, John Ono, 1940-1980</**skos:altLabel**>
<**skos:altLabel** xml:lang="en">Lennon, John Winston, 1940-1980</**skos:altLabel**>
<**skos:altLabel** xml:lang="en">Ono Lennon, John, 1940-1980</**skos:altLabel**>
</rdf:RDF>

In this serialization (or data context), the 100 field of a MARC authority for a personal name is represented with the **prefLabel** element, while its 4XX See-From fields are found in **altLabel** elements.

Experiments

It is important to note that we used Refine’s Reconciliation service in an unintended way. Many of the how-to guides and screencasts demonstrating Refine’s Reconciliation service -- including van Hooland and Verborgh’s guide at freeyourmetadata.org(16) -- conclude by instructing the user to extract URIs (or URLs) from the reconciled data. Our intention, however, was not to find or extract URIs; apart from the Library of Congress headings compiled for our thesaurus, no URIs would exist for the majority of our data. Instead, we hoped to use the

Reconciliation service to automatically verify headings in our collection, and change any headings that did not match. For our purposes, URIs did not matter (or so it seemed).

The initial experiment was a proof-of-concept: a test RDF thesaurus was run against a small set of data to confirm that correct headings would automatically match to the thesaurus. Due to our limited interaction with the SKOS and RDF, the test RDF file was created using the Open Metadata Registry's (OMR) sandbox website.(17) Inside the sandbox, we uploaded a handful of locally-generated names and subjects from the HI Thesaurus and exported the resulting RDF file. In Refine, sample headings from the HI collection data were uploaded, along with some that were purposely misspelled or incorrectly entered.

The experiment worked: Refine successfully validated the correct headings and suggested correct headings for the misspelled. Unfortunately, the OMR offered no batch-upload mechanism to generate a fully-formed RDF file; entering more than 3,000 headings by hand was not an effective use of time or resources. This, along with the success of the proof-of-concept, presented us with two goals for further experimentation:

1) Develop a method to batch-convert headings into RDF, and
2) Use the RDF to verify and clean as close to 100 percent of the collection's existing headings as possible.

The first goal was achieved with a spreadsheet using CONCATENATE commands. Individual headings were wrapped with appropriate RDF-XML data and the resulting spreadsheet values were copied to Oxygen XML for formatting and validation. Once saved, the RDF file was uploaded to Refine and reconciled against collection data in pursuit of the second goal.

In the ensuing attempts at reaching the second goal, gradual tweaks to the process led to a 90 percent match of existing headings to thesaurus headings. Perplexingly, upon further inspection, the remaining 10 percent of unmatched headings were found to exist in the RDF file, but for some reason were missed by Refine during the reconciliation process.

The culprit turned out to be URIs. When the initial test RDF file was created, the OMR sandbox generated URIs for the headings, shown below in bold:

<skos:Concept **rdf:about="http://sandbox.metadataregistry.org/uri/TV/1007"** xml:lang="en">

<skos:prefLabel xml:lang="en">Brick, Shirley Eclipse, 1898-1929</skos:prefLabel>

<skos:altLabel xml:lang="en">Brick, Shirley E., 1898-1929</skos:altLabel>
</skos:Concept>

As stated previously, only a handful of SKOS Concepts were created in the OMR, since no batch-upload function was available. This meant that the vast majority of our local headings -- more than 99 percent -- would not have URIs. Instead, we defaulted to a single, blanket URI for all entities, a modification of the base URI generated by OMR:

<skos:Concept **rdf:about="http://sandbox.metadataregistry.org/uri/TV/"** xml:lang="en">
<skos:prefLabel xml:lang="en">Brick, Shirley Eclipse, 1898-1929</skos:prefLabel>
<skos:altLabel xml:lang="en">Brick, Shirley E., 1898-1929</skos:altLabel>
</skos:Concept>

We later learned that when a new RDF file uploaded into Refine, the full data of the file is converted and stored in the program as Turtle, another flavor of RDF. When each entity is given its own unique URI, the RDF extension correctly translates them into separate Turtle RDF entries:

<http://sandbox.metadataregistry.org/uri/TV/3118>
a skos:Concept ;
reg:status <http://metadataregistry.org/uri/RegStatus/1001> ;
skos:inScheme <http://sandbox.metadataregistry.org/uri/TV> ;
skos:prefLabel "Rice University--Presidents"@en .

<http://sandbox.metadataregistry.org/uri/TV/1022>
a skos:Concept ;
reg:status <http://metadataregistry.org/uri/RegStatus/1001> ;
skos:inScheme <http://sandbox.metadataregistry.org/uri/TV> ;
skos:prefLabel "Automobiles--Texas--Houston"@en .

<http://sandbox.metadataregistry.org/uri/TV/1019>
a skos:Concept ;
reg:status <http://metadataregistry.org/uri/RegStatus/1001> ;
skos:inScheme <http://sandbox.metadataregistry.org/uri/TV> ;

skos:prefLabel "American National Red Cross"@en . [..]

Unfortunately, using the same URI for every entity resulted in Refine mashing all headings from the uploaded RDF into a single Turtle entity:

<http://sandbox.metadataregistry.org/uri/TV/>

a skos:Concept ;

reg:status <http://metadataregistry.org/uri/RegStatus/1001> ;

skos:inScheme <http://sandbox.metadataregistry.org/uri/TV> ;

skos:prefLabel "Presidents"@en , "Automobiles--Texas--Houston"@en , "American National Red Cross"@en , "Presidents--Texas--Houston"@en , "American Red Cross"@en , "Presidents--Texas"@en , "Presidents--United States"@en , structures, etc."@en , "Bathrooms--Texas--Houston"@en , "Architecture--Texas--Houston"@en [...]

Unsurprisingly, this caused matching errors in the reconciliation process. To remedy our rookie mistake, we added unique (but otherwise meaningless) URIs to the CONCATENATE formula and regenerated the RDF file. When the new RDF file was reconciled against the collection data, 99.6 percent of the collection's headings matched the thesaurus headings. The remaining 0.4 percent were identified as headings that needed to be fixed or added to the thesaurus.

Implementation

The described experiments commenced in late Spring of 2015. Shortly after the process was developed, HI team members from Technical Services and Woodson Special Collections agreed to implement the process on a provisional basis. As of mid-2016, the reconciliation process was deemed enough of a success to be run on a monthly basis. The agreement also allows for periodic (provisionally, annual) review of the entire collection and thesaurus, to ensure both the collection and the thesaurus contain the same headings.

Each month, members of the Technical Services Copy Cataloging Unit work to enrich the metadata for a set of HI photos. After all enrichments are submitted to the Copy Cataloging Unit Team Lead, new additions to the thesaurus are sent to the Metadata Coordinator, who generates the valid SKOS RDF-XML entities:

Heading: **Brick, Shirley Eclipse, 1898-1929**

Auto-Incremented Dummy URI ID: **1402**

CONCATENATE Formula to generate URI:
=CONCATENATE("http://sandbox.metadataregistry.org/uri/TV/",[Dummy URI ID])

Result: **http://sandbox.metadataregistry.org/uri/TV/1402**

CONCATENATE Formula to Generate SKOS Entities:
=CONCATENATE("<skos:Concept rdf:about=","""",[Generated URI],""""," xml:lang=","""","en","""",""><skos:inScheme rdf:resource=","""","http://sandbox.metadataregistry.org/uri/TV","""","/><reg:status rdf:resource=","""","http://metadataregistry.org/uri/RegStatus/1001","""","/><skos:prefLabel xml:lang=","""","en","""",">",[Heading],"</skos:prefLabel></skos:Concept>")

Result:

```
<skos:Concept rdf:about="http://sandbox.metadataregistry.org/uri/TV/1402" xml:lang="en">
        <skos:inScheme rdf:resource="http://sandbox.metadataregistry.org/uri/TV"/>
        <reg:status rdf:resource="http://metadataregistry.org/uri/RegStatus/1001"/>
        <skos:prefLabel xml:lang="en">Brick, Shirley Eclipse, 1898-1929</skos:prefLabel>
</skos:Concept>
```

Once appended, the RDF file is saved on a departmental server, accessible by Technical Services, Special Collections, and Digital Scholarship. Alerted that the updated file is available, the Assistant Head of Special Collections then uploads the RDF file to their local Refine app and reconciles it against the enriched metadata submitted by the Copy Cataloging Unit. After all headings are verified as correct, the enriched metadata is overlaid on the HI collection in the RDSA.

Implications for Future Work

Despite positive reaction from team members, the procedure is not without caveats that prevent it from being a permanent solution. Most immediately, the Refine app will not auto-update the RDF file with headings found in the collection data but not the thesaurus. New headings must be manually added to the RDF file in Oxygen; the previous version of the RDF has to be removed from the Refine app and re-added each time the process is run.

The process of generating the RDF after each update -- from CONCATENATE commands in Excel to Oxygen to its upload into Refine -- also seems a bit convoluted at times. In fact, the RDF extension can generate a valid RDF file from an active Refine project by allowing the user to construct a template "RDF skeleton." This templating function has successfully generated a valid SKOS RDF-XML file from the thesaurus data; this method could eventually replace the CONCATENATE method of generating our HI RDF. Unfortunately, this deep-dive into the RDF extension highlighted a number of concerns -- namely, its longevity and ability to play well with the Refine app. The RDF extension was previously available for download from the Digital Enterprise Research Institute (part of the National University of Ireland) at **http://refine.deri.ie**, along with in-depth documentation and installation instructions. In 2013, the Digital Enterprise Research Institute was absorbed by the Insight Centre for Data Analytics; not long after, the website was largely deactivated, including the RDF extension pages. The authors of this article cannot confirm a successful installation of the latest version (v0.9) of the extension available from several Github repositories, raising concerns about its compatibility with future releases of the Refine app. (Note: the last version of the extension available from the DERI website was v0.8, which can be successfully installed in the stable-but-legacy v2.5 of the app, then known as Google Refine, as well as the latest release candidate for OpenRefine, v2.6-rc2. A working download link for v0.8 of the extension can be found in the References section of this article.)

Finally, our Reconciliation experiments may serve as a step in the right direction for Fondren's immediate authority control needs, but putting "dummy" URIs in an RDF file on a closed server does not make our unique, local metadata shareable with the rest of the world. In 2006, Tim Berners-Lee published his four "expectations of behavior" for the creation of linked data:

1. Use URIs as names for things;
2. Use HTTP URIs so that people can look up those names;
3. When someone looks up a URI, provide useful information using RDF and SPARQL; and
4. Include links to other URIs so people can discover more things.(18)

In the time since Berners-Lee's post, the call for library-led linked data creation has only increased. The W3C's 2011 Library Linked Data Incubator Group Final Report noted: "The official owners of resource data and standards should assign URIs as soon as possible, since

application developers and other users of such data will not delay their activities, but are more likely to assign URIs themselves, outside of the owning institution. When owners are not able to assign URIs in good time, they should seek partners for this work or delegate the assignment and maintenance of URIs to others in order to avoid the proliferation of URIs for the same thing and to encourage the re-use of URIs already assigned."(19)

Our project does not currently have a method to generate URIs for thesaurus headings -- or, more accurately, URIs that are meaningful and can be reached by people looking for them. After the HI thesaurus project was put into production in late 2015, the authors realized this necessity and began investigating the University of North Texas's Name App project(20) as a potential solution. Written in the Django web framework, the Name App is used to create and edit authorities metadata for people, organizations, and events. Potential benefits of the Name App include a public web portal allowing anyone (including our cataloging and special collections staff) to see or use data, as well as providing metadata fields to record alternate authority keys (such as ORCID data). Most attractively, the Name App automatically generates URIs for each authority record, providing the infrastructure needed to establish linked data.(21)

While the Name App is not yet a reality for Fondren, the authors have pressed onward to fulfill the fourth tent pole of Berners-Lee's expectations. Since the HI thesaurus blends valid LC authorities with local headings, we set out to enrich our RDF data with URIs for all valid LC headings. This was initially accomplished by using the Link Identifiers tool in MARCEdit(22), a program for creating and manipulating MARC Records. The Link Identifiers tool was designed to automatically generate URIs of LC and VIAF name and subject access points in bibliographic records, depositing the links in subfield-zero (ǂ0) fields. By creating dummy MARC records populated with HI name and subject headings, we were able to resolve a portion of the access points, export the URIs, and deposit them in our RDF file within **skos:exactMatch** elements. Later on, we verified our exported URIs (and filled in some gaps) using *lc-reconcile*(23), a Refine reconciliation service for LCNAF and LCSH headings through id.loc.gov searches. The discovery of *lc-reconcile* also allowed us to find the root LCSH headings that formed many of our subdivided local authorities, and add them to our RDF file in **skos:closeMatch** elements.

Ultimately, this enrichment is preparatory work for using the HI thesaurus as the basis for a Rice-centric thesaurus, hosted on the web as linked open data. Such a thesaurus would serve multiple functions in tracing the development of names from within the campus, as well as personal names not found in LCNAF. There is also an exciting potential for such a thesaurus to track traditional authority sources next to new forms of identification (e.g., VIAF, ORCID, ISNI, and even social media handles). The development of the Reconciliation process, along with the

ongoing maintenance of the HI thesaurus, are the first steps toward a more integrated conceptualization of authority work for both traditional and digital collections.

References

1. Robert Estep (2014), “How to Build a Local Thesaurus [Poster],” *DC-201414: International Conference on Dublin Core and Metadata Applications*, Oct. 8-11, 2014.
2. Jessica Milstead (2002), “Thesaurus Management Software,” *American Society for Indexing*, http://www.asindexing.org/about-indexing/thesauri/thesaurus-management-software/
3. Patricia Dragon (2009), “Name Authority Control in Local Digitization Projects and the Eastern North Carolina Postcard Collection,” *Library Resources & Technical Services*, 53:3, 185-196.
4. Dorothea Salo (2009), “Name Authority Control in Institutional Repositories,” *Cataloging & Classification Quarterly*, 47:3-4, 253.
5. Ibid, 261.
6. Dragon, 193.
7. https://wiki.duraspace.org/display/DSDOC5x/Authority+Control+of+Metadata+Values
8. C. Skourlas, A. Tsolakidis, P. Kakoulidis, and G. Giannakopoulos (2015), "DSpace and customized controlled vocabularies," *AIP Conference Proceedings* 1644, 105-112.
9. https://wiki.duraspace.org/display/DSDOC5x/Batch+Metadata+Editing
10. Mark Phillips (2013) “Metadata Analysis at the Command-Line,” *Code4Lib Journal*, 19, http://journal.code4lib.org/articles/7818
11. Laura Waugh, Hannah Tarver, Mark Phillips (2014), “Introducing Name Authority Into an ETD Collection,” *Library Management* 35:4-5, 271-283.
12. https://github.com/vphill/metadata_breakers/blob/master/dc_breaker.py
13. https://github.com/OpenRefine/OpenRefine/wiki/Reconciliation-Service-API
14. https://github.com/downloads/fadmaa/grefine-rdf-extension/grefine-rdf-extension-0.8.0.zip
15. https://www.w3.org/TR/2009/NOTE-skos-primer-20090818/
16. http://freeyourmetadata.org/reconciliation/
17. http://sandbox.metadataregistry.org/
18. https://www.w3.org/DesignIssues/LinkedData.html
19. https://www.w3.org/2005/Incubator/lld/XGR-lld-20111025/
20. https://github.com/unt-libraries/django-name
21. Waugh, Tarver, and Phillips: 273.
22. http://marcedit.reeset.net/downloads

23. https://github.com/cmh2166/lc-reconcile