

Latimer Core Workshop

Orientation and Exercises

Author: Matt Woodburn

Version: 1.1

Date: 2024-08-28

Prepared for the Latimer Core (LtC) Workshop (WKSH04) at SPNHC-TDWG 2024 in Okinawa, 2024-09-04 as part of NHM's work on data standards for the DiSSCo Transition project¹.

Latimer Core (LtC) information

To familiarise with the TDWG Latimer Core collection descriptions data standard, we suggest taking a look at the following online resources:

[Latimer Core wiki](#): General guidance for what the standard is, what the components are, and how to use it.

[TDWG Latimer Core homepage](#) (including the formal [term list](#), [quick reference guide](#), and other [resources](#)): The official reference for the standard's term definitions.

Aims of the material

This workshop material has been developed to provide people with a hands-on sandbox for working with structured Latimer Core data, trying out different collection description use cases with reference to a real-world example, and seeing what data outputs look like in common shared file formats. We anticipate that the material should continue to be available beyond the event and made available to a wider audience if found to be useful.

Airtable overview

What is Airtable?

Airtable (<https://airtable.com/>) is a cloud platform that functions like a spreadsheet-database hybrid. Users can create multiple databases (called 'bases' in Airtable), create tables within those bases, and add fields of defined data types, including formulae, to those tables. Relationships can also easily be configured within a base by adding lookup fields that reference other tables within the same base. However, like a spreadsheet, users can work directly with table data in a grid-like interface. The aim is to provide users with a

¹ DiSSCo Transition. Funding Source: HORIZON-INFRA-2023-DEV-01 | Grant Number: 101130121

spreadsheet-like feel for direct data management, while also providing richer database-like functionality (such as relationships and data types) through no-code configuration. Over this, Airtable provides functionality for configuring interfaces over the data (again without the use of code) for different user experiences.

Why are we using it for Latimer Core prototyping?

LtC is a fairly structured and relational standard, and its potential applications can therefore be quite difficult to explore and demonstrate using flattened structures like spreadsheets or CSVs. In Airtable however it's quite straightforward to configure bases that mirror the LtC standard and structure, using tables that represent the LtC classes, fields that represent the LtC properties, and lookup fields that represent the LtC data model relationships. In particular, it's very easy to configure many-to-many relationships in Airtable, of which there are a number in LtC (for example, a single ObjectGroup record may relate to multiple Taxon records, and a single Taxon record may relate to multiple ObjectGroup records).

Airtable's functionality for developing custom user interfaces also helps us to prototype and demonstrate ideas for how software might be developed to provide a cohesive user experience in working with structured LtC data. In addition, Airtable also provides an API that enables us to prototype the extraction of data in LtC-compliant, standardised formats (such as JSON and CSV).

Limitations

Airtable is commercial software, although with a free tier that anyone can sign up to without cost. The free tier includes a lot of the core functionality, but has restrictions in areas like the maximum number of records per base (1,000), number of editors, and API calls per month. Paid tiers have less restrictions and more functionality, but come with a monthly per-user licensing cost. Further details can be found at <https://airtable.com/pricing>. The free tier used for this demonstration should have everything needed for an LtC sandbox environment, but an upgrade may be required for LtC datasets of any significant size.

Latimer Core example bases

For this workshop, we've created two example LtC bases.

1. TDWG: Latimer Core Template

<https://airtable.com/appPnlzRZyJ5dkTdV/shr59oUbbXh1M0FJr>

This base is effectively an empty template (containing no data) with the data structure and interfaces configured to represent the full scope of the LtC standard. It's intended as a starting place that people can replicate (as many times as desired) to try out different LtC use cases.

- The base is configured to have a table for each LtC class, and within them a field for each LtC property. These include lookup fields for all 'relationship' properties in LtC

(the 'has...' properties, such as 'hasTaxon' or 'hasIdentifier'), in order to configure the relevant relationships between the tables in the base.

- Additional lookup tables have also been created for some properties (e.g., 'preservationMethod', 'baseTypeOfObjectGroup') where LtC recommends that users apply controlled vocabularies to those properties. This enables us to demonstrate interface features for controlled vocabulary maintenance through the demonstration interfaces.
- Interfaces have been configured that should allow users to browse, add, edit and delete data in any part of the data model.
- The base contains no data, so that it provides a clean template for users to enter and experiment with their own Latimer Core schemes.

2. TDWG: Latimer Core Digitisation Example

<https://airtable.com/apphqjDuC81ffdAiS/shrnqR2n9ZoxJ0noa>

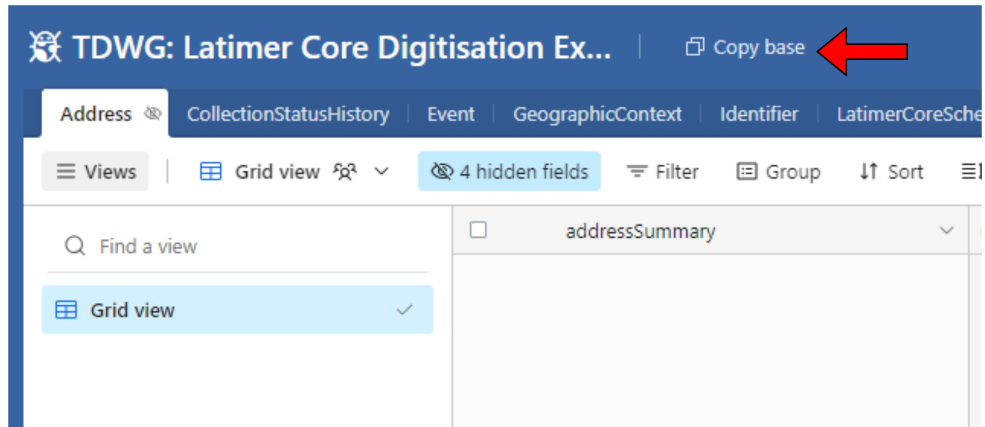
This base represents a real-world example of an LtC use case for planning and tracking the digitisation of the NHM London's British and Irish Coleoptera dry collection. It uses the same template as above, but with some modifications:

- The base contains example data, added as part of the initial stages of the NHM pilot.
- Some of the tables and properties have been hidden using table and field show/hide functionality. These represent LtC terms that aren't needed for this particular use case, and so have been hidden to simplify the user experience (although are still present in the data model behind the scenes).
- For the same reason, related elements of the custom user interfaces (pages and form elements) have also been hidden.

Getting started

The steps below should guide you through the process of creating an Airtable account (if required) and copying the example bases to your own workspace.

1. Click this link to go to the '[TDWG: Latimer Core Digitisation Example](https://airtable.com/apphqjDuC81ffdAiS/shrnqR2n9ZoxJ0noa)' base.
 - a. At the top of the screen, click "Copy base".



b. This will prompt you to sign in to Airtable.

- i. If you don't already have an Airtable login, you'll be prompted to sign up for a free Airtable account. Follow the process to create your account.

Note - if you don't wish to sign up for an Airtable account, you can explore the tables and data in the example bases in read-only mode.

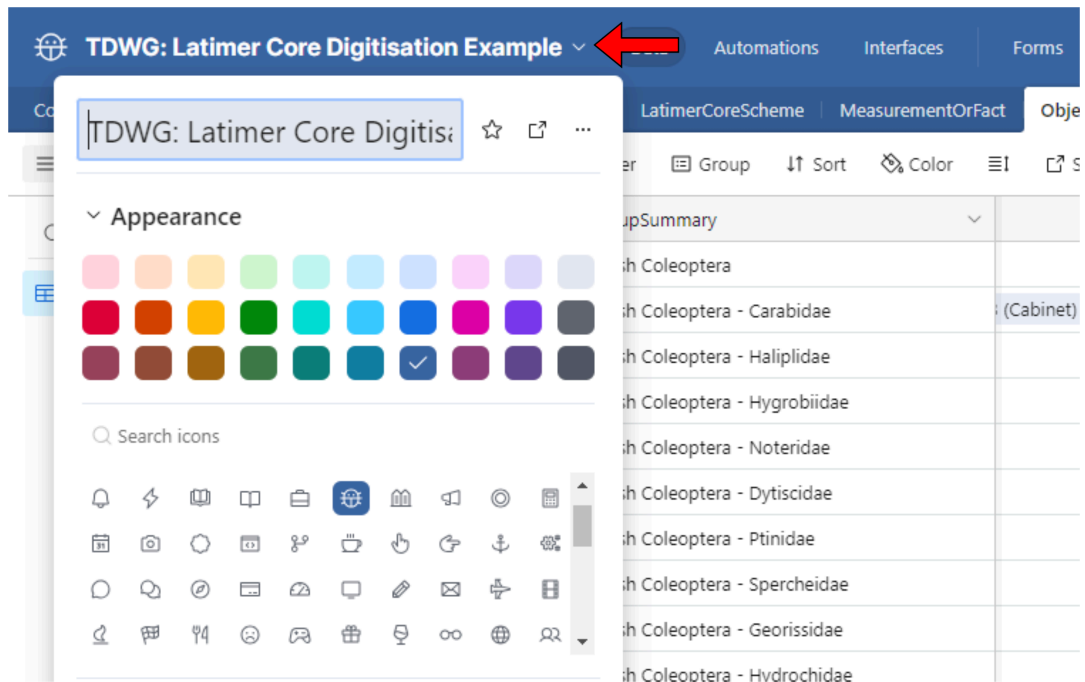
- ii. If you already have an account, sign in as usual.

c. Once logged into Airtable, a prompt should ask you which workspace you would like to copy the base into. If you already have Airtable workspaces, select whichever workspace you prefer. Otherwise, just click "Add base" to add it to your "My First Workspace" workspace.

2. Repeat steps a. and c. using this '[TDWG: Latimer Core Template](#)' link to copy the second LtC example base into your workspace.

Notes:

- These steps will result in your own copies of the bases and data, so no changes that you make will impact the originals or anyone else's copies.
- Once copied, it may help to rename or restyle (e.g. change the colour or icon) of your copies of the bases, to help to distinguish them from the shared templates. These options can be accessed by clicking on the down arrow next to the title of the base.

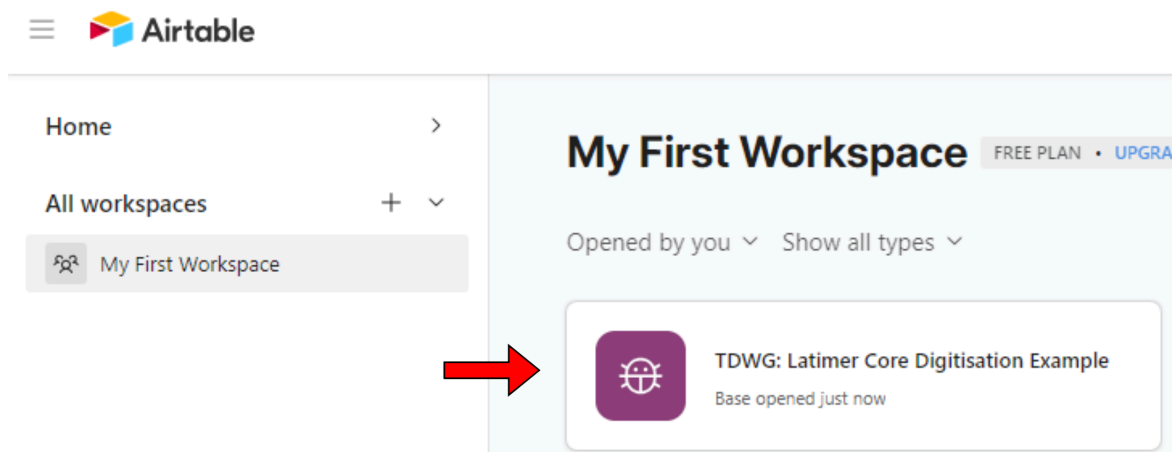


- Your copies of the bases will be private to your user by default. If you would like to share them with others, sharing options can be found by clicking the 'Share' button towards the top right of the screen.

Exploring the data and interfaces

Extensive guidance on how to use Airtable can be found at <https://support.airtable.com/docs/getting-started-with-airtable>. The guidance below covers some basic information to get you started.

To start exploring the LtC bases, click on one of them in your workspace home page.



Data view

The default view when you enter the base is the data view. This shows you each of the tables as a grid, and you can navigate between tables using the tabs at the top of the screen.

objectGroupSummary	MaterialEntity	collectionManagementSystem	collectionName
1 British and Irish Coleoptera	Axiell EMU 9.1	British and Irish Coleoptera	
2 British and Irish Coleoptera - Carabidae	Axiell EMU 9.1	British and Irish Coleoptera - Carabidae	
3 British and Irish Coleoptera - Halpiliidae	Axiell EMU 9.1	British and Irish Coleoptera - Halpiliidae	
4 British and Irish Coleoptera - Hygrobiidae	Axiell EMU 9.1	British and Irish Coleoptera - Hygrobiidae	
5 British and Irish Coleoptera - Noteridae	Axiell EMU 9.1	British and Irish Coleoptera - Noteridae	
6 British and Irish Coleoptera - Dytiscidae	Axiell EMU 9.1	British and Irish Coleoptera - Dytiscidae	
7 British and Irish Coleoptera - Ptinidae	Axiell EMU 9.1	British and Irish Coleoptera - Ptinidae	
8 British and Irish Coleoptera - Spercheidae	Axiell EMU 9.1	British and Irish Coleoptera - Spercheidae	
9 British and Irish Coleoptera - Georissidae	Axiell EMU 9.1	British and Irish Coleoptera - Georissidae	
10 British and Irish Coleoptera - Hydrochidae	Axiell EMU 9.1	British and Irish Coleoptera - Hydrochidae	
11 British and Irish Coleoptera - Helophoridae	Axiell EMU 9.1	British and Irish Coleoptera - Helophoridae	
12 British and Irish Coleoptera - Hydrophilidae	Axiell EMU 9.1	British and Irish Coleoptera - Hydrophilidae	
13 British and Irish Coleoptera - Histeridae	Axiell EMU 9.1	British and Irish Coleoptera - Histeridae	
14 British and Irish Coleoptera - Hydraenidae	Axiell EMU 9.1	British and Irish Coleoptera - Hydraenidae	
15 British and Irish Coleoptera - Ptilidae	Axiell EMU 9.1	British and Irish Coleoptera - Ptilidae	
16 British and Irish Coleoptera - Leiodidae	Axiell EMU 9.1	British and Irish Coleoptera - Leiodidae	

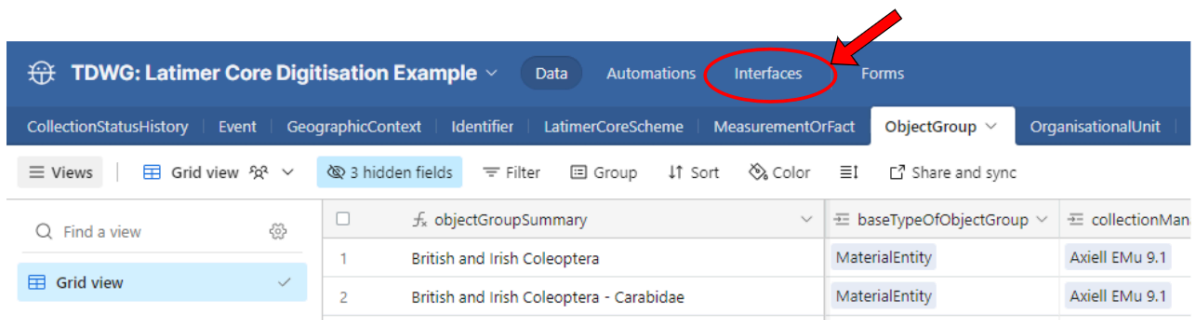
Data can be edited directly in the table grids, or you can click on the value in the first column to edit a record through a form view. You can type directly into a text or numeric field, while lookup fields will prompt you to select a record to link from a list (or add a new record to the related table). New records can be added by just clicking into the bottom empty row of a table grid, and records deleted by right-clicking on the relevant row and selecting 'Delete record'. There are also options for grouping, filtering and sorting the tables, as well as adding custom views.

Notes:

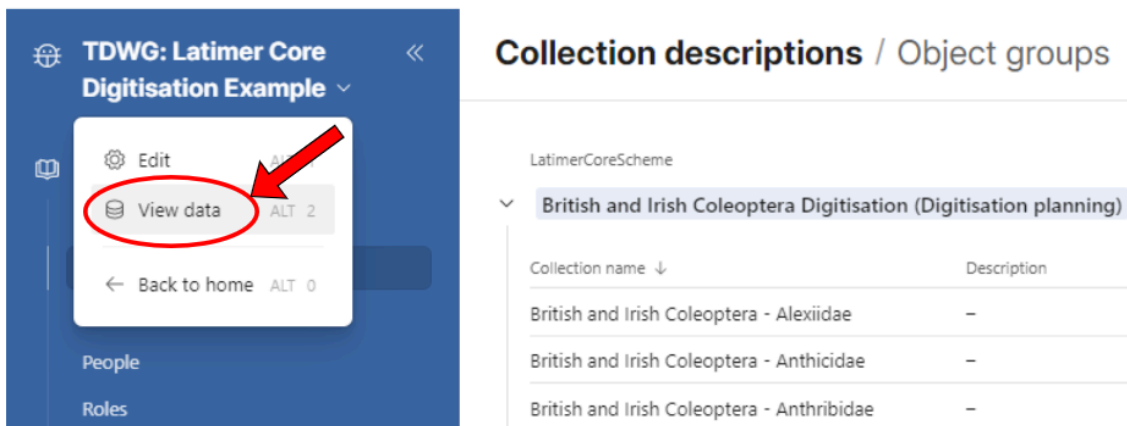
- The first two fields ('...Summary' and 'recordID') of each table are formula fields, which are auto-calculated so can't be edited. These fields are not part of the LtC standard, but are used to support the Airtable implementation.
- As these are your own copies of the bases, you have full control over the configuration. However, changes to table names and structure (e.g. adding/removing/renaming fields) would impact the LtC interfaces (see below) and data exports. If you'd like to explore more of the Airtable functionality but want to keep the example bases LtC-compliant, it's recommended that you create a new base or additional copies of the LtC example bases (via the home page) to experiment with.

Interface view

The LtC bases also come with example interfaces configured. These can be accessed from the data view by clicking 'Interfaces' in the top menu bar.



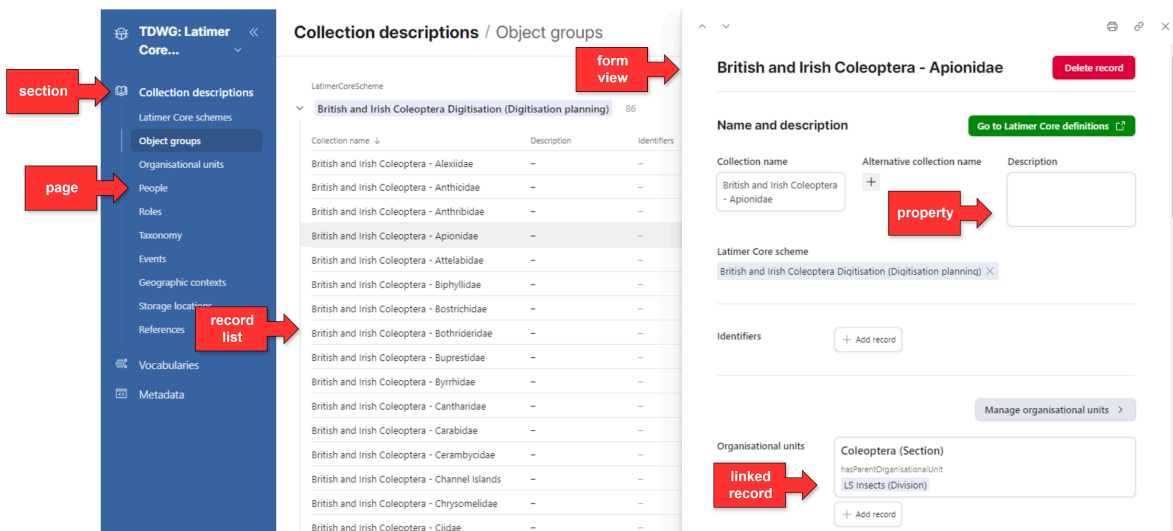
When in the 'Interface' view, you can return to the 'Data' view by clicking the base name in the top left corner and selecting 'View data'.



The example LtC interface has three main sections (shown in the menu on the left of the screen):

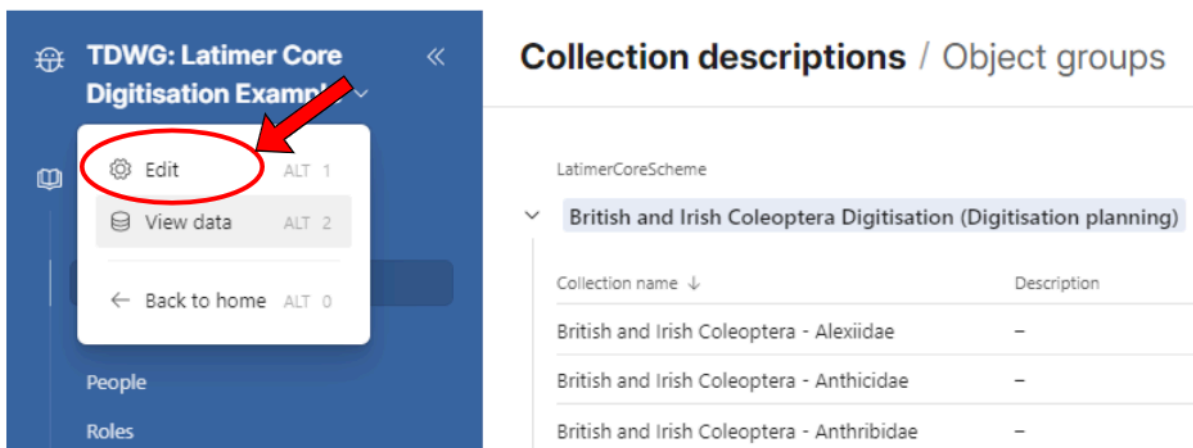
- **Collection descriptions:** Interfaces for viewing, searching and editing data in the main LtC classes.
- **Vocabularies:** Simple list interfaces for building up controlled vocabularies for relevant terms (e.g. preparationType, preservationMethod).
- **Metadata:** Interface for adding record-level metadata, and linking it to the relevant ObjectGroups.

The pages in the 'Collection descriptions' section each provide a list view, with a row for each record in the underlying table. Clicking into a row opens up a form view for that record.



The form view can be used to edit data, and navigate to additional forms representing linked records in other tables. Any changes to records made through the interfaces are immediately reflected in the underlying tables.

If you'd like to explore how the interfaces are configured, you can access 'edit' mode by clicking the base name in the top left corner and selecting 'Edit'.



Note: For technical reasons, the 'new record' forms that you see when adding records in the interface may only include a subset of the fields that are available for that record. If you create the basic record first using that form, and then click into the 'edit' form for that record, then all of the fields should be available for editing.

Creating Latimer Core schemes in the sandbox

There are three mandatory elements to an LtC dataset:

1. **LatimerCoreScheme:** There must be at least one record in the LatimerCoreScheme table, and it must have a 'schemeName' entered and 'isDistinctObjects' set to true or false.

2. **ObjectGroup**: there must be at least one record in the ObjectGroup table, it must be linked to a LatimerCoreScheme via the LatimerCoreScheme's 'hasObjectGroup' property, and it must have a value for 'baseTypeOfObjectGroup'.
3. **RecordLevel**: there must be at least one record in the RecordLevel table, which must have a value for 'license' and a related Identifier record (using the 'hasIdentifier' property) providing an identifier for the dataset.

Examples of these records can be found in the 'TDWG: Latimer Core Digitisation Example' base, by using the 'Data' view and exploring the 'LatimerCoreScheme', 'ObjectGroup' and 'RecordLevel' tables.

These elements are required for an LtC-compliant dataset, and are also needed for the data extract script (see below) to function correctly. The LtC standard does define some further mandatory terms when adding records for specific classes (see <https://ltc.tdwg.org/quick-reference/index.html>), and these would need to be included to create a fully compliant LtC dataset. It's not currently possible to set mandatory table fields in Airtable, so these constraints aren't automatically applied in the example bases.

Once these elements are in place, you're free to experiment with adding further data and trying out different ways of using Latimer Core. The [Latimer Core Wiki](#) provides further guidance for understanding the standard and designing Latimer Core schemes, and term definitions can be found in the [Terms page](#) and [Quick Reference Guide](#).

Example exercises

Below are some example exercises which may help you to work through the initial creation of the LtC dataset and introduce some of the different aspects of the data. It should be possible to work through these examples in either the 'Data' view or the 'Interface view', or even to swap between them to see what the different experiences are like.

Most of the steps involved have already been carried out already in the 'TDWG: Latimer Core Digitisation Example' base, so it may be helpful to refer to the data in there for examples if any steps are unclear.

Exercise 1: Creating a new Latimer Core Scheme

1. Create a LatimerCoreScheme record
 - Specify the scheme name
 - Specify whether it's allowed for the same individual object to be included in more than one ObjectGroup within the scheme ('isDistinctObjects' = false) or not ('isDistinctObjects' = true).

Note: If isDistinctObjects = true, we know that we should be able to safely add up MeasurementOrFact metrics (like object counts) for multiple object groups within the scheme without the risk of counting the same physical object more than once. For more details, visit the [LtC wiki](#).

2. Create an ObjectGroup record

- Specify the base type of the ObjectGroup
- 3. Link the ObjectGroup record to the LatimerCoreScheme (using the latter's 'hasObjectGroup' property)
- 4. Add a RecordLevel record
 - Specify a license
 - Create an Identifier record to represent a Latimer Core dataset identifier
- 5. Link the ObjectGroup record to the RecordLevel record (using the latter's 'hasObjectGroup' property)

Exercise 2: Building a basic collection description

1. Give your ObjectGroup a collection name
2. Add a brief description of the collection
3. Add some details of the nature of the objects (types of material, preservation methods etc) that the collection contains
4. Add some more information on the scope of the collection:
 - Add one or more Taxon records and link them to the ObjectGroup
 - Add one or more GeographicContext records and link them to the ObjectGroup
 - Add one or more GeologicalContext records and link them to the ObjectGroup
 - Add one or more EcologicalContext records and link them to the ObjectGroup
5. Add an OrganisationalUnit representing your institution, and link it to the ObjectGroup

Exercise 3: Annotating the collection with metrics and narratives

1. Add MeasurementOrFact records linked to the ObjectGroup to represent
 - a. the estimated number of objects in the collection, and
 - b. the percentage of objects that currently have a digital record in the collection management system.
2. Add a MeasurementOrFact record linked to the ObjectGroup to summarise the history of the collection
3. Add a MeasurementOrFact record linked to the ObjectGroup to provide a note about the collection's current state of housing conditions

Exercise 4: Incorporating people

1. Add yourself to People

2. Add yourself as the data creator to the RecordLevel metadata
3. Link yourself with your current job title to the OrganisationalUnit representing your institution
4. Add one or more curators to the ObjectGroup
5. Add one or more notable collectors to the ObjectGroup

Exercise 5: Adding references to external identifiers

1. For one of the Taxon records created earlier, look up the taxon in the Catalogue of Life (<https://www.catalogueoflife.org/>), and add that to the Taxon record as a linked Identifier record
2. Look up your institution in ROR (<https://ror.org/>) and add the identifier using an Identifier record linked to your institution's OrganisationalUnit
3. Add an Identifier record for your ORCID (if you have one) and link it to your Person record

Using Google Colab to download the LtC data in json and csv format

An experimental data extractor is under development that allows you to extract data from LtC-conformant Airtable bases in JSON and CSV formats. The aims of the script are:

- to explore and demonstrate how LtC data might be extracted from a source system/database in which they're stored (in this case, Airtable) and converted into formats that are suitable for sharing with other people (e.g. CSV) and/or other systems (e.g. JSON)
- to allow people to explore what LtC datasets that they work on in Airtable sandboxes look like in these other formats

The extractor is currently in the form of Python code in a Jupyter Notebook, and shared using Google Colab. This enables anyone (with a Google account) to view and run the script against their Airtable bases, and to take their own copies if they want to try out any edits to the script itself.

The tool is available to anyone that would like to try it out in its current form. For more information and access, visit the '[Airtable to Latimer Core Data Extractor](#)' documentation.