# Data Capture Exercise

**Start session** - leader (5 minutes)

**Darwin Core** - trainer (5 minutes)
 -choose some key points to focus on
 - remind of GBIF data quality requirements from the Darwin Core for data
 - GBIF EML profile for metadata
 (each dataset type has it's own set of requirements)
 https://www.gbif.org/data-quality-requirements
(perhaps use your own experience of learning about/using Darwin Core and/or going to TDWG conferences)

**Darwin Core Q&A** (15 minutes) trainer moderate

**Data Origins and Types** - trainer (5 minutes)
Many different data origins fit into the three dataset types that GBIF allows (ask if any one remembers what those three types are - have them type in chat)

Ask if the Data types picture quiz was hard for them? It definitely causes you to think and you realize that data can be published multiple ways...when you add the ability to use Darwin Core extensions into that you add more possibilities

Can use example of a checklist with associated occurrences?
That could be published as two different datasets. One occurrence dataset and one checklist dataset.
But the better way which allows the data to be linked in one dataset would be to publish a checklist and add the occurrence extension to publish the occurrences along with it in one dataset.

**Breakout discussion** (20 minutes)

Discuss
1. Is your data type different than you originally thought?
2. With what kind of data do you work?
3. How would you publish your data to GBIF (using which Core and/or extension)?

Select reporter. Tell the group which dataset types individuals in your group will be using to publish data.

Send to breakout discussion

**Reporting** (10 minutes)
trainer call on the different groups (Sophie translate if needed)

**Data capture exercise** demonstrate one record - trainer (15 minutes)

Discuss occurrenceIDs and the need to create one if one doesn't exist. GUIDs preferred. If GUIDs are created using tools on the internet, those GUIDs should be added permanently to the source data ~~dataset~~. If that is not an option for the project, discuss how to create a "triplet."
urn:catalog:institutionCode:collectionCode:catalogNumber
urn:catalog:FMNH:Herps:12345

**Breakout for exercise** (30)
https://docs.gbif.org/course-data-mobilization/en/exercise-2.html

Example of one completed record: USE CASE 1 - Exercise 2 - Completed template (Occurrences)

Mentors share images for specimens (see link above for downloads to the base data), have group pick out the data and data entry on their spreadsheets. Try to get through at least one more record (different that then example used in group)

**Conclusion**
Final thoughts?

**Homework** for Thursday