

# Fundamentals for Data Cleanup - Terminology

Facilitation Guide - 90 minutes

## Description

This module is the first in a two-part series. In this module, participants will learn important terminology that will give them a background on data, help them understand why data cleaning is important, and begin to address how to clean data. The terminology will cover terms related to data, data storage, data cleaning, standards, encoding, and more.

## Learning Objectives

Participants will...

- Become familiar with a wide range of data-related terms
- Learn how the terms usefully apply to data and the participants' work
- Gain the needed background to learn data cleaning in the next presentation

## Materials

Presentation

Pre and Post Assessment

*Optional:* Virtual Question Parking Lot (i.e. Jamboard)

## Facilitator Guide

*Welcome and Getting Started (Slides 1-2) - Introduce yourself and your role. Share the purpose of the session and discuss the Learning Objectives.*

5 Minutes

### Tips & Tricks

- Key Ideas to Highlight
  - This is going to be fun! You'll learn a lot.
- Resources to Share
  - Pre-Assessment

- Cover the learning objectives:
  - Become familiar with a wide range of data-related terms
  - Learn how the terms usefully apply to data and the participants' work

- Gain the needed background to learn data cleaning the next presentation

Take a few minutes to go through the pre-assessment and see what you know. If you're unsure/have questions about the content, those are great places to start when engaging with the presentation.

*Consider – Using a virtual tool like Jamboard to serve as a question parking lot for participants to share questions that come to mind throughout the module.*

*Cleanliness is King (Slides 3-8) - In this section, we'll cover important terms that describe the aims and parameters of data cleanliness.*

10 Minutes

Tips & Tricks

- Link learning the vocabulary to practical application and data cleaning

- Explain each of the terms on the slides as they relate to data cleaning.

*Let's Agree to Agree (Slides 9-15) - In this section, we'll cover the definition of standards and related terminology. We'll also go over their purpose and some examples of standards related to the Field Museum.*

10 Minutes

Tips & Tricks

- Link learning the vocabulary to practical application and data cleaning and storage

- Explain each of the terms on the slides as they relate to data cleaning/storage
- Guide participants through the DarwinCore example slides.

*A Rose by Any Other Name (Slides 16-26) - In this section, we'll cover a set of related terms and definitions and the distinctions between them.*

10 Minutes

Tips & Tricks

- Link learning the vocabulary to practical application and data storage

- Explain each of the terms on the slides as they relate to data.

- Focus on the differences between similar/related words.

*Planes, Trains, and Automobiles (Slides 28-61) - In this section, we'll cover vocabulary related to data storage and formatting.*

40 Minutes

Tips & Tricks

- Link learning the vocabulary to practical application, data storage, and formatting

- Explain the terminology on the slides
- Establish the overarching categories, such as field types, structural terms, and integrity, before focusing on the individual terms in each category.
- Most of the slides have examples to help demonstrate the concepts.

*Here be Dragons (Slides 62-71) - In this section, we'll cover the preliminary steps for data cleaning and related vocabulary, including the data relationships.*

15 Minutes

Tips & Tricks

- Resources to Share
  - Post Assessment

- Show the example data set and practice identifying the field types.
- The next slide shows an example of the recipient dataset, where the data from the previous slide needs to go, with the data relationships indicated under the new fields.
- Explain the different data relationships.
- Show the example mapped data and go over the considerations for data mapping.
- Data mapping is the first step in cleaning data; the next presentation will go over the following steps.
- **Questions?** - Check to see if your audience has any questions about the presentation. If used, refer back to the virtual question parking lot (i.e. Jamboard) to answer any last questions.
- **Thank you!** – Give yourself a round of applause! Share link to post-assessment for the final Check for Understanding.