# Dataset documentation guidance

This is the first draft of a reference document defining the **dataset documentation** to be used for “The Congruence Engine: Digital Tools for New Collections Based Industrial Histories ” TaNC-AHRC-funded project.

**The Congruence Engine: technical infrastructure and data management**

v.0.1 - October 2022

Authors: Anna-Maria Sichani, Arran Rees, Jamie Unwin

This document aims to provide documentation guidance to the Congruence Engine team for various datasets we are going to ingest and (re)use from data providers and project partners throughout the project.

## (Inspiration from [Datasheets for datasets](https://cacm.acm.org/magazines/2021/12/256932-datasheets-for-datasets/fulltext))

## **Dataset Name**

## **Description**

Size

Format

### 

### **Storage**

### 

### **Copyright**

### **Source URL**

### **Use cases - mini-investigations**

### **Linked datasets**

### **Potential datasets to use alongside this dataset**

### **Known History / things to be aware of**

### **Example projects using this data**

### **Known Biases**

### **Known History / Things to be aware of**

### **Is this a derivative or intermediate dataset**

Can it be regenerated from other datasets or sources elsewhere? If so how?

### **Life of dataset / Can or should it ever be deleted**

### **Owner(s)**

### **PID**

### **ORCID**