**Stages of DH Projects: Data and Collections**

Having explored the nature and history of the humanities, both traditional and digital, we now move into a focused study of how a digital humanities project is conceived, constructed, and executed.

Remember: in our course we will use simple, free, “no code” DH tools that allow us to easily analyze humanities data to reveal patterns and insights that traditional studies might not afford.

We review the concept of data, where to find and collect data, how humanists view data, and how we can begin to think of datasets that we might use for our own final projects.

**What are the Collection Stages of a DH Project?**

The word “data” is plural and means “pieces of information.” Datum is the singular form, but in common parlance, we use the noun data as either singular or plural.

Data are everywhere. It’s the humanists job to identify a dataset of interest, then collate, transform, study, analyze, interpret that data, and finally publish their results of that analysis.

Here are a few datasets:

* About yourself:

age, height, date of birth, place of birth, current and past addresses, color of eyes, color of hair.

* About the speeches of Dr. Martin Luther King, Jr.:

date, duration, unique words, order of words

* About the butterflies in your neighborhood:

genus, species, common name, wingspan, host plant, feeding preferences, colors, habitat, predators

* About your social network:

number of friends, type of relationship with friends, geographic location of friends, statistics of

your own posting behaviors

The first step in any DH project is the collation of a dataset of interest. Think about a humanities subject that interest you, and begin to compile data about it for possible use in your final DH project for this class.