

As noted in the Arches [Implementation Considerations](#), it is highly recommended that any adopters of the Arches Platform consider the utilization of controlled vocabularies to reinforce structure and consistency of terminology for data entry.

A controlled vocabulary is the preferred list of terms to describe proper names—including people, organizations, and companies— locations, objects, concepts, etc. Similar to a thesaurus, in which synonyms or related terms are grouped together, the controlled vocabulary acts as the authoritative list of terms that ensures

the most accurate word, spelling, and/or phrase is used for consistency and precision of indexing within a database. For example, there is a variation of terms that can be used to describe a *paintbrush*, referring to the tool used to apply paint to a large surface area. It can be spelled with a single word *paintbrush* or broken up into two words, *paint brush*. They might also be called *painting brushes* or could fall under the label *brushes, paint*. There are also many languages with their own words and phrases for this object, such as *schilderskwasten* (Dutch), *brosse* (French),

塗抹刷 (traditional Chinese), or *brocha* (Spanish). The controlled vocabulary will establish an authority record for this object, will indicate the preferred term and spelling, and indicate that any additionally listed terms and spellings refer to and identify the same object. The authority record also defines the preferred term to differentiate it from any similar, but distinct, objects; for example, *paintbrush* is purposefully distinct from *artists' brush*, due to the slight variation of purpose and usage situations for each brush. A controlled vocabulary is not restricted to objects; they can be used to describe geographic locations, concepts, people, organizations, and more. The primary intention is to organize data into predefined words and phrases that will ensure consistency and accuracy in records.

The benefits to using a controlled vocabulary is not only for indexing consistency, it also establishes the basis for powerful information search and retrieval for data stored within the database. These words and phrases can be tagged and labeled within the content of the database, and then organized in a way to facilitate the findability of these records. When a controlled vocabulary is structured within a hierarchy—for example, the record for *paintbrush* will fall structurally underneath the larger record for Tools and Equipment which, in turn, falls underneath the top-level record for Objects— this forms a *taxonomy*. A taxonomy establishes order and classification for individual records for information storage and retrieval. To move

Terms:

paintbrushes ([preferred](#), [C,U,English-P,D,U,PN](#))

paintbrush ([C,U,English,AD,U,SN](#))

paint brush ([C,U,English,AD,U,SN](#))

paint brushes ([C,U,English,UF,U,N](#))

brushes, paint ([C,U,English,UF,U,N](#))

painting brushes ([C,U,English,UF,U,N](#))

塗抹刷 ([C,U,Chinese \(traditional\)-P,D,U,U](#))

畫刷 ([C,U,Chinese \(traditional\),UF,U,U](#))

kwasten (schildersgereedschap) ([C,U,Dutch,UF,U,U](#))

schilderskwasten ([C,U,Dutch,UF,U,U](#))

brosses (paintbrushes) ([C,U,French-P,D,U,PN](#))

brosse (paintbrush) ([C,U,French,AD,U,SN](#))

















pinceau (paintbrush) ([C,U,French,AD,U,N](#))

pinceaux (paintbrushes) ([C,U,French,UF,U,N](#))

brochas ([C,U,Spanish-P,D,U,PN](#))

brocha ([C,U,Spanish,AD,U,SN](#))

one step further, an *ontology* establishes relationships between records that may not exist within a traditional tree-structured taxonomy. For example, in a taxonomy, the *paintbrush* record will fall underneath *Tools and Equipment* in the hierarchy; however, an ontology can

-  [Top of the AAT hierarchies](#)
-  [... Objects Facet](#)
-  [..... Furnishings and Equipment \(hierarchy name\)](#)
-  [..... Tools and Equipment \(hierarchy name\)](#)
-  [..... equipment](#)
-  [..... <equipment by process>](#)
-  [..... <surface working and finishing equipment>](#)
-  [..... <surface working and finishing tools and accessories>](#)
-  [..... brushes \(implements\)](#)
-  [..... artists' brushes \[N\]](#)
-  [..... <artists' brushes by form>](#)
-  [..... <artists' brushes by function>](#)
-  [..... <artists' brushes by material>](#)
-  [..... <brushes for personal use> \[N\]](#)
- [..... hairbrushes \(grooming tools\)](#)
- [..... nail brushes](#)
- [..... shaving brushes](#)
- [..... toothbrushes](#)
-  [..... <brushes: maintenance tools> \[N\]](#)
- [..... dusting brushes \(maintenance tools\)](#)
- [..... file brushes](#)
- [..... scrub brushes](#)
-  [..... paintbrushes](#)
- [..... fountain brushes](#)
- [..... wallpaper brushes](#)

establish a non-hierarchical relationship between *paintbrush* and the American painter Georgia O'Keeffe, who used a paintbrush for her work.

A controlled vocabulary can be created from scratch or adopted from another cultural institution, depending on the needs of the organization. There may be several reasons why an institution may choose to develop their own controlled vocabulary: the terms that will be included in the vocabulary are specific or unique to the institution; there is a small number of terms needed for the database; there is a need for control over the definition

or spelling of words to be included in the vocabulary.

Even if the vocabulary is built from scratch, there are several established controlled vocabularies created and maintained by cultural heritage institutions or other such professional organizations that can either be fully adopted, modified and adapted, or referred to for specific use cases within a wide range of data management systems, including the Arches Platform. For example, the Getty in Los Angeles creates and maintains Controlled Vocabularies under multiple topics: the *Art & Architecture Thesaurus* (AAT), which details objects, concepts, dates, materials, and more related to art and architecture, the *Union List of Artist Names* (ULAN), which details names of individuals and corporate bodies related to art and architecture, or the *Thesaurus of Geographic Names* (TGN), which details places and geographic sites, both modern and historical. Other examples of controlled vocabularies are the [Virtual International Authority File](#) (VIAF) and the [Library of Congress Subject Headings](#) (LCSH).

Vocabularies can be shared across systems, including Arches instances, or can be adapted depending on the needs of the project or organization. The decision of which controlled vocabulary to choose depends heavily on the type of data collected and the material an organization wants to describe; implementers are not restricted to the use of a single vocabulary but can pick and choose from a variety of sources.

Controlled vocabularies are managed on the Arches Platform within the [Reference Data Manager](#) (RDM) in conjunction with the [Arches Designer](#). A primer for using controlled vocabularies and the RDM on the Arches Platform is coming soon.

For further reading on controlled vocabularies, we recommend reviewing *Introduction to Controlled Vocabularies: Terminology for Art, Architecture, and Other Cultural Works* by Patricia Harpring (a free version is [available online here](#)).