

# Controlled Vocabularies

## *What is it? Why does WSDOT Need One?*

---

When we converse, we speak in “natural language.” This is language in all its raw, rich, ever-changing glory. But when we organize our information and label it, this natural language has so much variability that it can make it difficult to find information that we know exists. For example, what one engineer might record as Superpave another might record as asphalt or hot-mix asphalt. Similarly, chip seals and bituminous treatments can refer to the same thing. But if these items are entered into a database differently, relevant information may not be found if only one search term is used.

A controlled vocabulary imposes some order to ease the finding of the stored information, reduce the duplication of effort that arises because existing information can't be found, and increase the flow of information between offices that have divergent “professional” vocabularies yet need access to the same information sets in the performance of their work. In a controlled vocabulary, synonyms are identified and linked so information searches using one term can still find relevant information.

### **Takes the Guesswork out of Searching**

A controlled vocabulary makes a database easier to search. Since we have many different ways of describing concepts, drawing all of these terms together under a single word or phrase in a database makes searching the database more efficient as it eliminates guesswork. However, arriving at this efficiency requires consistency on the part of the individual indexing the database and the use of pre-determined terms.

### **A Familiar Concept**

Whether we realize it or not, we have all dealt with controlled vocabularies. Phonebook Yellow Page listings are arranged using controlled vocabulary. For example, a search for "Car Dealers" leads you to a note to “see Automobile Dealers.” At a basic level, this is how a controlled vocabulary system works, pointing users from commonly accepted “natural language” phrases to a phrase or phrases used in the controlled vocabulary.

### **One Search is All it Takes**

Conducting a search in a database that uses controlled vocabulary or indexing terms is efficient and precise. The biggest advantage to

### **Basic Elements of a Controlled Vocabulary**

A well-developed vocabulary is usually called a *Thesaurus*, and the terms which constitute the Thesaurus are usually listed in alphabetical order to facilitate finding them. Controlled vocabulary terms in a Thesaurus are usually listed in **boldface type** to distinguish them from certain uncontrolled terms that may also be included for the sake of pointing to the controlled term. Thesauri are usually hierarchical or faceted in structure, and will have the following basic elements:

- **TERM:** An accepted element in the controlled vocabulary
- **[DEFINITION/QUALIFIER]:** An optional descriptive paragraph that serves to either distinguish terms that are homographs or precisely define terms that may have multiple meanings.
- **RELATED TERM(S):** An element or elements also in the controlled vocabulary that an indexer may prefer to the one listed.
- **BROADER TERMS:** Elements in the controlled vocabulary that are higher in the hierarchy than the element listed.
- **NARROWER TERMS:** Elements in the controlled vocabulary that are lower in the hierarchy than the element listed.
- **USE:** This follows an uncontrolled listing, and points the indexer to an accepted term in the controlled vocabulary
- **USED FOR:** This is the reciprocal of the USE reference and accompanies the term to which the USE reference refers.

controlled vocabulary is that once you do find the correct term, most of the information you need is grouped together in one place, saving you the time of having to search under all of the other synonyms for that term.

## **The Role of Controlled Vocabulary at WSDOT**

In large organizations like WSDOT, controlled vocabularies could help improve technical communication. The use of controlled vocabulary ensures that everyone is using the same word to mean the same thing. This consistency of terms is one of the most important concepts in technical writing and knowledge management, where effort is expended to use the same word throughout a document or organization instead of slightly different ones to refer to the same thing.

The most effective communication occurs when all parties involved agree on the meaning of the terms being used. Consequently, finding the right words to index the content of a database is one of the most time-consuming and difficult aspects of developing it. Fortunately, a well-developed yet nascent controlled vocabulary for transportation information already exists in the form of the Transportation Research Thesaurus (TRT) developed by the National Transportation Library (<http://trt.trb.org/>). The TRT promises to be the best controlled vocabulary for indexing all WSDOT documents. An oversight committee has been established to review the addition of terms to the TRT as they are recommended.

Controlled vocabularies are employed in the WSDOT Library catalog (<http://www.secstate.wa.gov/library/catalog.aspx>) and the WSDOT Data Catalog. As well, effort is underway to improve indexing of web pages using controlled vocabularies to guide indexing. The need for controlled vocabularies in Document Management has also been discussed in the Document Management committee.

## **How Do We Begin?**

Because controlled vocabularies take some skill to apply, it is recommended that initial actions to develop the use of a controlled vocabulary be developed by a specialist (indexer or cataloger). This individual would build on the work already underway in the Department to develop a more robust controlled vocabulary capable of encompassing project plans, samples, photos, reports, and other relevant information. As well, the individual can provide advice on selecting terms from the TRT for use in designing the new Project Management & Reporting System.

It is recommended that this vocabulary be incorporated into agency information systems using tools such as drop down menus to standardize language.