

Database Content Best Practices

MSU Libraries DAWS Team

July, 2009

Overview

Database Content Best Practices is a document for those who are creating database content. It is a guide and reference to help reduce server errors, database content administration and development, and increase end-user and search-engine usability.

These best practices are focused on the types of content input into database fields, and adheres a variety of technical issues such as database structure and security, and related PHP calls to the database.

Special Characters and Character Strings

Across computer technology, there are several characters and string structures used in programming. These characters include ; " { } | \ / ? # and when used within a database field can cause either or both the database and webpage browser code to choke and not return the information being requested.

When a browser sends a request string of characters to a database and asks the database to retrieve any record with that string in a specified field, database security will stop the request if it sees a string it's been programmed to recognize as potential hacking. Character strings databases are programmed to stop are any with three or less characters such as 13c or L. or 4-D.

Dates

Databases have different date field formats which are designed to be used for the storage of different types of dates. For example: the date or date-range of the object's origination, or date it was scanned, or the date it was entered into the database. Therefore, it is important to know what date format is being used and for what purpose.

If the date is for an object's origination, a determination of what style of date needs to be decided before inputting record content. Types to be determined include: Will it be only a year? Will it have a season with a year? Will it have a day, month and year?

Because we are in the 21st century and many objects or citations are often from the 19th or 20th centuries, we can no longer use date abbreviations such as '89 or '00. This style of date has become ambiguous.

Additionally, it is important to remember the length of character strings and make sure they are not less than or equal to only 3-characters.

Not Advised	Best Practices
Su '88	Summer 1988
Ja-F '96	Jan.-Feb. 1996
S-O 1992	Sep.-Oct. 1992
leaving the field blank	Unknown

about 1900 or maybe 1900 or 1900?	ca. 1900 or circa 1900 or c. 1900
in the 1880's or 1900's	Late 19 th Century, Late 1800s or Early 1900s (with no apostrophe)

Keywords

Typically, the Keyword field has comma-delimited keywords and keyword phrases. Keywords are those that the end-user might use to find an object, and not necessarily the same as traditional library catalog topics (subject-headings), although subject-headings can be used as keywords.

Author(s) are not input as keywords unless the author is the subject of the object.

Not Advised	Best Practices
mix, tom, and oke, pok	Tom Mix, Pok Oke
13c, 2,4-D	carbon (13c), 2-4-D
bridges [water type] and rivers	water bridge, rivers

Abstracts, Notes, Summaries

These fields are the most forgiving with respect to database content. This is because they are "text" formatted fields with a larger number of allowable characters within the field and, when requested by the browser, the entire field is returned instead of a specific parsed string of characters from within a field.

Authors or Creators

To increase searching usability, multiple authors are input in the same field with the word "and" separating each author's name. This allows each name to be entered in the Last, Name and Last, I.D. format, while at the same time enabling author name searching, regardless of the number of authors.

Not Advised Multiple Name Fields	Best Practices Single Author(s) Field
using separate fields for individual authors within a group of authors using separate fields for Last and First	Last, First I. and Last, I. I. and Last, First

Titles

The Title field is typically programmed to be searched, along with abstract, keywords, and note fields, for a word or phrase requested by the Search or Advanced Search function on the webpage. Therefore, to increase the results back to an end-user's search request, a clearly defined and unambiguous descriptive title, without problematic characters is preferred.

Not Advised	Best Practices
Hineman Cold Spring Flour Mill of the 1880's & 90's; 5 miles West of Sheridan MT	Hineman Cold Spring Flour Mill 5 miles West of Sheridan, MT (1880s - 1890s)

Ice jam '00	Ice Jam, 1900
Equip 1	Bulldozers and Grading Machines in Riverbed Water (Equip 1)
Fam 1 or Family 1	Women and Children in Turn-of-Century Clothing by Pole Fence (Fam 1)
mr. & mrs. Owsley '89	Mr. and Mrs. Owsley 1889
Point of Rock, Looking West across the valley from the top of the hill.	Point of Rock Looking West Across Valley From Hill Top