Metadata in Drupal

Office of Environmental Information
Office of Information Analysis and Access
Information Access Division
Policy and Program Management Branch

Same 6 required fields

- Title ("DC.title")

- Description ("DC.description")

- Keywords ("keywords")

- Publisher ("DC.creator")

- Channel ("DC.Subject.epachannel")

- Type ("DC.type")

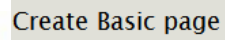# Metadata in Drupal

- 5* Required fields must be completed on the first screen

- Publisher is a drop down field that you select once

- Default for language is English

- Optional metadata - expiration date

- Additional metadata on other tabs

- At this time, you have to manually select the additional metadata in each facet by each term until our purchase of third party software is approved

- Once our purchase is approved, we will integrate third party software as we have already tested it in our sandbox

- This software will auto-suggest terms for all of the optional fields

Primary Metadata *

Taxonomy Topics

Taxonomy Facets

Create Basic page

Title *

Body *(Edit summary)

Text format Filtered HTML

More information about text formats

- Allowed HTML tags: <a> <abbr> <b> <blockquote>
 <caption> <cite> <code> <col> <colgroup> <dd> <div> <dl> <dt> <em> <figcaption> <figure> <h2> <h3> <h4> <h5> <h6> <hgroup> <hr> <i> <img> <ins> <mark> <p> <pre> <small> <span> <strong> <sub> <sup> <table> <tbody> <th> <td> <tfoot> <thead> <tr>
- Global tokens will be replaced with their respective token values (e.g. [site:name] or [current-page:title]).

☐ Use wide template
If your content is very wide, check this option to use the wide template.

PDFs in Drupal - being linked to

- Will be attached to a Document Node page

- This document page will have metadata on it

- You can put a group of similar PDFs on the same node page (50 states, report by years, etc) with good node metadata

- Our librarians may review the metadata

The searchable collection

- Is comprised of Non-linked html or pdf files **with good metadata** that are housed in Drupal

- These files can be housed in your topical area or in a searchable collection area of your choice (eg, OECA searchable collection)

- You are responsible for the files in your area(s)

- These files can be surfaced by Search

- These files can be part of our automated lists made by Drupal based on specific criteria

Questions?

Title: Metadata tag: "DC.title

Do:

- Include your most important and/or critical search terms in the title.

- Make your title understandable without any other context. Include the Area Name, if needed **for context**

- Match DC.title to the left-most portion of the HTML title (<title> tag). Only include the Area Name if it provides needed context, such as "Basic Information | Asbestos."

Do Not:

- Do not use acronyms unless you spell it out within the description or keyword fields.

Metadata tag: "DC.description"

Do:

- Write a short statement describing the content on your page, highlighting key concepts or issues.

    - "This page is…" is a poor description; "Water pollution is…" is much better.

- Include important search terms **not** already in your title.

- Check your first paragraph - you have probably already written a good description of your content, and can edit it further for brevity and/or to add in key search terms.

Do Not:

- Do not exactly copy the title into the description field.

- Do not use the same description for every page in a web area or TSSMS.

DC.Description examples

EPA

Examples:

	Description	Why?
Good	As part of efforts to improve preparedness and the ability to respond to terrorist attacks, EPA has been called upon to play a strategic role in homeland security	Provides context beyond the page title, includes key terms ("preparedness", "terrorist").
Bad	EPA Homeland Security	Repeats page title.
Good	Learn about carbon dioxide, methane, nitrous oxide, fluorinated gases, and how human activities add greenhouse gases to our atmosphere	Includes important search terms too long for the the title field.
Bad	This page is about greenhouse gas emissions.	Provides very little context, and almost no searchable terms.

Metadata tag: "keywords"

Do:

- Pull keywords from the **actual text of your content**.

    – Look for terms in your headings, table of contents, anchor links, introductory paragraphs, etc..

- Be selective with keywords. In most cases, **ten or less keywords** per document are sufficient.

- Separate keywords with a comma.

Do not:

- Do not repeat terms from the title or description.

- Do not create keywords for every possible combination of terms, or for capitalization, plurals, etc.

- Do not use the same keywords for entire set of pages in a web area or TSSMS.

- Do not use general terms, such as "EPA" and "environment".

Metadata tag: "DC.creator"

Do:

- Only list **the web-owner that *publishes* the document to the web** (your office, branch, lab, etc.), not the original author, or even the originating EPA office, branch, or lab.

- Follow this specific format:

  - US EPA, AA/RA, Real Owner (Lab/Center/Office/Division/Branch)

  - With the last and real owner spelled out.

  - The regional office name should always be spelled out

Do not:

- Do not list the original author, document creator, publisher, or even the originating EPA office, branch or lab.

Metadata tag:"DC.Subject.epachannel"

Do:

- Read the scope notes below to determine the channel that best fits your content.

- Find more information about <u>Channels in the Metadata FAQs</u> or view <u>EPA Channels in the Web Taxonomy</u>.

- Select at least one channel for every page. You can select multiple channels.

  - If your page has both scientific and regulatory content, apply both channels. If you have educational content, scientific content, and information about the Agency, apply three channels. Content that is specific to one channel should only be tagged for that channel.

  - If all four channels apply, you may want to re-think the content on your page.

Metadata tag: "DC.Subject.epachannel"

Do Not:

- Do not apply "Learn the Issues" to every page unless that content is actually appropriate. Content that is specific to "Laws & Regulations" should not also be tagged "Learn the Issues."

EPA Channels

Channel Name	Scope Notes
Laws & Regulations	Materials and content related to the legal and regulatory responsibilities and programs of the agency. Including, but not limited to, compliance and enforcement activities, guidance, regulatory development, permitting programs, etc.
Science & Technology	Materials, tools and content and related to the scientific, technical and research activities of the agency. Including, but not limited to, methods, models, research programs and plans, laboratories, software and databases, science products, etc.
Learn the Issues	Educational and consumer information as well as general or basic information related to all topics. Including, but not limited to, health and safety information, environmental emergency information and contacts, household management information (e.g. energy efficiency, recycling and waste reduction, chemical use and storage info, etc.), local information, etc.
About EPA	Information about the agency itself. Including, but not limited to, information about its leadership, its organization, its budget, its strategic plans, etc.