



magnolia®



Search Engine Optimization with Magnolia CMS

Best Practices and Built-in Support

A Magnolia Tech Brief

Table of Content

Introduction	3
Creating W3C Compliant Code	5
Providing a Sitemap	8
Enabling Secure Crawler Access	10
Creating Search-Engine Friendly URLs	13
Providing a Shallow Site Architecture	15
Inserting HTML Links	17
Preventing Broken Links	18
Categorizing Content	20
Redirecting Traffic	22
Optimizing Images	24
Managing Multi-Language and Multi-Regional Sites	27
Inserting Keywords in Page Titles and Description Meta Elements	30
Enhancing Your Site's Link Popularity	32
Attracting Visitors to Your Site	33
References	34
Learn More About Magnolia CMS	35

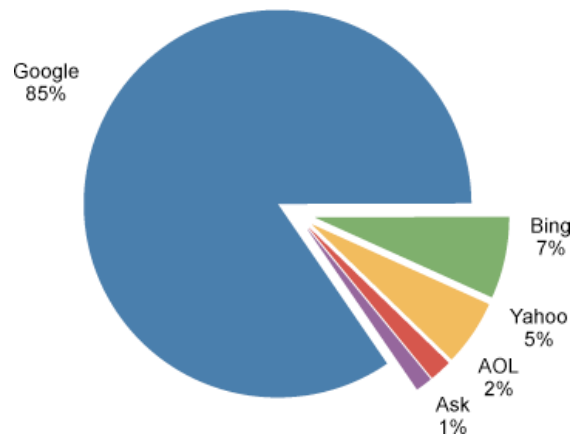
Introduction

This Tech Brief shows how to apply search engine optimization (SEO) to enterprise websites by employing best practices together with the built-in tools in Magnolia CMS. We take the key SEO factors identified by the [Moz.org \[1\]](#) consulting group, and pair them with various Magnolia features. Our intended audience is online marketers, people employing web communications, and anyone interested in increasing their site rank.

The purpose of SEO is to enhance the volume and quality of traffic directed to a website via search engine results. Search engines aim to provide users with the most relevant results, while webmasters attempt to tweak their sites in order to move it up in the rankings and attract profitable, organic traffic. SEO began in the early 1990s, and quickly degenerated into a tug of war between search engines and webmasters. For each new weighting factor introduced by the engines, crafty webmasters invented means to circumvent it. As a result, many undeserving sites achieved high rank at the expense of great sites providing valuable information. The Web became cluttered with content designed to satisfy the engines as opposed to users.

Today, SEO is practiced in a more cooperative environment in which search engines provide advice to webmasters. The engines will probably never fully disclose the algorithms that drive their crawlers but the fundamentals of SEO are firmly established.

Google is the dominant player, with more than 80% of the search market, followed by Bing, Yahoo, AOL and Ask, according to the findings of [statowl.com \[2\]](#):



Here we will show how Magnolia can help optimize your site with the following SEO best practices:

- Creating W3C compliant code
- Providing a sitemap
- Enabling secure crawler access
- Creating search-friendly URLs
- Providing a shallow site architecture
- Inserting HTML links
- Preventing broken links
- Categorizing content
- Redirecting traffic
- Optimizing images
- Managing multi-language and multi-regional sites
- Inserting keywords in page titles and description meta elements
- Enhancing your site's link popularity
- Attracting visitors to your site

Creating W3C Compliant Code

Best Practice

In order to crawl and index a website, a search engine has to be able to read and interpret its code. A page coded in compliance with standards developed by the World Wide Web Consortium ([W3C \[3\]](#)), is simpler for search engines to parse.

The [standards page at W3C \[4\]](#) lists more than 100 specifications, ranging from accessibility to XSLT. While this can seem overwhelming, four of the standards are key:

Standard	Magnolia Compliance
Valid code (HTML, XML, XHTML)	✓ Full
Semantically correct code	✓ Full
CSS	✓ Full
DOM and ECMAScript	✓ Full

1. Valid Code

The Magnolia Templating Kit is a best-practice framework that produces valid XHTML. The templates can be used to build compelling websites quickly and efficiently without the need for technical expertise.

The use of XHTML markup is declared with a DOCTYPE element, as you can see in the source code of any page created using a Magnolia template.

```

<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="en" lang="en"
class="js-on">
  <head>
    <style type="text/css"></style>
    <meta http-equiv="Content-Type" content="text/html; charset=utf-8">
    <meta name="keywords" content="Standard Article">
    <meta name="description" content="This is the abstract for the
article page "Standard Article". It is a brief résumé on the content
of this page. The abstract on the beginning of a page is also used
in teasers and openers that reference this page. Li European lingues
es membres del sam familie. Lor separat existentie es un myth. Por
scientie, musica, sport etc, litot Europa usa li sam vocabular. Li
lingues differe solmen in li grammatica, li pronounciation e li plu
commun vocabules. ">
    <meta name="author" content>
    <meta name="robots" content="all">
    <meta name="generator" content="Magnolia Enterprise Edition 4.4">
    <title>Demo Project - Standard Article</title>
    <link rel="stylesheet" type="text/css" href="/demo-project/

```

The page validates against the Document Type Definition (DTD) declared in the same element. The structure of a Magnolia page conforms to the document type by having an `html` element with the XHTML namespace declared, a `head` element including the `title` element, and a `body` element. All elements and attribute names are written in lower case and attribute values are enclosed in quotes. All non-empty elements (e.g. `p`, `li`) are properly terminated with a closing tag. All empty elements (e.g. `br`, `hr`, `img`) are properly terminated with a trailing slash (`
`).

2. Semantically correct code

Semantic markup describes the meaning of a page in terms of content, rather than the design.

In Magnolia templates, heading elements (`h1`, `h2`, `h3`, etc.) are used in the correct order without skipping levels, lists are used to list items, and tables are only used to present tabular data, not to create page layouts. Try removing CSS, JavaScript and images from your code. Is the code still understandable? Does it make sense to a reader? If yes, then it is semantically good code.

3. CSS

Cascading Style Sheets change the appearance of page elements by assigning styles to them. A style sheet gives a consistent appearance to an entire site. Semantic XHTML and CSS can improve your rankings, as it leads to better crawling, faster website response, and better accessibility and usability, which gives you higher conversions and increases your linking chances.

The look and feel of a Magnolia-based website is controlled by a single CSS file. Since the XHTML structure is designed to be styled with CSS, appearances can be changed without touching templates. Designers working with CSS don't necessarily need to know anything about Magnolia in order to create a compatible site design. They can start with the static HTML prototype that ships with the system, pick an HTML structure that is a close match to their original design, and then modify it further.

Magnolia CSS includes some browser-specific properties such as `-moz-border-radius` that are not part of the official CSS specification. Such properties are not used in areas critical to usability and fall back to standard properties gracefully.

4. DOM and ECMAScript

Document Object Model (DOM) provides a scripting language such as ECMAScript, the standardized version of JavaScript, easy access to the structure, content and presentation. DOM is future proof; it will allow any scripting language interact with the document. As long as you site has useful content, JavaScript code will play virtually no role in SEO optimization. The only thing that matters is where the code resides.

Magnolia uses jQuery, the most popular JavaScript library. jQuery is a cross-browser library designed to simplify client-side scripting of HTML. It supports DOM element selection, traversal and modification. The entire library is stored in single external JavaScript file rather than embedded in the page HTML. An external file saves bandwidth by reducing page length, yielding faster downloads.

	Title	Resource type
Templating Kit		
Themes	templating-kit	
Site Definitions	js	
Template Definitions	all	Processed Javascript
Paragraph Definitions	themes	
Dialog Definitions	sample-css	Processed CSS
Templates		
Resources		

Providing a Sitemap

Best practice

While most crawlers will find your site on their own, best practice is to keep them informed by providing an XML sitemap that lists your site's URLs. The document includes additional metadata about each page, such as when it was last updated, how often it changes, and its relative importance to other pages on your site.

The Sitemaps XML protocol (sitemaps.org [5]) is supported by all major search engines. You should include a reference to your sitemap's location in a robots.txt file.

Magnolia compliance

Magnolia's [Google Sitemap module](#) [6] generates your sitemap XML document automatically. It finds all the pages on your site and allows you to easily set priority and change frequency in the user interface. No XML editing is necessary. Default values are provided should you prefer to tweak just the critical pages.

Page title	URL	Changefreq	Priority
Home	/demo-project	yearly	1.0
About	/demo-project/about	weekly	0.5
Articles	/demo-project/about/subsection-articles	weekly	0.5
Standard Article	/demo-project/about/subsection-articles/article	weekly	0.5
Article with TOC	/demo-project/about/subsection-articles/large-article	weekly	0.5

The generated XML is compliant with the Sitemaps protocol so that it is ready for submission. If you manage multiple sites in the same Magnolia instance, you have the option to create a separate sitemap for each site.

Magnolia's Sitemap template is an additional component, unrelated to the XML sitemap. It allows you to include a sitemap page on your site without coding or configuration. It serves as a further navigational element, by offering an overview of content. This feature follows SEO best practices by assuring that each page is accessible by at least one static text link.

News and Events
➤ Events Overview
➤ New Year's Eve
➤ Another event
➤ News Overview
➤ Annual Report Presented
➤ Apprenticeship Training
➤ Another News

Enabling Secure Crawler Access

Best practice

While SEO is mostly about enticing that crawlers to your content, there are some pages you may not want them to access. These include administrative, password protected, and paid access pages.

A robots.txt file restricts crawler access on a site-wide basis. The file is typically placed in the root directory of your site. All legitimate crawlers check for the presence of a robots.txt file, and respect the directives it contains. Even if you do not want to restrict access, it is advisable to include a robots.txt file because the absence of one has been known to produce a 404 error, resulting in none of your pages being crawled. This can also be the case if the file is not constructed correctly. Specific crawlers can also be denied access by entering their names or IP addresses in the file.

The following example instructs all crawlers to keep out of the `private` and `temp` directories and denies access to a specified crawler.

```
User-agent: BadBot # replace 'BadBot' with the actual
user-agent of the bot
Disallow: /private/
Disallow: /temp/
```

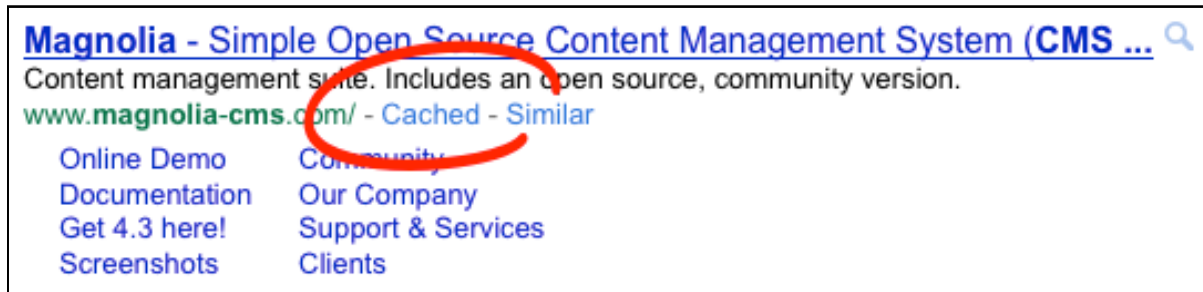
It is also possible to exclude individual pages with a robots meta element in the page HTML header.

```
<html>
<head>
<title>...</title>
<META NAME="ROBOTS" CONTENT="NOINDEX, NOFOLLOW">
</head>
```

Relevant values for the robots meta element are:

- NOINDEX prevents the page from being included in the index.
- NOFOLLOW prevents the crawler from following any links on the page. Note that this is different from the link-level `rel` attribute which prevents the crawler from following an individual link. For anti-spam reasons, it is generally better to use the `rel` attribute with value NOFOLLOW for individual links.

- NOARCHIVE instructs the crawler not to keep the page in the engine's public index. With Google, this prevents a cached copy of the page from being available in search results.



Preventing crawlers from following undesirable links can help combat automated spam. This method is endorsed by all major engines. Adding a `rel="nofollow"` attribute to a link allows control on an individual link basis. It can be used to reference external content while not endorsing it, and overcomes the necessity to exclude the crawler from all links on a page.

```
<a href="http://www.badsite.com" rel="nofollow">
```

Magnolia compliance

You can use all the above access restriction methods in Magnolia:

- Editing a robots.txt file is vastly simplified in that Magnolia produces clean, human-readable URLs such as `http://www.magnolia-cms.com/support-and-services`.
- At a single-page level, templates set the `robots` meta element in the HTML header to "all" by default, allowing full crawler access. Options to change this behavior can be offered to editors with minimal configuration directly in AdminCentral. Customization would involve adding a field in the Page Info dialog definition and a minor tweak to the paragraph template.

The image shows a screenshot of the 'Article Info' dialog box in Magnolia CMS, specifically the 'META DATA' tab. The dialog has a dark grey header with the title 'Article Info' and a close button. Below the header are four tabs: 'ARTICLE INFO', 'META DATA', 'OUTPUT CHANNELS', and 'SHOW ALL'. The 'META DATA' tab is active. It contains three main input areas: 'Meta Keywords' (a large empty text box), 'Meta Description' (a larger empty text box), and 'Robots' (a dropdown menu). The 'Robots' dropdown is currently open, showing a list of options: 'all', 'index, nofollow', 'noindex, follow', and 'noindex, nofollow'. The 'all' option is selected and highlighted. At the bottom right of the dialog, there are two buttons: 'CANCEL' and 'SAVE CHANGES'. The 'SAVE CHANGES' button is highlighted in green. A breadcrumb trail at the bottom left reads 'About > Articles > Stand...'. The overall interface is clean and modern, with a light grey background and dark text.

Creating Search-Engine Friendly URLs

Best practice

URLs are an important SEO element for search engines as well as users. Search engines are more likely to index static, short, clean URLs. Users appreciate their readability too. While most crawlers can follow dynamically generated URLs, there is evidence that when dynamic parameters are appended to a URL, the crawler only follows them to sites considered to have high importance. Request parameters should be avoided in the URL, as the crawler may view several URLs pointing to the same content as duplicates, lowering page rank.

Good

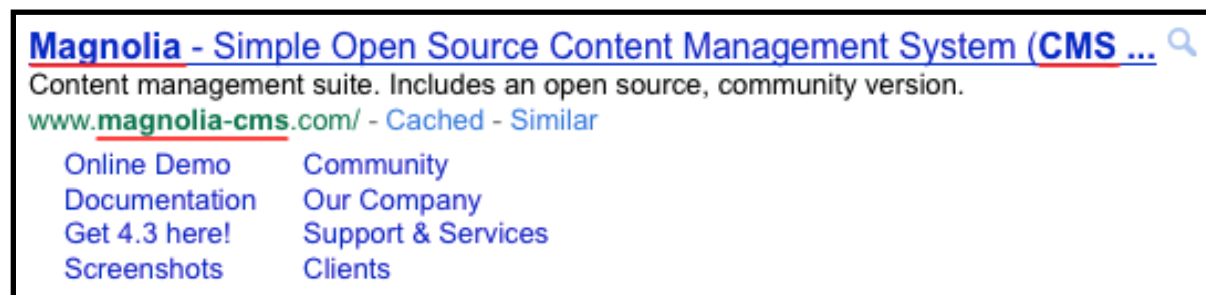
`http://www.yoursite.com/books/hardcover`

Bad

`http://www.yoursite.com?p=72&s=0`

Page URLs should:

- Include important keywords, as keywords are highlighted in search results. Most SEO experts agree that including them in a URL is important. Keywords also increase the likelihood of a visitor clicking through to your site.



- Have short, case-insensitive words, separated with hyphens.
- Be descriptive of page content. Both from the SEO and usability perspectives, an obvious URL is a good URL. URLs get pasted, shared and written down, so the more obvious the content, the better.
- Avoid the use of meaningless numbers and extraneous information.

Magnolia compliance

Magnolia produces clean, human-readable URLs by default. Dynamically generated URLs are converted to crawler-friendly static ones. Request parameters are not appended to URLs.



Some systems append session IDs to a URL. A session is typically needed when a user logs in. A user's shopping cart is an example of content that would be stored in a session. Appending session IDs to URLs is a security issue. A URL that contains an ID cannot be copied without the risk of inadvertently handing the session to another user.

Good

```
http://www.yoursite.com/cart
```

Bad

```
http://www.yoursite.com/cart?jsessionid=undn2i5014
```

Magnolia is session-less until the user logs in. At that point, a session ID is stored in a cookie instead of appending it to the URL. This avoids the duplicate content issue and makes URLs safe to copy and share.

When creating internal links, Magnolia uses the title of the target page as anchor text by default. The title typically includes the most relevant keywords about the target page.

Minimizing length and trailing slashes will make URLs easier to copy and paste, and fully visible in search results. A Magnolia URL is equally valid, regardless of whether you include an .html suffix or slash:

```
http://www.magnolia-cms.com/clients/references
```

```
http://www.magnolia-cms.com/clients/references.html
```

```
http://www.magnolia-cms.com/clients/references
```

Providing a Shallow Site Architecture

Best practice

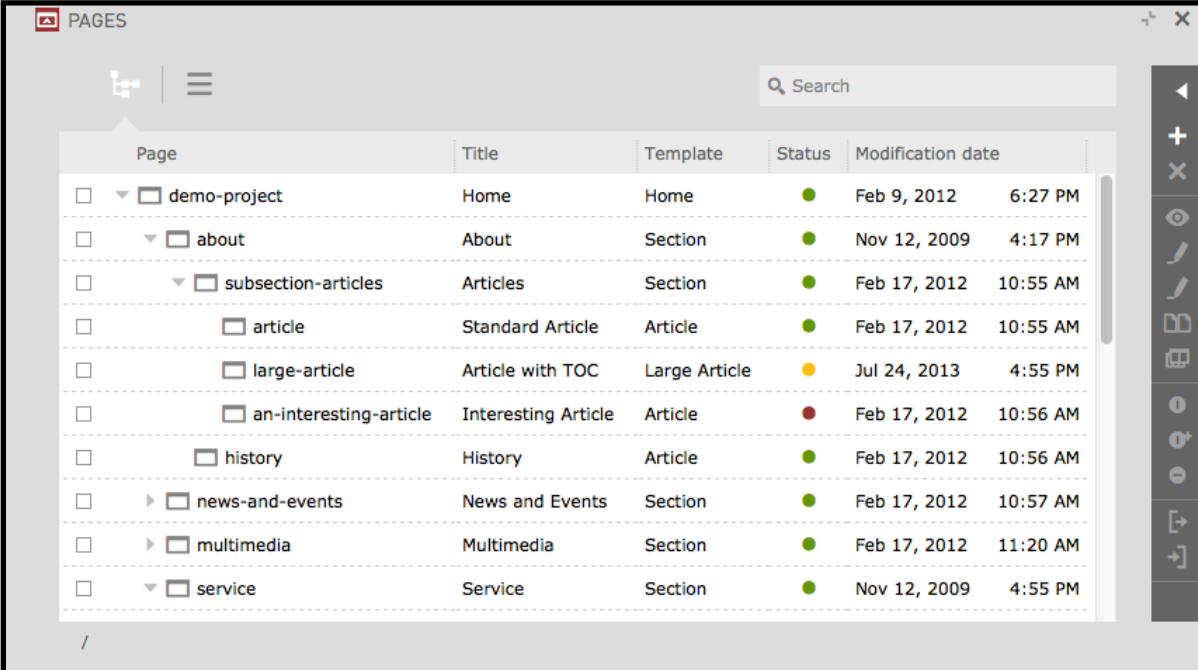
Site architecture is critical to SEO. Many sites grow organically and their architecture becomes a maze-like, illogical complex. A clear structure provides both the user and web crawler with a positive navigational experience.

As a rule of thumb, the flatter the structure, the better. The higher a page appears in the site structure, the more likely it will be ranked high by a crawler. Crawlers seem to determine the depth of a page by the number of directories present in the URL, giving more relevance, for example, to `/books/art` than to `/books/bestsellers/hardcover/art`. A one or two layer limit is recommended but for larger sites this is not always feasible.

It is also important to maintain consistency throughout your site. All pages should follow the same format, and once established, you should endeavor to stick to the URL guidelines. Where possible, avoid the use of sub-domains as crawlers may view them as separate from your main domain.

Magnolia compliance

Magnolia stores content in a hierarchical content repository. Editors see the hierarchy as a tree where each page belongs to a branch and has a parent node:



The screenshot shows the Magnolia CMS Pages interface. At the top, there is a search bar and a navigation menu. Below, a table displays a hierarchical tree structure of pages. The table has columns for Page, Title, Template, Status, and Modification date. The tree structure is as follows:

Page	Title	Template	Status	Modification date
demo-project	Home	Home	●	Feb 9, 2012 6:27 PM
about	About	Section	●	Nov 12, 2009 4:17 PM
subsection-articles	Articles	Section	●	Feb 17, 2012 10:55 AM
article	Standard Article	Article	●	Feb 17, 2012 10:55 AM
large-article	Article with TOC	Large Article	●	Jul 24, 2013 4:55 PM
an-interesting-article	Interesting Article	Article	●	Feb 17, 2012 10:56 AM
history	History	Article	●	Feb 17, 2012 10:56 AM
news-and-events	News and Events	Section	●	Feb 17, 2012 10:57 AM
multimedia	Multimedia	Section	●	Feb 17, 2012 11:20 AM
service	Service	Section	●	Nov 12, 2009 4:55 PM

Organizing a site as a hierarchy enforces consistent structure. Pages and branches can be rearranged via drag-and-drop.

Magnolia allows you to manage an unlimited number of sites in AdminCentral. Each site has its own site tree, to keep the hierarchy shallow. You can map unique domain names to each site, creating microsites for campaigns, for example. To a crawler, these appear either as unique sites (www.winter2010.com) or as parts of your domain (`winter2010.example.com`).

Sometimes it is difficult to avoid deeper structures. For example, you may want to create a campaign inside your main site at `/example/marketing/campaigns/winter2010`. Adhering to the shallow hierarchy principle is possible even in this case. Magnolia allows you to shorten the URL with URI mapping. Visitors and crawlers access the campaign site at `www.example.com/winter2010` while the content is served from the true location deeper in the hierarchy.

Inserting HTML Links

Best practice

Navigation links should be rendered as HTML to make them easy for crawlers to read. Search engines take the anchor text of a link into consideration when indexing it. Graphics and Flash objects, used by many designers to control look and feel, normally cannot be crawled. Alternative text descriptions should be provided. Use of uncrawlable elements also hinders users who do not have the required browser plug-ins.

Since the navigation system is a site-wide collection of links to topics, links should be consistent throughout the site. Navigation is also a perfect opportunity to include descriptive keywords.

Magnolia compliance

The Magnolia Templating Kit offers four navigation structures out-of-the-box: horizontal, vertical, breadcrumbs, and meta-navigation. All of them render links as HTML. Consistent navigation is enforced throughout the site using the same template.

The screenshot displays a Magnolia CMS website interface with four navigation structures highlighted on the left side:

- Horizontal:** A dark navigation bar with the Magnolia logo and tagline "Simple Open-Source Content Management". It includes a search bar with a "SEARCH" button and a menu with items: Home, About (highlighted), News and Events, Multimedia, Service, and Members Area. In the top right corner, there are links for "Contact", "Sitemap", and "Login".
- Breadcrumb:** A blue banner with the text: "This is the banner text for the 'Article' section. It is inherited by all child pages unless these are configured differently. A banner can be disabled in the template configuration settings." and "Section: Article". Below it, a breadcrumb trail reads: "You are here: Home > About > Articles > Standard Article".
- Vertical:** A yellow sidebar menu with the heading "Articles" and a list of items: Standard Article (highlighted), Article with TOC, and Interesting Article. Below this is a "History" section.
- Meta:** The main content area shows the title "Standard Article", a "Print" button, a "Bookmark" button, and categories: "Family Culture Weather".

Preventing Broken Links

Best practice

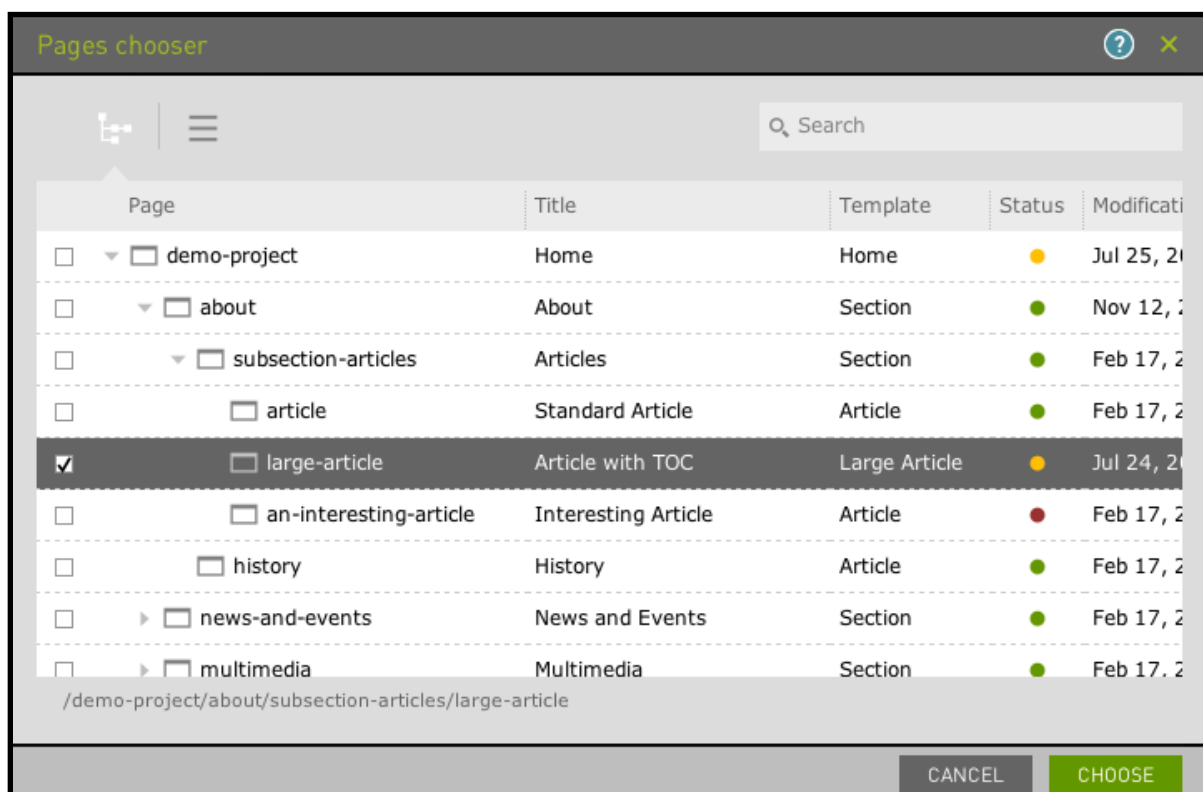
Broken links can lower your site's rankings. Google mentions this specifically in its [guidelines \[7\]](#). Broken links are also a source of irritation for users.

While you have little control over changes to external links, you can inadvertently break internal links in a variety of ways, such as when changing site structure or deleting content that has incoming links. A process should be in place to regularly validate all links on a site.

Magnolia compliance

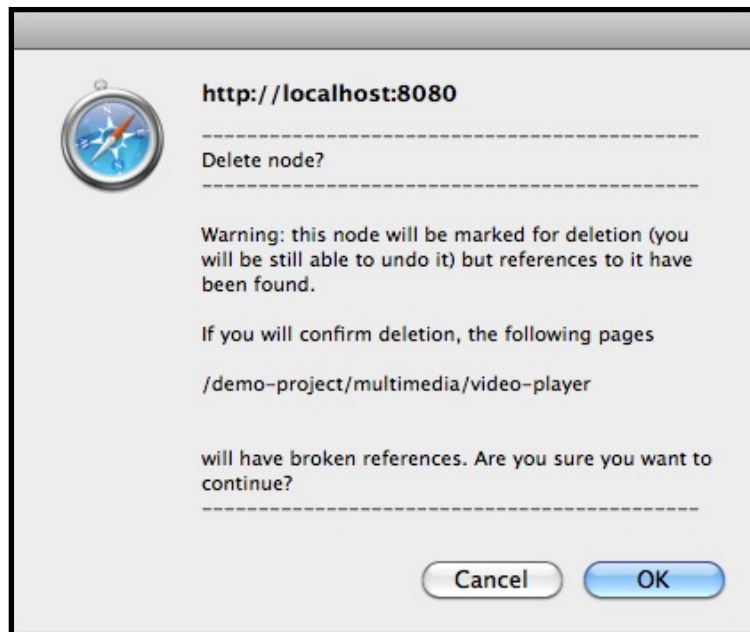
Magnolia has safeguards to prevent broken links:

- A hierarchical content store ensures that there are no orphaned pages that have no incoming links. All pages belong to a hierarchy, and each page has at least one incoming link - from its parent.
- Each page has a unique content ID for use as an internal link target. Dialogs provide a user-friendly browser for locating the target. Authors don't need to know the unique ID.

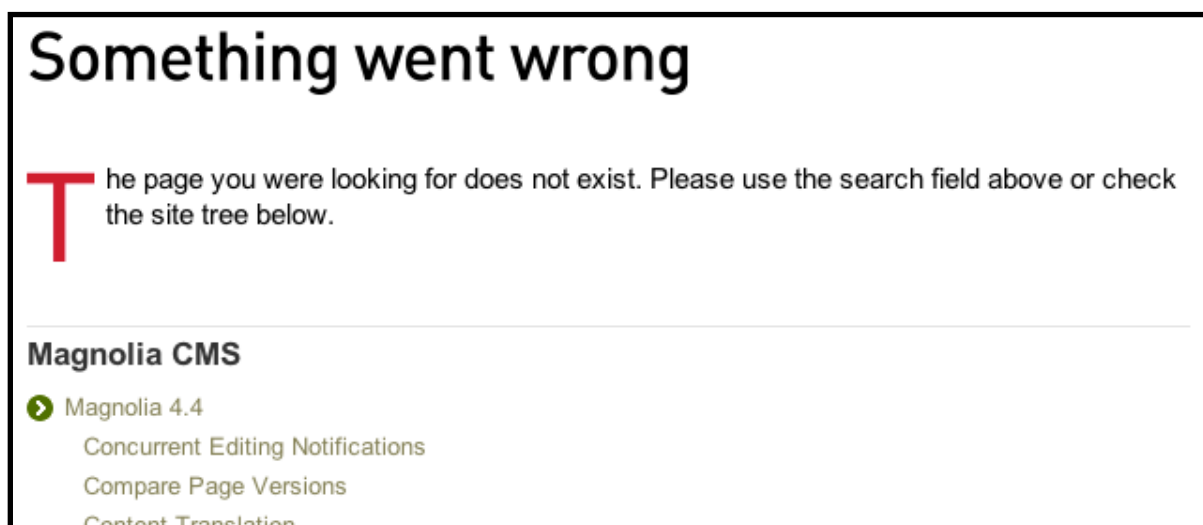


The unique ID stays the same even when the target page is moved or renamed. This allows for flexible site reorganization without breaking the links. The same mechanism is used when linking to documents in the document management system (DMS). Internal links to non-existing pages and documents will not be rendered, avoiding broken links.

- A reference view lists any incoming and outgoing links. A message on shows dependent content on deletion and deactivation.



- A custom 404 "not found" page can be displayed for broken incoming links. You can provide helpful alternatives such as sitemap tree or a search box to find the desired content:



Magnolia does not check for outgoing broken links but third party link checker tools for this purpose exist.

Categorizing Content

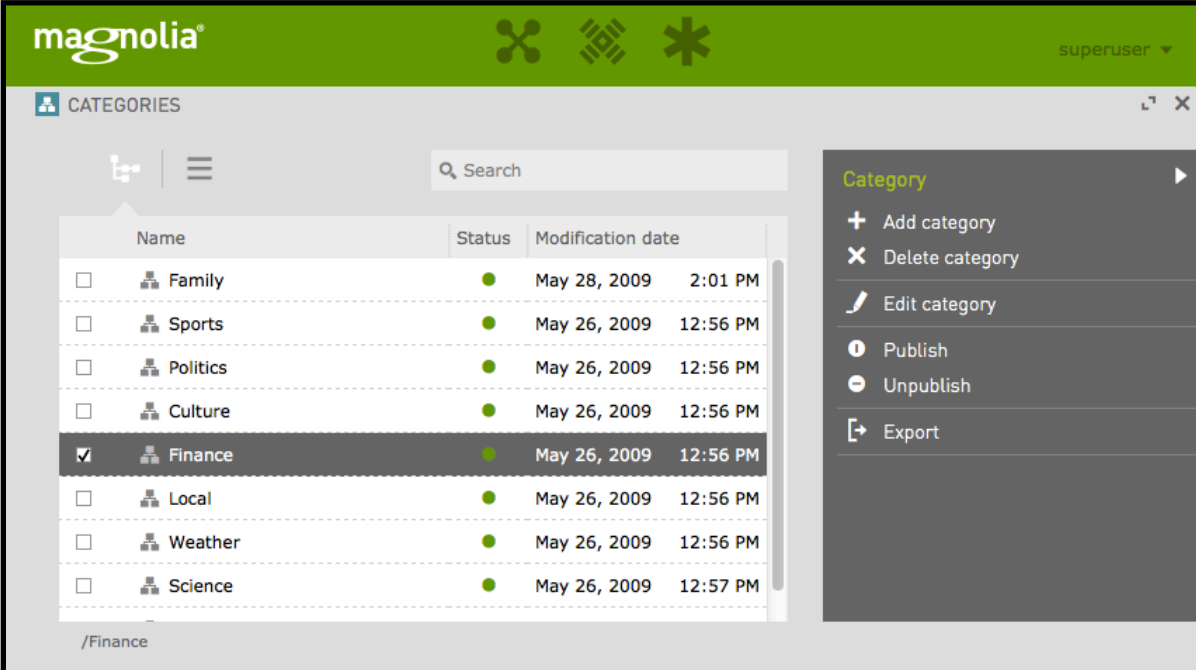
Best practice

Usability and SEO go hand-in-hand. Categorizing content allows visitors to find related and similar content with ease. Logical, keyword-rich categories provide intuitive navigation of a large collection of content. Categorization offers a range of opportunities to improve internal linking, visibility, and keyword density. Common uses of categories are tag clouds, and links to overview and landing pages.

Note! Although content tagging is an effective technique, tag clouds in particular have been abused. Bloggers saw them as a keyword-stuffing opportunity, and search engines have developed techniques to combat this trend.

Magnolia compliance

Magnolia allows you to create custom categorization taxonomies. You can tag articles with categories to support browsing of related content. This feature can be extended and adapted to suit your needs.





The screenshot displays the Magnolia CMS 'CATEGORIES' management interface. At the top, the Magnolia logo and user 'superuser' are visible. Below the header, there is a search bar and a table of categories. The table has columns for 'Name', 'Status', and 'Modification date'. The 'Finance' category is selected. To the right, a sidebar provides actions for the selected category: 'Add category', 'Delete category', 'Edit category', 'Publish', 'Unpublish', and 'Export'.


Name	Status	Modification date
<input type="checkbox"/> Family	●	May 28, 2009 2:01 PM
<input type="checkbox"/> Sports	●	May 26, 2009 12:56 PM
<input type="checkbox"/> Politics	●	May 26, 2009 12:56 PM
<input type="checkbox"/> Culture	●	May 26, 2009 12:56 PM
<input checked="" type="checkbox"/> Finance	●	May 26, 2009 12:56 PM
<input type="checkbox"/> Local	●	May 26, 2009 12:56 PM
<input type="checkbox"/> Weather	●	May 26, 2009 12:56 PM
<input type="checkbox"/> Science	●	May 26, 2009 12:57 PM

By default, you can categorize any page based on the Article template. Categories are displayed as related links in the article header:

Standard Article

 Print
  Bookmark

Categories: **Family** **Culture** **Weather**



Magnolia provides an editorial Category Cloud which allows editors to assign weights to categories. The more weight a category has, the bigger it shows in the cloud:

Hot Topics

Culture **Family** **Finance** International **Local** **Politics** Science Sports

The Category Overview template lists all articles within a category and suggests related categories:

Family

Standard Article

This is the abstract for the article page "Standard Article". It is a brief résumé on the ... [▶ read on](#)

Article with TOC

This is the abstract for the article page "A Large Article". It is a brief résumé on the content ... [▶ read on](#)

Related categories

[▶ Sports](#)

[▶ Local](#)

Redirecting Traffic

Best practice

Redirecting is the practice of sending requests to alternative destinations. Typical reasons to redirect are:

- Pages have moved due to site re-organization but search engines and users still have the old URLs in their indexes and bookmarks. Without redirection, traffic going to the moved pages is lost to 404 errors.
- Identical URLs serving the same content lead to duplicate content issues, as crawlers mistake each URL for a separate site. This is known as the "canonical URL issue" in SEO parlance. It is not uncommon to have `example.com`, `www.example.com`, and `example.com/index.html` deliver the same content. Sometimes there is a valid need to display the same content on multiple pages.

These factors adversely affect SEO but they can be corrected by redirecting the incoming request to an alternative URL. There are several ways to redirect but the only search engine friendly way is a 301 permanent redirect, which ensures that your hard-earned page rank is transferred to the new location. Contrast this with a temporary 302 redirect, which is disastrous for SEO, as page rank is not applied towards the new location.

Magnolia compliance

Redirects are cumbersome to implement on a Web server. Responsibility over Web server configuration often falls to another team and is out of reach for authors and editors, even for technical CMS team members. Fortunately, Magnolia allows you to configure redirects directly in AdminCentral.

A permanent 301 redirect is easy to do with [virtual URI mapping \[8\]](#). The request can be redirected to the same domain or another one. For example, when visitors request `winter2010.com`, you can redirect them to `example.com/marketing/campaigns/winter2010`, the actual location of the content. Short addresses are easier to market and remember.

virtualURIMapping	—
winter2013	—
class	info.magnolia.cms.beans.config.DefaultVirtualURIMapping
fromURI	/winter2013
toURI	permanent:/marketing/campaigns/winter2013

In addition, the Templating Kit includes a Redirect template that allows authors to set up a redirect page and define where the incoming request should go. Both internal and external pages can be identified as redirection targets:

Redirect ? X

REDIRECT

Path SELECT ANOTHER...

Page in Navigation Hide in Navigation

CANCEL
SAVE CHANGES

If you read more about the [canonical URL issue \[9\]](#) on SEO blogs, you will come across the "canonical URL tag". It is a `link` element in the page HTML head that tells crawlers that the page is actually a copy of the page given in the `href` attribute, and that page rank should be attributed to that page instead.

```
<link rel="canonical" href="http://www.example.com" />
```

The canonical URL is almost the same as 301 redirect in terms of SEO with the following exceptions:

- 301 redirect redirects all traffic (crawlers and visitors). The canonical meta element is just for crawlers.
- 301 is a stronger signal that multiple pages have a single, canonical source.
- 301 can redirect traffic to another domain, while still counting existing page rank towards the new location.

Canonical URL is sometimes favored over 301 because you don't need to ask developers to implement it. But as seen above, this is not an issue in Magnolia since editors can do their own virtual URI mapping in AdminCentral.

Optimizing Images

Best practice

Great images enhance your site visually and enhance SEO rankings, provided they are properly optimized. Google and other search engines offer image-specific searches. A high rank in these searches increases your visibility. Search engines cannot read the content of images, and therefore must rely on accompanying text and tags.

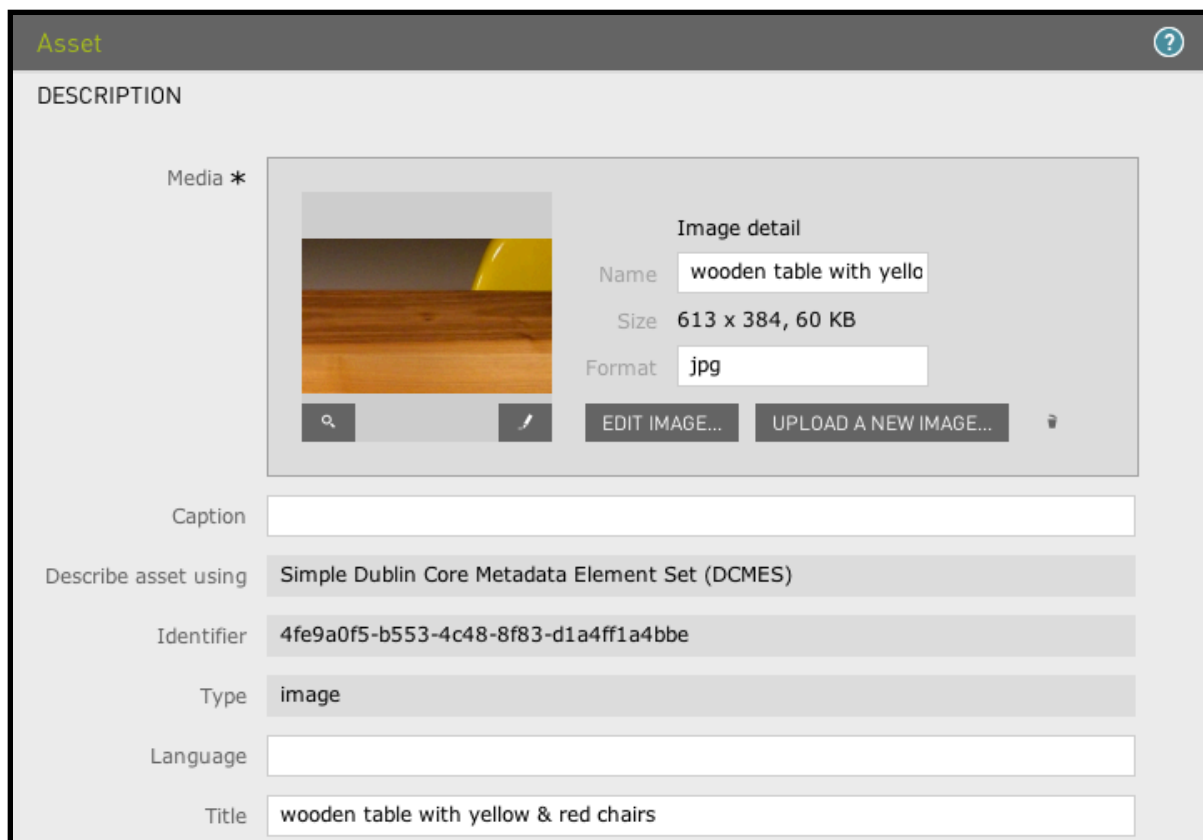
Best practices include:

- Image alt attributes with a short description of the image, for example ``.
- Keyword-rich image file names. `yellow-roses.jpg` is far more descriptive and relevant than `img003.jpg`.
- Size and quality optimized for quick loading. Crisp, sharp images look better in thumbnails displayed in search results, and are more likely to be linked to by other webmasters.
- Captions that aptly describe the image assist vision-impaired users.
- Width and height specified in the element. Many browsers render a page before images are downloaded. When dimensions are provided, other page elements wrap around them correctly, improving the initial user experience.

Magnolia compliance

Images can be stored in Magnolia's built-in document management system or uploaded at editing time. When the DMS is used, a two-phase failsafe mechanism ensures that alt tags, appropriate file names, and captions are included:

1. At upload time, details such as image source, copyright, subject line and size are captured. The fields provided for this purpose in the DMS are based on the [Dublin Core Metadata Standard \[10\]](#). If the original file name contains spaces, Magnolia replaces them with hyphens. In the event that an editor fails to include the information requested in the dialog, the source and subject-line are used as fallback.



The screenshot displays the 'Asset' editor interface. At the top, the word 'Asset' is in green, and a help icon is in the top right. Below is a 'DESCRIPTION' section. A 'Media *' field contains a thumbnail of a wooden table with a yellow chair. To the right of the thumbnail is an 'Image detail' section with the following fields: 'Name' (wooden table with yello), 'Size' (613 x 384, 60 KB), and 'Format' (jpg). Below these are buttons for 'EDIT IMAGE...', 'UPLOAD A NEW IMAGE...', and a trash icon. Further down are fields for 'Caption', 'Describe asset using' (Simple Dublin Core Metadata Element Set (DCMES)), 'Identifier' (4fe9a0f5-b553-4c48-8f83-d1a4ff1a4bbe), 'Type' (image), 'Language', and 'Title' (wooden table with yellow & red chairs).

2. When inserting an image to a page, the information read from the DMS can be changed once again. If the fields are configured as mandatory, an editor cannot insert an image without providing suitable alt attribute and caption. This enforces best practice compliance.

As a result, the page template renders crawler-accessible HTML code:



Large wooden table from above
Magnolia International Ltd.

The main part of the news starts here.

Li European lingues es membres del sam familie. Lor separat existentie es un myth. Por scientie, musica, sport etc, litot Europa usa li sam vocabular. Li lingues differe solmen in li grammatica, li pronunciation e li plu commun vocabules. Omnicos directe al desirabilite de un nov lingua franca: On refusa continuar payar custosi traductores. At solmen va esser necessari far uniform grammatica,

```

Resources Scripts Timeline Profiles Storage Console
▼ <dl class="photo">
  ▼ <dt>
    
  </dt>
  ► <dd class="zoom">...</dd>
  <dd class="caption">Large wooden table from above</dd>
  <dd class="copyright">Magnolia International Ltd.</dd>
  <dd class="longdesc">Image Description</dd>

```

Magnolia Imaging module ensures that high-quality, web-optimized images are used uniformly throughout the site. The module generates multiple variants from a single source image, resulting in considerable time-savings for editors. Sizing and cropping happens automatically, based on CSS rules. Built-in filters that scale and crop images can also be used for effects such as text overlays, which allow crawlers to read image text. It is possible to configure a standard output format and resolution for all images, and control the behavior of images in each area of a template.

Managing Multi-Language and Multi-Regional Sites

Best practice

Multi-language sites offer content in multiple languages, and allow you to extend your reach to a global audience. Estimates are that 73% of web users do not speak fluent English. Regional sites target specific geographic locations, and normally have a country-specific domains such as `.de`, `.co`, or `.uk`. For multilingual sites, best practice is to ensure that your content appears in search results for the appropriate language:

- Use the targeted language consistently throughout the page. Render all content and navigation in that language.
- Serve language-specific content at a separate URL. Use Unicode encoding.
- Avoid cookies to display translated versions.
- Avoid automated redirection based on the user's perceived language, as this could prevent a crawler from viewing all language variations.

For multi-regional sites, best practice is to ensure that your site appears in search results for the targeted country:

- Use a country-code top-level domain name such as `.de` or `.cn`. This is a strong indicator to users and search engines that the site is intended for a specific country
- Include local contact details, currency, and language in the site's content.
- Host the site on a server within the targeted country.

Magnolia compliance

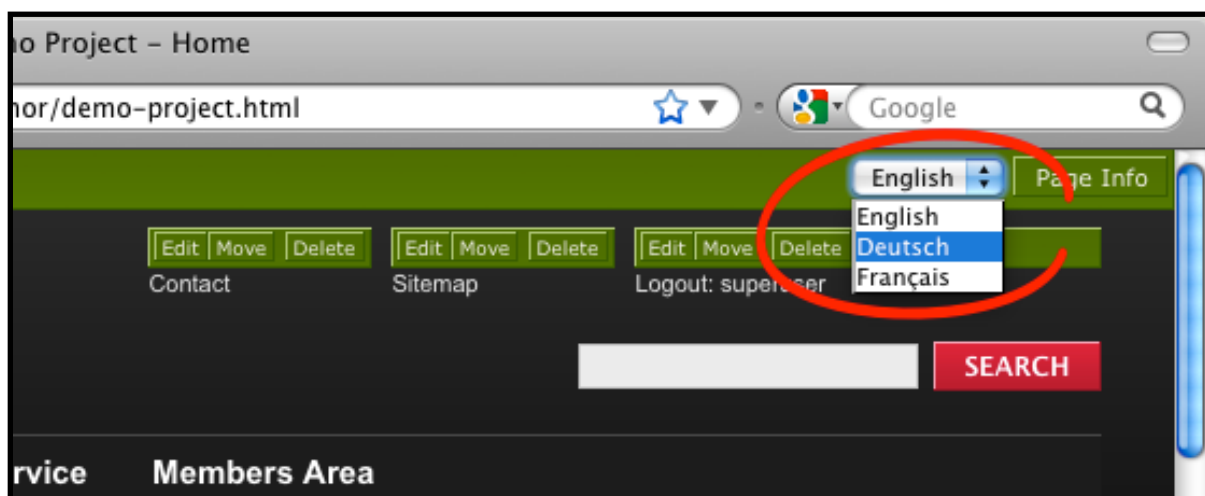
You can manage multi-language and multi-regional sites in a single Magnolia instance, while adhering to all best practices. A multisite configuration ensures that language is used consistently across multilingual sites. Redirection based on user's perceived language is not used.

Individual sites are configured in site definitions. In the example below, `demo-project` is a multi-lingual site that serves content in English and German, while `demo-project-de` is a site targeted at a German speaking audience only. Internationalization (i18n) configuration accommodates multiple languages. Both of

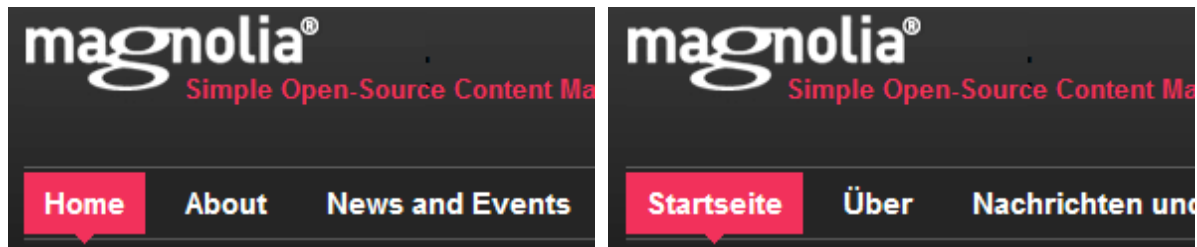
these sites [extend \[11\]](#) the default site, inheriting common elements, facilitating rapid site deployment.

▼	demo-project	—
▶	domains	—
▼	i18n	—
▼	locales	—
	en	—
	de	—
◆	class	info.magnolia.cms.i18n.DefaultI18nContentSupport
◆	defaultLocale	en
◆	enabled	true
◆	fallbackLocale	en
▶	mappings	—
◆	extends	../default
▼	demo-project-de	—
▶	domains	—
▼	i18n	—
◆	defaultLocale	de
◆	extends	../demo-project

A language selection dropdown allows editors to select the content authoring language:



Since navigation is generated from content, links display in the correct language, provided the content is translated:



Where built-in precautions fail due to editor error, a fallback language can be used to ensure that at least some of the content is rendered.

From SEO perspective, it is best when a crawler finds all language-specific content in one directory. Magnolia accomplishes this by injecting a language identifier in the URL.

Locale	URL
Fallback Language	<code>www.example.com/article.html</code>
German	<code>www.example.com/de/article.html</code>
French	<code>www.example.com/fr/article.html</code>
Spanish	<code>www.example.com/es/article.html</code>

Inserting Keywords in Page Titles and Description Meta Elements

Best practice

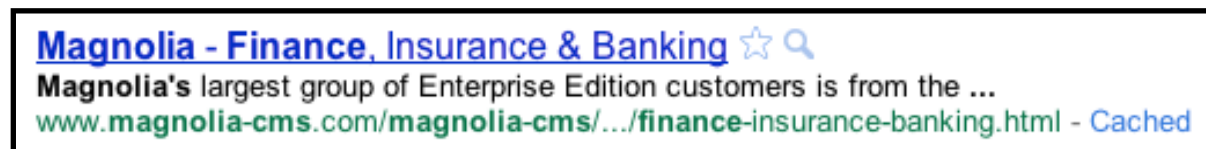
Page titles are an extremely important SEO factor. Search engines give titles considerable algorithmic weight and use them as headings on search results. Titles should include keywords and, where appropriate, your company name for branding purposes:

```
<title>Magnolia - Finance, Insurance & Banking</title>
```

Description meta elements on the other hand do not add much to page relevance but should not be neglected:

```
<meta name="description" content="Magnolia's largest group of Enterprise Edition customers is from the...">
```

Their content often appears below the title in search results:

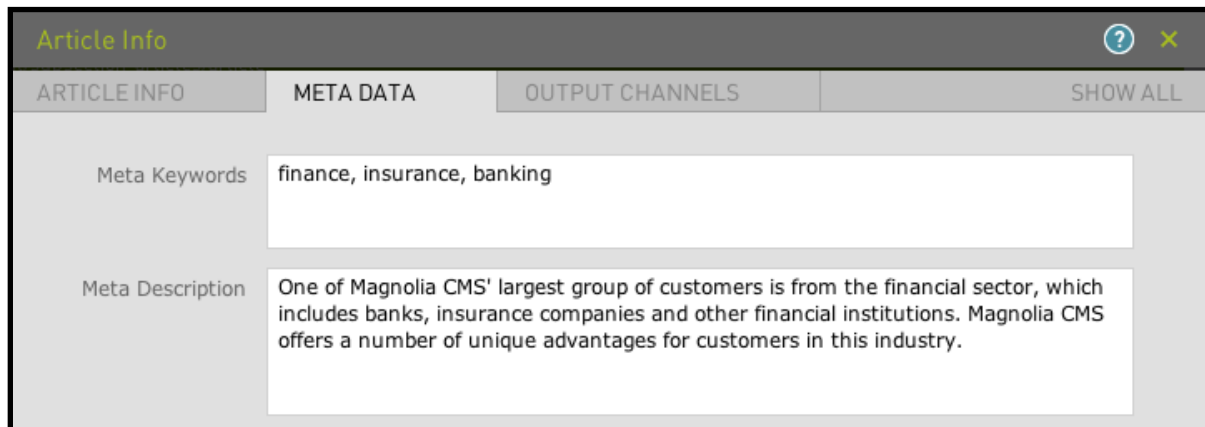


The relevance of keywords stipulated in the keyword tag once had great impact but this is no longer true. Actual page content is now far more relevant. Still, a keyword tag should be included:

```
<meta name="keywords" content="Finance, Insurance, Banking">
```

Magnolia compliance

Magnolia produces the `title` element from the page title automatically, so this aspect is never overlooked. Inserting keywords and descriptions is done in the page info dialog:



Article Info		?		X	
ARTICLE INFO	META DATA	OUTPUT CHANNELS	SHOW ALL		
Meta Keywords	finance, insurance, banking				
Meta Description	One of Magnolia CMS' largest group of customers is from the financial sector, which includes banks, insurance companies and other financial institutions. Magnolia CMS offers a number of unique advantages for customers in this industry.				

These fields can be made mandatory with a simple configuration change in AdminCentral. If an editor fails to provide the description, page abstract is used automatically as a fallback. Authors don't need to fill the field to be SEO compliant.

Enhancing Your Site's Link Popularity

Best practice

Search engines place strong emphasis on your site's link popularity. Links from external sites, especially highly ranked ones, increase your site's rank. The more links targeting your site from external sources the better.

Magnolia compliance

There are countless ways to build links and a comprehensive discussion on this is beyond the scope of this paper. Some ways that Magnolia can help to build incoming links are:

- By developing a blog on an external domain to promote your main site. The Templating Kit allows you to manage multiple sites in a single instance.
- A forthcoming [social media module \[12\]](#) will facilitate effortless integration with social media networks that have become a must for success on the Web.
- Social bookmarking provided in many Magnolia templates.
- Friendly URLs make creating external links intuitive and easy to create.

Attracting Visitors to Your Site

Best practice

SEO is like a merry-go-round: the more popular you are with users, the higher search engines rank you. And the higher your rank, the more visitors you attract. For SEO and visitors, developing and continually expanding a great site are key to attracting and retaining users.

Magnolia compliance

Magnolia offers a variety of features to help you create an attractive, enterprise-level website:

- Instant inline editing for fast changes.
- Move, copy and reorder site structures via drag-and-drop without breaking links.
- Imaging engine scales, crops and filters images to generate optimized variants.
- Multi-language content.
- Multisite support for microsites, campaigns, corporate subsidiaries.
- Human readable URLs.
- W3C accessibility.
- Templates for news, events, forms, images galleries, and categories.
- Podcasts, webcasts, video, audio, and RSS feeds.
- User-generated content, forums, and commenting.
- Content formatting for mobile devices.
- Campaign monitor integration for newsletters.
- Dublin Core based metadata.
- Full-text search.

References

- [1] <http://moz.com>
- [2] <http://www.statowl.com>
- [3] <http://www.w3.org>
- [4] <http://www.w3.org/TR>
- [5] <http://www.sitemaps.org/index.html>
- [6] <http://documentation.magnolia-cms.com/modules/google-sitemap.html>
- [7] <http://www.google.com/support/webmasters/bin/answer.py?answer=35769>
- [8] <http://documentation.magnolia-cms.com/templates/uri-mapping.html>
- [9] <http://moz.com/blog/canonical-url-tag-the-most-important-advancement-in-seo-practices-since-sitemaps>
- [10] <http://dublincore.org/>
- [11] http://documentation.magnolia-cms.com/technical-guide/configuration-mechanisms.html#Extending_configuration
- [12] <http://wiki.magnolia-cms.com/display/WIKI/Magnolia+Twigs+Module>

Learn More About Magnolia CMS

To learn more about Magnolia CMS:

- [Download](#) a free trial of Magnolia CMS.
- Visit magnolia-cms.com.
- Read [official documentation](#).
- Subscribe to the Magnolia [users mailing list](#) and [newsletter](#).
- Follow us on [Twitter](#), like us on [Facebook](#), circle us on [Google+](#) or find us on [LinkedIn](#) and [Xing](#).

Contact Information

Magnolia International Ltd. - Head Office

St.Johanns-Vorstadt 38
4056 Basel
Switzerland

+41 61 228 90 00

info@magnolia-cms.com

Magnolia España Software and Computer Applications S.L.

Paseo de la Castellana 153, Bajo
28046 Madrid - España

+34 66 263 43 36

info-es@magnolia-cms.com

Magnolia Americas, Inc.

168 SE 1st Street
Suite 1007 - Miami, FL 33131
United States of America

(305) 267-3033

info-us@magnolia-cms.com

Magnolia Software & Services CZ s.r.o.

Chobot 1578
76701 Kroměříž
Česká Republika

+420 571 118 715

info-cz@magnolia-cms.com

For a complete list of contact options, visit <http://www.magnolia-cms.com/company/contact-us.html>.

Legal Notices

©Magnolia International Ltd. 2014.

Magnolia is a registered trademark of Magnolia International Ltd.

All rights reserved. No part of this document may be reproduced in any form without prior permission in writing from Magnolia International Ltd.