MyFlorida.com	State Of Florid	A
my	STATE TECHNOLOGY OFFICE 4030 Esplanade Way, Suite 180 Tallahassee, Florida 32399-0950	
Doc Ref Number:	STO-2-72-007	Revision Number: 0.0 Revision Date:
Document Type:	Enterprise Standard	Page: 1 of 17
Policy Title:	Portal MetaTag Standard	

Synopsis:	This policy addresses issues related generating meta data on all web pages throughout the State Portal (MyFlorida.com).	
Authority:	F.S. 282.102 (23)(a) provides for the State Technology Office (STO) to develop an integrated electronic system which reflects cost- effective deployment strategies in keeping with industry standards and practices, including protections and security of private information as well as maintenance of public records.	
Applicability:		
Effective Date:	February 4, 2003	Expiration Date:
POC for Changes:	Grant Sellars	
Approval By:	Kimberly Bahrami, State CIO	
Approved on:	February 4, 2003	

Doc Ref Number:	STO-2-72-007	Revision Number: 0.0
Document Type:	Enterprise Standard	Page: 2 of 17
Policy Title:	Portal MetaTag Standard	

I. General Information

- 1. In order to provide the most efficient search engine results for web information and services accessible through MyFlorida.com, all Florida state government .html web pages, accessed from MyFlorida.com, will be properly coded with meta data information as described in the "MyFlorida.com Search Engine Optimization Strategies for State Agency Web Developers" document.
- 2. The document titled "MyFlorida.com Search Engine Optimization Strategies for State Agency Web Developers" will be available at the State Technology Office Web site (www.myflorida.com/myflorida/sto) under Plans & Policies.
- 3. The Portal MetaTag Standard will be review annually and is subject to change by the State Technology Office as necessary.

Contact Information

Any questions regarding the Portal MetaTag Standard may be addressed to the Portal Management Team at: <u>info@myflorida.com</u>.

II. Definitions

None

III. Implementing Procedures

- 1. All Florida state government Web sites will develop web content to comply with published policy.
- 2. Agency Webmasters and Agency Chief Information Officers should be familiar with Portal policies and standards prepared by the State Technology Office and should ensure that all published web content, accessible through MyFlorida.com, meets requirements as established.
- 3. Florida government Web sites (and web content) will periodically be reviewed by State Technology Portal Team members for compliance with published policies and standards. Notices of non-compliance or suggested improvement(s) will be addressed with agencies as necessary.

IV. Initial Distribution

- 1. STO internal Organization
- 2. Agency Chief Information Officers
- 3. Agency Webmasters
- 4. MyFlorida.com Portal Team
- V. Development and Revision History

Doc Ref Number:	STO-2-72-007	Revision Number: 0.0
Document Type:	Enterprise Standard	Page: 3 of 17
Policy Title:	Portal MetaTag Standard	

 $\ensuremath{\mathsf{N/A}}$ – this is the original document with no previous versions.

VI. Approval Signature Block

Kumherly Bahrami	2/4/03
Name & Title:	Date
Kimberly Bahrami	
State Chief Information Officer	

VII. Listing of Appendices

None

Doc Ref Number:		Revision Number: 0.0
	Enterprise Standard Portal MetaTag Standard	Page: 4 of 17

MetaTag Descriptions

The purpose of meta tagging strategies is to gain optimum results when using the MyFlorida portal search engine. The portal search engine is currently powered by Inktomi and uses the following HTML tags, as shown in the table below:

Тад	Content Index & used	
Tag	to locate page	in search results
HTML	4.0 Tags:	
Meta Tag - Description	X	Х
Meta Tag - Keywords	X	
HTML Tag - Title	X	Х
HTML Tag - Body	x	(only if Description is missing)
HTML Tag - IMG Alt= Description	X	
HTML Tag - <a (linked="" href="" phrase)<="" td=""><td>X</td><td></td>	X	
Dublin Core Meta Tags:		
dc.date	X	
dc.subject	X	
dc.publisher	X	Х

Currently, the search engine is configured to scan agency web sites each weekend for new and updated pages. It will store meta tag, page title, and page body information and make this information available for searching. However, when attempting to sort the results of a search, it will rank the pages where the search term(s) were found in the following tags higher:

<u>Tag</u>	<u>Weight</u>	
Keywor	ds	10
Page Ti	le	7
Descrip	tion	5
Remote	e Anchors	4
Alt Tag	S	1

For instance, if the word "employment" is entered on the MyFlorida search page, it will return all pages with this term. However, it will sort all of these pages in the following order:

Ranking	Number of Times Search Term "Employment" was found in:
1st	Meta Tag Keywords, Page Title, Meta Tag Description, Anchor Tag (link
	contains search term), Image Alt Tag Description contains search term,
	Page Body.
2nd	Meta Tag Keywords, Page Title, Meta Tag Description, Anchor Tag
	(link), Page Body.
3rd	Meta Tag Keywords, Page Title, Meta Tag Description, Page Body.
4th	Meta Tag Keywords, Page Title, Page Body.
5th	Meta Tag Keywords, Page Body.
6th	Page Body.

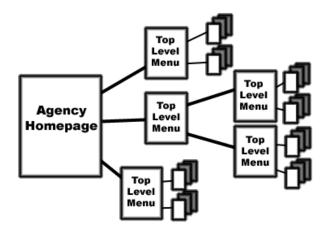
Doc Ref Number:	STO-2-72-007	Revision Number: 0.0
Document Type:	Enterprise Standard	Page: 5 of 17
Policy Title:	Portal MetaTag Standard	

To continue this example, let's say that 1,000 pages were found with the term "employment". Of the 1,000 pages, only 10 also contained the term "employment" in the Meta Tag Keywords, Page Title, Meta Tag Description, Anchor Tag (link contained the word "employment"), Image Alt Tag (image description contained "employment"), and the Page Body (text on the page contained "employment). Then these 10 pages would fall under the first category in the chart above and be ranked according to the number of occurrences of the word "employment." The page containing the most occurrences would be listed first.

If, on the other hand, someone searched for "Kalamazoo" and only five pages contained this word in the body text (this word is not contained in meta tags or other elements), then the search engine would list all five pages and rank them based on number of times the word Kalamazoo appeared on the page.

It is important to note that the page with NO KEYWORDS is ranked last. It is also important to realize that the above chart also applies to most other search engines (such as Google, AltaVista, etc). Adding meta tags to your pages will not only help them to be found in the MyFlorida search engine, but also in most other search engines.

In order to optimize the MyFlorida search engine results, and help MyFlorida visitors locate government information quickly and easily, agency webmasters are encouraged to adopt the following meta tagging strategy:



At a minimum, the pages on your website represented in the above example by a white background should contain meta tags. A good meta tagging strategy includes:

- (1) Applying appropriate meta tags to all top-level menu pages.
- (2) Applying appropriate meta tags to all intro/index pages to areas of your web site. These areas could be bureau homepages or menu pages that link to pages on a specific subject.
- (3) Using Meta Tag Keywords thoughtfully so that only those terms appropriate for the page are selected.
- (4) Using Meta Tag Keyword phrases when possible. For instance, someone searching for employment information may also be interested in a specific area of employment and might narrow their results by searching for "health care employment" or "law enforcement employment". Therefore, including the keyword phrase "law enforcement

Doc Ref Number:		Revision Number: 0.0
Document Type:	Enterprise Standard	Page: 6 of 17
Policy Title:	Portal MetaTag Standard	

employment" as well as the keyword "employment" would be appropriate on a menu page for employment information in one of the law enforcement occupations.

MetaTag Coding Example

Below is an example of the meta tags that might be applied to a top-level menu for employment in a law enforcement related occupation (Corrections):

<title>Working for Corrections - Employment, Volunteer and Internship Information</title>

<META name="description" content="Learn more about the employment, volunteer and internship opportunities available with the Florida Department of Corrections.">

<META name="keywords" content="academies, applicants, application, apply, benefits, certification, certified, community supervision, correctional officer, correctional probation officer, correctional, corrections officer, corrections, Department of Corrections, duties, employee, employees, employment application, employment, exams, Florida prisons, florida, government, guard, health care, health, healthcare, internship, job list, job, law enforcement, minimum qualifications, minimum salary, officer, opportunities, position, positions, posts, prison guard, prison, probation officer, probation, qualifications, salary, selection process, state, statewide, supplemental application, training centers, training, vacancies, vacant, volunteer, volunteers">

<meta name="DC.Publisher" content="Florida Department of Corrections">

<meta name="DC.Subject" scheme="LCSH" content="Florida Dept. of Corrections -- Job Vacancies">

</head>

<body bgcolor="#FFFFFF" text="#000000" link="#0000FF" vlink="#800080" alink="#000080">

{. . page body text, images with ALT descriptions, Linked phrases . .}

</body> </html>

Meta Tag Standards & Examples

For more information on any of the tags listed in this section, please see the following technical documents:

Dublin Core: http://dublincore.org/

HTML 4.0 Meta Tag and Title Standards:

Doc Ref Number:	STO-2-72-007	Revision Number: 0.0
Document Type:	Enterprise Standard	Page: 7 of 17
Policy Title:	Portal MetaTag Standard	

http://www.w3.org/TR/REC-html40/struct/global.html

W3C "Helping Search Engines Index Your Website" http://www.w3.org/TR/REC-html40/appendix/notes.html#recs

Reference for HTML 4.0 Standards

While the HTML 4.0 standards for meta tagging are less specific than the newer Dublin Core, they are much more universally accepted and used by other search engines. It is therefore recommended that developers use the HTML 4.0 tags for Keywords and Descriptions instead of the Dublin Core dc.description and dc.keywords.

In addition, the HTML page title tag, <title>, receives the second highest weighting within MyFlorida search results. It is not necessary to repeat the page title information in the Dublin Core dc.title tag, as the MyFlorida search engine will use the HTML title tag (which is a required tag under HTML 4.0 specifications).

Keywords

Keywords are the most important element for controlling search results on MyFlorida because the search engine places the highest weight on words and phrases contained in this tag.

Example:

<META name="keywords" content="vacation, Greece, sunshine"> <META name="keywords" content="vacation,Greece,sunshine">

Tips & Considerations:

- Consider what a person might enter in a search engine to find your information.
- Limit your keywords (and phrases) list to no more than 1,000 characters, including spaces.
- Place the least important keywords last because some search engines will not read the full 1,000 characters.
- Test your keywords by searching MyFlorida (as well as other search engines) using individual keywords as well as word combinations and phrases to see what types of pages show up in the results. This test will give you an idea whether the keywords you chose are appropriate for the topic of your page. If certain combinations of keywords seem to target pages with similar information better, move these keywords to the beginning of your keyword list.
- If you have web traffic reporting software, review the reports that show which search terms were used by visitors who found your pages. Are some keywords used more frequently (and need to be added to more pages)? Are terms reported with less usage due to visitor usage or the fact that the term is only found in the body text of a few pages?

Policy Title:	Portal MetaTag Standard	
Document Type:	Enterprise Standard	Page: 8 of 17
Doc Ref Number:	STO-2-72-007	Revision Number: 0.0

Description

The first 150-180 characters (with spaces) of this description will appear in the MyFlorida search results. When a description IS NOT provided, then the MyFlorida search engine will display the first characters from the words and phrases found on the page (including Alt tag descriptions of images).

Example:

<META name="description" content="Idyllic Florida vacations">

Tips & Considerations:

- Keep the text between 150 to 180 characters, including spaces, so it won't be cut short.
- Provide a different title and description for each page on the site.
- Make sure each description matches the first sentences in the body text (for engines that don't look at META tags).

Example of MyFlorida search results when Description tag IS NOT provided:

Working for Corrections - Employment, Volunteer and Internship	xx%
Information	
Click to skip menus and jump to page body MyFlorida.com State	
Government Portal Page Border Blue Line directory contact us 411	24 May 02
	Find
Florida Department of Corrections http://www6.myflorida.com/employ/	Similar
- 12.4KB	

Example of MyFlorida search results when Description tag IS provided:

Working for Corrections - Employmen t, Volunteer and Internship	xx%
Information	
Learn more about the employment , volunteer and internship	
opportunities available with the Florida Department of Corrections	24 May 02
Florida Department of Corrections http://www6.myflorida.com/employ/	Find
- 12.4KB	Similar

Note: Descriptions can be displayed for Microsoft Word, Powerpoint & Excel documents as well as Acrobat PDF documents through completion of the document's property sheet. For Microsoft, look under "File" then "Properties". For Acrobat look under "File" then "Document Information" then "General".

Doc Ref Number:	STO-2-72-007	Revision Number: 0.0
Document Type:	Enterprise Standard	Page: 9 of 17
Policy Title:	Portal MetaTag Standard	

Title

Every HTML document **must** have a <u>TITLE</u> element in the <u>HEAD</u> section. In addition, this Title must be unique among all other pages on your website.

Authors should use the <u>TITLE</u> element to identify the contents of a document. Since users often consult documents out of context, authors should provide context-rich titles. Thus, instead of a title such as "Introduction", which doesn't provide much contextual background, authors should supply a title such as "Introduction to Medieval Bee-Keeping" instead.

For reasons of accessibility, user agents must always make the content of the $\underline{\text{TITLE}}$ element available to users (including $\underline{\text{TITLE}}$ elements that occur in frames). The mechanism for doing so depends on the user agent (e.g., as a caption, spoken).

Titles may contain <u>character entities</u> (for accented characters, special characters, etc.), but **may not contain other markup** (including comments). Here is a sample document title:

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN"

"http://www.w3.org/TR/html4/strict.dtd">

<HTML>

<HEAD>

<TITLE>A study of population dynamics</TITLE>

... other head elements...

</HEAD>

<BODY>

... document body...

</BODY>

</HTML>
```

Tips & Considerations:

- Remember -- the title tag is one of the most important elements to search engines that collect meta tag data.
- It will appear when the site is bookmarked.
- Enter words that will differentiate your site from others.
- Limit the text's size to 69 characters, including spaces.

Example of MyFlorida search results when the TITLE tag IS NOT provided:

http://www6.myflorida.com/employ/	xx%
Learn more about the employment , volunteer and internship	
opportunities available with the Florida Department of Corrections	
Florida Department of Corrections http://www6.myflorida.com/employ/	24 May 02
- 12.4KB	Find
	Similar

Example of MyFlorida search results when the TITLE tag IS provided:

Working for Corrections - Employmen t, Volunteer and Internship	xx%
Information	
Learn more about the employment , volunteer and internship	

Doc Ref Number:	STO-2-72-007	Revision Number: 0.0
Document Type:	Enterprise Standard	Page: 10 of 17
Policy Title:	Portal MetaTag Standard	

opportunities available with the Florida Department of Corrections	24 May 02
Florida Department of Corrections http://www6.myflorida.com/employ/	Find
- 12.4KB	Similar

Doc Ref Number:	STO-2-72-007	Revision Number: 0.0
Document Type:	Enterprise Standard	Page: 11 of 17
Policy Title:	Portal MetaTag Standard	

Reference for Dublin Core

DC.Date

Name:DateIdentifier:DateDefinition:A date associated with an event in the life cycle of the resource.Comment:Typically, Date will be associated with the creation or availability of the
resource. Recommended best practice for encoding the date value is
defined in a profile of ISO 8601 [W3CDTF] and follows the YYYY-MM-DD
format.

Date (of creation or availability of resource; [WTN8601] recommended)

Qualifiers that refine *Date*:

Created

Name: created Label: Created Definition: Date of creation of the resource.

Valid

Name: valid Label: Valid Definition: Date (often a range) of validity of a resource.

Available

Name: available Label: Available Definition: Date (often a range) that the resource will become or did become available.

Issued

Name: issued Label: Issued Definition: Date of formal issuance (e.g., publication) of the resource.

Modified

Name: modified Label: Modified Definition: Date on which the resource was changed.

Doc Ref Number:	STO-2-72-007	Revision Number: 0.0
Document Type:	Enterprise Standard	Page: 12 of 17
Policy Title:	Portal MetaTag Standard	

Encoding Schemes for *Date***:**

DCMI Period

Name: Period Label: DCMI Period Definition: A specification of the limits of a time interval. See also: <u>http://dublincore.org/documents/dcmi-period/</u>

W3C-DTF

Name: W3CDTF Label: W3C-DTF Definition: W3C Encoding rules for dates and times - a profile based on ISO 8601 See also: <u>http://www.w3.org/TR/NOTE-datetime</u>

Examples: <meta name = "DC.Date" content = "1820"> <meta name = "DC.Date" content = "1935"> <meta name = "DC.Date" content = "1998-05-14"> <meta name = "DC.Date" scheme = "WTN8601" content = "1998-05-14">

DC.Subject

Name:	Subject and Keywords	
Identifier:	Subject	
Definition:	The topic of the content of the resource.	
Comment:	Typically, a Subject will be expressed as keywords, key phrases	s or
	classification codes that describe a topic of the resource.	
	Recommended best practice is to select a value from a	controlled
	vocabulary or formal classification scheme.	

Subject (topic or keyword)

The qualifiers below are recommended for the Subject element:

Encoding Schemes for Subject:

LCSH

Name: LCSH Label: LCSH Definition: Library of Congress Subject Headings

MeSH

Name: MESH Label: MeSH

Doc Ref Number:	STO-2-72-007	Revision Number: 0.0
Document Type:	Enterprise Standard	Page: 13 of 17
Policy Title:	Portal MetaTag Standard	

Definition: Medical Subject Headings

See also: http://www.nlm.nih.gov/mesh/meshhome.html

DDC

Name: DDC Label: DDC Definition: Dewey Decimal Classification See also: <u>http://www.oclc.org/dewey/index.htm</u>

LCC

Name: LCC Label: LCC Definition: Library of Congress Classification See also: http://lcweb.loc.gov/catdir/cpso/lcco/lcco.html

UDC

```
Name: UDC
Label: UDC
Definition: Universal Decimal Classification See also:
http://www.udcc.org/
```

Example:

```
<meta name = "DC.Subject"
content = "heart attack">
<meta name = "DC.Subject" scheme = "MESH"
content = "Myocardial Infarction; Pericardial Effusion">
<meta name = "DC.Subject"
content = "vietnam war">
<meta name = "DC.Subject" scheme = "LCSH"
content = "Vietnamese Conflict, 1961-1975">
<meta name = "DC.Subject" scheme = "LCSH"
content = "Friendship">
<meta name = "DC.Subject"
content = "Friendship">
```

Note: The MyFlorida search engine will collapse information contained in the DC.Subject tag with the Keywords tags. This means that words and phrases contained in this tag will receive the highest possible ranking in the search results, but the same results will be achieved by using the KEYWORDS tag.

Doc Ref Number:	STO-2-72-007	Revision Number: 0.0
Document Type:	Enterprise Standard	Page: 14 of 17
Policy Title:	Portal MetaTag Standard	

DC.Publisher

Name: Identifier:	Publisher Publisher	
Definition: Comment:	An entity responsible for making the resource available Examples of a Publisher include a person, an organization, service. Typically, the name of a Publisher should be used to indicate the entity.	or a

Publisher (entity that made the resource available)

<meta content="O'Reilly" name="DC.Publisher"/>
<meta content="Digital Equipment Corporation" name="DC.Publisher"/>
<meta content="University of California Press" name="DC.Publisher"/>
<meta content="State of Florida (USA)" name="DC.Publisher"/>

When the DC.Publisher information is provided the MyFlorida search engine will display the publisher information. See examples:

Exa	mp	ole o	of M	yFlorida	searc	h	results	w	hei	ı DC	C.Publisher	tag	IS NO	T provided:
			-	_		-								

Working for Corrections - Employmen t, Volunteer and Internship	xx%
Information	
Learn more about the employment , volunteer and internship	
opportunities available with the Florida Department of Corrections	24 May 02
http://www6.myflorida.com/employ/ - 12.4KB	Find
	Similar

Example of MyFlorida search results when DC.Publisher tag IS provided:

Working for Corrections - Employmen t, Volunteer and Internship	xx%
Information	
Learn more about the employment , volunteer and internship	
opportunities available with the Florida Department of Corrections	24 May 02
Florida Department of Corrections http://www6.myflorida.com/employ/	Find
- 12.4KB	Similar

Notice that the publisher information (in this case Florida Department of Corrections) appears on the last line with the URL. Publisher information should always be the agency name and not the bureau or unit. If needed, use the DC.Creator tag to provide author information.

DC.Creator

Name:CreatorIdentifier:CreatorDefinition:An entity primarily responsible for making the content of the resource.

Doc Ref Number:	STO-2-72-007	Revision Number: 0.0
Document Type:	Enterprise Standard	Page: 15 of 17
Policy Title:	Portal MetaTag Standard	

Comment: Examples of a Creator include a person, an organization, or a service. Typically, the name of a Creator should be used to indicate the entity.

Guidelines for usage:

Creators should be listed separately in the same order that they appear in the publication. Personal names should be listed surname or family name first, followed by forename or given name. When in doubt, give the name as it appears, and do not invert.

In the case of organizations where there is clearly a hierarchy present, list the parts of the hierarchy from largest to smallest, separated by full stops. If it is not clear whether there is a hierarchy present, or unclear which is the larger or smaller portion of the body, give the name as it appears in the item.

If the Creator and Publisher are the same, use only the Publisher tag. If the nature of the responsibility is ambiguous, the recommended practice is to use Publisher for organizations, and Creator for individuals. In cases of lesser responsibility, other than creation, use dc.Contributor.

Examples for Creator (entity that created the intellectual content)

<meta name = "DC.Creator" content = "Gogh, Vincent van"> <meta name = "DC.Creator" content = "van Gogh, Vincent"> <meta name = "DC.Creator" content = "Mao Tse Tung"> <meta name = "DC.Creator" content = "Mao, Tse Tung"> <meta name = "DC.Creator" content = "Plato"> <meta name = "DC.Creator" content = "Plato">

Doc Ref Number:	STO-2-72-007	Revision Number: 0.0
Document Type:	Enterprise Standard	Page: 16 of 17
Policy Title:	Portal MetaTag Standard	

HTML 4.0 Reference for Robots

For more information, please see: http://www.w3.org/TR/REC-html40/appendix/notes.html#idx-search robot

Provide robots with indexing instructions

People may be surprised to find that their site has been indexed by an indexing robot and that the robot should not have been permitted to visit a sensitive part of the site. Many Web robots offer facilities for Web site administrators and content providers to limit what the robot does. This is achieved through two mechanisms: a "robots.txt" file and the <u>META</u> element in HTML documents.

The robots.txt file

When a Robot visits a Web site, say http://www.foobar.com/, it firsts checks for http://www.foobar.com/robots.txt. If it can find this document, it will analyze its contents to see if it is allowed to retrieve the document. You can customize the robots.txt file to apply only to specific robots, and to disallow access to specific directories or files.

Here is a sample robots.txt file that prevents all robots from visiting the entire site

User-agent: * # applies to all robots Disallow: / # disallow indexing of all pages

The Robot will simply look for a "/robots.txt" URL on your site, where a site is defined as an HTTP server running on a particular host and port number. Here are some sample locations for robots.txt:

Site URL	URL for robots.txt
http://www.w3.org/	http://www.w3.org/robots.txt
http://www.w3.org:80/	http://www.w3.org:80/robots.txt
http://www.w3.org:1234/	http://www.w3.org:1234/robots.txt
http://w3.org/	http://w3.org/robots.txt

There can only be a single "/robots.txt" on a site. Specifically, you should not put "robots.txt" files in user directories, because a robot will never look at them. If you want your users to be able to create their own "robots.txt", you will need to merge them all into a single "/robots.txt". If you don't want to do this, your users might want to use the Robots META Tag instead.

Some tips:

URL's are case-sensitive, and "/robots.txt" string must be all lower-case. Blank lines are not permitted within a single record in the "robots.txt" file.

There must be exactly one "User-agent" field per record. The robot should be liberal in interpreting this field. A case-insensitive substring match of the name without version information is recommended.

Doc Ref Number:	STO-2-72-007	Revision Number: 0.0
Document Type:	Enterprise Standard	Page: 17 of 17
Policy Title:	Portal MetaTag Standard	

If the value is "*", the record describes the default access policy for any robot that has not matched any of the other records. It is not allowed to have multiple such records in the "/robots.txt" file.

The "Disallow" field specifies a partial URL that is not to be visited. This can be a full path, or a partial path; any URL that starts with this value will not be retrieved. For example,

Disallow: /help disallows both /help.html and /help/index.html, whereas **Disallow: /help/** would disallow /help/index.html but allow /help.html.

An empty value for "Disallow" indicates that all URLs can be retrieved. At least one "Disallow" field must be present in the robots.txt file.

Robots and the META element

The <u>META</u> element allows HTML authors to tell visiting robots whether a document may be indexed, or used to harvest more links. No server administrator action is required. In the following example a robot should neither index this document, nor analyze it for links.

<META name="ROBOTS" content="NOINDEX, NOFOLLOW">

The list of terms in the content is ALL, INDEX, NOFOLLOW, NOINDEX.