

Web Crawlability in Certain Events

Overview

This document is about the visibility of your event to search engine web crawlers, also known as robots or 'bots'.

There are three options:

1. **All Private:** No events are crawlable. This is the default, so you don't need to take any action.
2. **All Public:** You want *all* events to be crawlable. Ask Certain to enable this for your domain.
3. **Some Private:** You want only *some* events to be crawlable. Follow step 2, and then use the META tag described below to make an event private (that is, *not* crawlable).

How to structure the contents of your site to optimize it for search engines (SEO) is **not** in the scope of this document.

Default Behavior in Certain

By default, events created within Certain are **not** crawlable. If you would like all your events to be public, Certain can enable web crawlers across your domain and block-allow your events to be indexed. Contact your Customer Success Manager, who can facilitate this request.

Once this is enabled for your domain, there are additional HTML META tags that can be added to the display shell to enable **better** crawlability; however, as stated above, that is outside the scope of this document.

META Tags

This section is applicable if you've asked Certain to enable web crawlers in your domain, as described above.

You can use a special HTML <META> tag to tell robots to index or *not* index the content of a page, and/or not scan it for links to follow. By simply adding this extra HTML tag into the head of your Certain event display shell, you can instruct the web crawlers to exclude your event's website(s) and form(s) from being indexed.

Private Events

In order for Certain events to be **excluded** from being crawled, you should add the **robots** META tag described below to the custom display shell of the events' display configuration.

How to write a Robots META Tag

When to include it

The default behavior of sites that exclude the **robots** META tag is that they are found and indexed by web crawlers. If you would like to exclude your event from being found by web

crawlers then you should add this extra tag into your event display shell with the following values.

What to put into it

The "NAME" attribute must be "ROBOTS".

Valid values for the "CONTENT" attribute are: "INDEX", "NOINDEX", "FOLLOW", and "NOFOLLOW". Multiple comma-separated values are allowed, but obviously only some combinations make sense. If there is *no* robots <META> tag, the default is "INDEX, FOLLOW", so there's no need to spell that out. That leaves:

```
<META NAME="ROBOTS" CONTENT="NOINDEX, FOLLOW">  
<META NAME="ROBOTS" CONTENT="INDEX, NOFOLLOW">  
<META NAME="ROBOTS" CONTENT="NOINDEX, NOFOLLOW">
```

Where to put it

Like any <META> tag, it should be placed in the **HEAD** section of an HTML page, as in the example below. You should put it in every page on your site, which you can do by placing it in the advanced display shell with Certain's display configuration. (Plan > Configure > Display.) This enables the HTML "wrapper" to be included in all the websites and forms in the event.

```
Shell Header (HTML to include above the main page <table>)  
  
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN"  
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">  
<html>  
<head>  
<meta name="ROBOTS" content="NOINDEX, NOFOLLOW" />  
<meta name="Description" content="{Event Notes}" />
```

Values for 'Content' Attribute

See overleaf for a list of the values for **content**, and the corresponding behavior that web crawlers, or "bots", will exhibit when they're included as part of your **ROBOTS <META>** tag.

This is a list of potential values for the **content** of `<meta name="robots">`, with the corresponding behavior that web crawlers, or "bots", will exhibit.

Value	Description	Used by
index	Allows the robot to index the page (default).	All
noindex	Requests the robot to not index the page.	All
follow	Allows the robot to follow the links on the page (default).	All
nofollow	Requests the robot to not follow the links on the page.	All
none	Equivalent to noindex, nofollow	Google
noodp	Prevents using the Open Directory Project description, if any, as the page description in search engine results.	Google, Yahoo, Bing
noarchive	Requests the search engine not to cache the page content.	Google, Yahoo, Bing
nosnippet	Prevents displaying any description of the page in search engine results.	Google, Bing
noimageindex	Requests this page not to appear as the referring page of an indexed image.	Google

List from “<meta>: The Document-level Metadata element”, by Mozilla Contributors:
<https://developer.mozilla.org/en-US/docs/Web/HTML/Element/meta>, licensed under CC-BY-SA 2.5.