

Cheat Sheet:

XPath für Nicht-XPath-Kundige

Ich möchte...	So mache ich das mit XPath
<p>Ich möchte den Inhalt eines Attributs (z. B. die Meta Description)</p> <p>HTML-Code: <tag attribut1="wert1" attribut2="wert2"></p>	<p>Schreibweise: //tag[@attribut1="wert1"]/@attribut2</p> <p>Beispiel: <meta name="description" content="Das will ich!"></p> <p>XPath: //meta[@name="description"]/@content</p>
<p>Ich möchte den Inhalt eines von <tag> und </tag> umschlossenen Tags</p> <p>HTML-Code: <tag>inhalt</tag></p>	<p>Schreibweise: //tag</p> <p>Beispiel: <h3>Das will ich!</h3></p> <p>XPath: //h3</p>
<p>Ich möchte den Inhalt eines von <tag> und </tag> umschlossenen Tags, das durch ein bestimmtes Attribut gekennzeichnet ist</p> <p>HTML-Code: <tag attribut="wert">inhalt</tag></p>	<p>Schreibweise: //tag[@attribut='wert']</p> <p>Beispiel: <h1 class="hell">Das will ich!</h1></p> <p>XPath: //h1[@class='hell']</p>

<p>Wenn es mehrere passende Objekte gibt, kann man über [index] auf das jeweilige Objekt zugreifen</p>	<p>Schreibweise: XPath-Ausdruck[index]</p> <p>Beispiel:</p> <pre> link 1 link 2 </pre> <p>XPath:</p> <pre>//ul/li[1] liefert „link1“ //ul/li[2] liefert „link2“</pre>
<p>Ich möchte ein bestimmtes Element, das von anderen Elementen umschlossen ist</p>	<pre>/element1/element2/element3/...</pre> <p>Beispiel:</p> <pre>/html/head/meta[@name="description"]/@content</pre> <p>Liefert den Wert des content-Attributs eines Meta-Tags, bei dem das Attribut „name“ „description“ entspricht. Das Meta-Tag wird vom <head>-Tag und dieses wiederum vom <html>-Tag umschlossen.</p>

Weitere Tipps:

- <https://builtvisible.com/seo-guide-to-xpath/>
- <https://www.screamingfrog.co.uk/web-scraping/>