

www.legislation.gov.uk Website Information Guide

A brief introduction to the www.legislation.gov.uk legislation API and website data organisation.

Types of page on the website:

- **Static pages:**

These pages have little or no dynamic text and change very rarely, e.g. homepage

<http://www.legislation.gov.uk/>, the Help and information pages e.g.

<http://www.legislation.gov.uk/understanding-legislation>, and forms for searching e.g.

<http://www.legislation.gov.uk/search> and browsing e.g. <http://www.legislation.gov.uk/browse>,

<http://www.legislation.gov.uk/browse/uk> and <http://www.legislation.gov.uk/changes>.

- **Document Pages:**

These are the pages that provide access to document content. There are a lot of variations and ways to view different document a summary of key points and options follows: The base format for most legislation document content on the website is XML conforming to the Crown Legislation Mark-up Language (CLML) schema. You can access the XML data via the legislation API by adding "/data.xml" to any legislation content page URI e.g. the XML version of the page

<http://www.legislation.gov.uk/ukpga/2010/1> is

<http://www.legislation.gov.uk/ukpga/2010/1/data.xml>. This data is stored in a native XML database

and transformed into several different formats for different uses. These include the default XHTML

view (/data.htm), an Akoma Ntoso XML version (Akoma Ntoso is an international legal XML standard - <http://www.akomantoso.org/>), e.g.

<http://www.legislation.gov.uk/ukpga/2010/1/data.akn> (/data.akn) and an HTML5 serialization of

the Akoma Ntoso XML, e.g. <http://www.legislation.gov.uk/ukpga/2010/1/data.html> (/data.html) as

well as generated PDFs, e.g. <http://www.legislation.gov.uk/ukpga/2010/1/data.pdf> (/data.pdf). We

also provide a plain view of the XHTML that displays pure legislation content without any additional functionality e.g. <http://www.legislation.gov.uk/ukpga/2010/1?view=plain> and an un-styled version

of the XHTML (/data.xht), e.g. <http://www.legislation.gov.uk/ukpga/2010/1/data.xht>. These URI

patterns apply to all documents and sections of documents (e.g.

<http://www.legislation.gov.uk/uksi/2018/288/article/1/made/data.xml>) except for the small

number of older documents that we only have as PDF images.

- **Document Versions:** Each document may have several versions, an original "enacted", "made" or "adopted" version and "revised" versions that show the content as it stands at different points-in-time following the incorporation of amendments. Welsh legislation will have versions in different languages. You can see all the versions available by viewing the underlying XML for the document, the <atom:link> tags at the beginning list document versions and associated page links. When you are viewing a document e.g. <http://www.legislation.gov.uk/ukpga/2016/15> the points on the timeline will correspond with these versions, e.g. <http://www.legislation.gov.uk/ukpga/2016/15/2016-12-05>. You can create a valid URI to any date on the timeline, even those that don't coincide with version changes e.g. <http://www.legislation.gov.uk/ukpga/2016/15/2017-01-01> but this will display the same text as the relevant version. If no date is specified in the URI you will always get the most recent version, i.e. today's date. The timeline can be switched on and off when viewing revised versions.

- **Tables of Contents:** When accessing a document you will generally start with the Table of Contents e.g. <http://www.legislation.gov.uk/ukpga/2010/1/contents> . There are a lot of different ways to view documents and sections of documents. You may consider it unnecessary to archive each individual section of a document as archiving the whole document e.g. <http://www.legislation.gov.uk/ukpga/2010/1> will include the individual sections e.g. <http://www.legislation.gov.uk/ukpga/2010/1/section/1> and <http://www.legislation.gov.uk/ukpga/2010/1/section/2> . The table of contents can be different for the different document versions, e.g. if sections have been added over time, e.g. <http://www.legislation.gov.uk/ukpga/1984/60/contents/2005-07-01> .
- **Additional Document Pages:** In addition to the document text and tables of contents there are additional pages for “More Resources” e.g. <http://www.legislation.gov.uk/ukpga/2010/1/resources> , Explanatory Notes e.g. <http://www.legislation.gov.uk/ukpga/2010/1/notes/contents> and other associated documents. Links to these pages will be listed in the XML <atom:link> tags for any version.
- **Plain View XHTML and Generated PDFs:** These pages can be accessed by entering the URI directly or through the “Plain View” and “Print Options” buttons at the top right-hand-side of any content page.

- **Search Result pages:**

Search results show lists of legislation matching the search criteria. The Browse pages for legislation types are the same as the Search Result pages for a particular legislation type, e.g. clicking on “UK Public general Acts” from the page <http://www.legislation.gov.uk/browse/uk> will take you to <http://www.legislation.gov.uk/ukpga> . You can find all content available via the Advanced Search menus through Document Pages (described below) the only difference is highlighting of matched search terms. One way to find all documents on the website is to track through the /[/type]/[year]/[number] URI permutations by paging through the browse pages e.g. <http://www.legislation.gov.uk/ukpga/2019> , <http://www.legislation.gov.uk/ukpga/2018?page=2> etc. There are about 300,000 distinct documents on the website. Each of which may have multiple temporal versions and formats.

- **New Legislation pages:**

These pages show new legislation published. They are essentially search results for new legislation grouped by publishing date <http://www.legislation.gov.uk/new> . There is a page for each weekday e.g. <http://www.legislation.gov.uk/new/2019-12-06> .

- **Changes to Legislation pages:**

These pages list tabulated changes to legislation over time. You can filter the lists in lots of different ways but to get all data you can track through permutations of years and document types but there is a lot of data here (around 1 million changes recorded), e.g. <http://www.legislation.gov.uk/changes/affected/ukpga/2018/affecting/uksi/2019> will list changes made to 2018 UK Public General Acts by 2019 UK Statutory Instruments.

Search result pages for document searches, e.g. <http://www.legislation.gov.uk/uksi> or <http://www.legislation.gov.uk/2016?title=finance> , New legislation page lists e.g. <http://www.legislation.gov.uk/new> and Changes to Legislation searches, as in <http://www.legislation.gov.uk/changes/affected/uksi/2005/735> , are available as HTML or ATOM

feeds, by adding "/data.feed" to the end of the URI
e.g. <http://www.legislation.gov.uk/uksi/data.feed> or
<http://www.legislation.gov.uk/2016?title=finance/data.feed> or
<http://www.legislation.gov.uk/new/data.feed> or
<http://www.legislation.gov.uk/changes/affected/uksi/2005/735/data.feed> . The raw XML for the feeds contains much richer metadata than that displayed in the regular HTML pages. You may need to right-click and "view source" when viewing a /data.feed page in your browser to see the XML. Each feed page is limited to around 20 results but you can navigate through paged results e.g. <http://www.legislation.gov.uk/uksi/2019/data.feed?page=2> to obtain all entries required.

- **Sitemaps:**

See <http://www.legislation.gov.uk/robots.txt> for links to the sitemaps. The majority of document and pages types should be present in the sitemaps. However, the draft legislation and Impact Assessments are not indexed.

- **Content Negotiation:**

In addition to the URIs above, the website supports HTTP content negotiation. This allows users to configure browsers and/or scripts to specify a certain format to be returned from the default URI rather than using the "/data.xxx" extension.