

Opening up legislation for re-use



Legislation.gov.uk,
developed with
The National Archives,
provides full access to the
statute book as open data

November 2010

www.tso.co.uk/opendata



Legislation.gov.uk is the result of a complex project to draw together the Office of Public Sector Information website, providing legislation as enacted, with the UK Statute Law Database, providing revised legislation, to create an integrated service that simplifies access to legislation and provides full access to the statute book as open data.

This is one of the first Government services to fully meet the linked data principles advocated by the recently established Public Sector Transparency Board.

“Legislation.gov.uk presents complex information in a clear and intuitive way. This groundbreaking work puts democracy at the heart of legislation and makes a major contribution to the government’s transparency agenda.”

Lord McNally, Minister of State and Deputy Leader of the House of Lords

TSO managed the development and maintenance of the Office for Public Sector Information (OPSI) website as part of an overall publishing contract. OPSI, now part of The National Archives (TNA) is at the heart of information policy in the UK, delivering access to and encouraging re-use of public sector information including legislation. The OPSI website had more than 1.5 million unique users and 50 million hits a month, the vast majority searching for legislation.

The new official home of legislation on the internet provides access to one of the world’s most comprehensive databases, bringing together content from both websites and providing effective navigation around the documents and between versions including an interactive browse and timeline facility. There are also multiple options for downloading and printing the legislation. In developing legislation.gov.uk, TSO worked with The National Archives to integrate semantic web technologies and to provide access to the underlying data through a powerful Application Programming Interface (API) which enables the data to be re-used.

The new website consists of 6.5 million distinct web pages, a further 6.5 million PDF documents and 1.8 billion RDF triples of Linked Data. It has been hailed as groundbreaking work, putting democracy at the heart of legislation and making a major contribution to the Government’s transparency agenda.

TSO’s open data services

TSO’s end-to-end open data services can help you to publish your information as open data to meet transparency requirements. Our experts create tools and processes that allow your data to be created in a structured way; enrich your data using text engineering techniques; convert your data into formats to publish as linked data on the web and provide and host web environments, allowing both citizens and machines to access your data.

Find out more at www.tso.co.uk/opendata

How legislation.gov.uk meets the draft principles of the Public Sector Transparency Board:

Principle	Legislation.gov.uk
<p>Public data policy and practice will be clearly driven by the public and businesses who want and use the data, including what data is released when and in what form – and in addition to the legal Right to Data itself this overriding principle should apply to the implementation of all the other principles.</p>	<p>Legislation.gov.uk has evolved from experience gained in publishing enacted legislation through www.opsi.gov.uk and revised legislation through www.statutelaw.gov.uk. Extensive research was carried out with users of both websites as part of the process of defining the new service. A key driver has been to create a single point of access for legislation in the UK bringing together these two previous sources and providing the underlying data all in one place.</p>
<p>Public data will be published in reusable, machine-readable form – publication alone is only part of transparency – the data needs to be reusable, and to make it reusable it needs to be machine-readable. At the moment a lot of government information is locked into PDFs or other unprocessable formats.</p>	<p>Legislation is available on legislation.gov.uk in both human and machine readable formats. Users can browse the content online in accessible HTML format or download in accessible PDF format. The underlying data is also available in re-usable XML, RDF and Atom formats through a published RESTful API. A single URI scheme is adopted for both human and machine readable formats to ensure that access is as easy as possible.</p>
<p>Public data will be released under the same open licence which enables free re-use, including commercial re-use – all data should be under the same easy to understand licence. Data released under the Freedom of Information Act or the new Right to Data should be automatically released under that licence.</p>	<p>Data released through the legislation.gov.uk API is released under an open licence which enables free re-use including commercial re-use.</p>
<p>Public data will be available and easy to find through a single easy to use online access point (data.gov.uk) – the public sector has a myriad of different websites, and search does not work well across them. It's important to have a well-known single point where people can find the data.</p>	<p>The service operates under two domains: legislation.gov.uk for the website and legislation.data.gov.uk for the data, i.e. the data is accessible directly through data.gov.uk</p>
<p>Public data will be published using open standards, and following relevant recommendations of the World Wide Web Consortium. Open, standardised formats are essential. However to increase reusability and the ability to compare data it also means openness and standardisation of the content as well as the format.</p>	<p>All of the content on legislation.gov.uk is published using open standards and following all relevant recommendations of the World Wide Web Consortium. A choice of open standard formats is available to maximise access and re-use; accessible HTML, accessible PDF, XML and the recommended linked data format: RDF. The XML is published using an open XML schema: the Crown Legislation Markup Language (CLML).</p>
<p>Public data underlying the Government's own websites will be published in reusable form for others to use – anything published on Government websites should be available as data for others to re-use. Public bodies should not require people to come to their websites to obtain information.</p>	<p>All of the content on legislation.gov.uk is available in reusable form for others to use, either in situ through the API or by downloading the underlying data in XML format to use offline.</p>
<p>Public data will be timely and fine grained – Data will be released as quickly as possible after its collection and in as fine a detail as is possible. Speed may mean that the first release may have inaccuracies; more accurate versions will be released when available.</p>	<p>The XML published on legislation.gov.uk is fine grained, providing a rich format for others to re-use. The XML schema is published and fully documented. New content is published daily and simultaneously in all formats. Established publishing processes ensure that the initial timely release is an accurate version of the published content in all formats, including the fine grained XML.</p>

Principle	Legislation.gov.uk
<p>Release data quickly, and then re-publish it in linked data form – Linked data standards allow the most powerful and easiest re-use of data. However most existing internal public sector data is not in linked data form. Rather than delay any release of the data, our recommendation is to release it ‘as is’ as soon as possible, and then work to convert it to a better format.</p>	<p>The publishing process for legislation is well established and was designed from the ground up using open data standards. This enables simultaneous publication in many different data formats. Therefore, legislation.gov.uk is an example of how you can successfully release complex public data quickly in a useful linked data form.</p>
<p>Public data will be freely available to use in any lawful way – raw public data should be available without registration, although for API-based services a developer key may be needed. Applications should be able to use the data in any lawful way without having to inform or obtain the permission of the public body concerned.</p>	<p>The raw data behind legislation.gov.uk is available without registration including the API-based services. Applications can consume the data without having to inform or obtain the permission of The National Archives or their publisher TSO.</p>
<p>Public bodies should actively encourage the re-use of their public data – in addition to publishing the data itself, public bodies should provide information and support to enable it to be re-used easily and effectively. The Government should also encourage and assist those using public data to share knowledge and applications, and should work with business to help grow new, innovative uses of data and to generate economic benefit.</p>	<p>Legislation.gov.uk also publishes detailed guidance on how to re-use the content. The URI scheme, API and XML schema are all fully documented to provide support to enable legislation to be re-used easily and effectively.</p>
<p>Public bodies should maintain and publish inventories of their data holdings – accurate and up-to-date records of data collected and held, including their format, accuracy and availability.</p>	<p>The legislation.gov.uk site is a live inventory of The National Archives’ database of legislation in both enacted and revised formats. Future plans include the addition of provenance information using an open provenance data model to give re-users further information on how each of the formats available are created so that an appropriate assessment can be made on the accuracy.</p>

Want to find out more?

To find out more about TSO’s open data services please visit www.tso.co.uk/opendata

To discuss your open data publishing requirements with a TSO expert please email opendata@tso.co.uk

To keep up to date with the latest information from TSO, follow us on www.twitter.com/TSOSolutions

75% recycled
This is printed on 75% recycled paper



CONFIDENTIALITY STATEMENT

The contents of this document together with all other information, data, materials, specifications or other related documents provided by Williams Lea (“WL”) (together “materials”) shall be treated at all times by the recipient as the confidential and proprietary information of WL. The recipient shall not disclose any such materials to any third parties without the express, prior written approval of WL. Where such express approval is granted by WL, the recipient shall ensure that all third parties to whom disclosure is made shall keep any such materials confidential and shall not disclose them or any part of them to any other person. All intellectual property rights in the materials shall remain the property of WL, or its third party licensors, and are protected by copyright.

© 2008-09 Williams Lea Group

DISCLAIMER

This document may be incomplete without reference to any oral briefing provided by WL, reflects current conditions and WL’s views as of this date and is subject to correction or change at any time. Although the information contained in this document is believed to be accurate in all material respects, neither WL nor any of WL’s advisers, agents, officers or employees accepts responsibility or liability for or makes any promise, representation, statement or expression of opinion or warranty, express or implied, with respect to the accuracy or completeness of the content of this document (to the extent permissible by law) unless and save to the extent that such promise, representation, statement or expression of opinion or warranty is later expressly incorporated into a legally binding contract.



Part of the Williams Lea Group