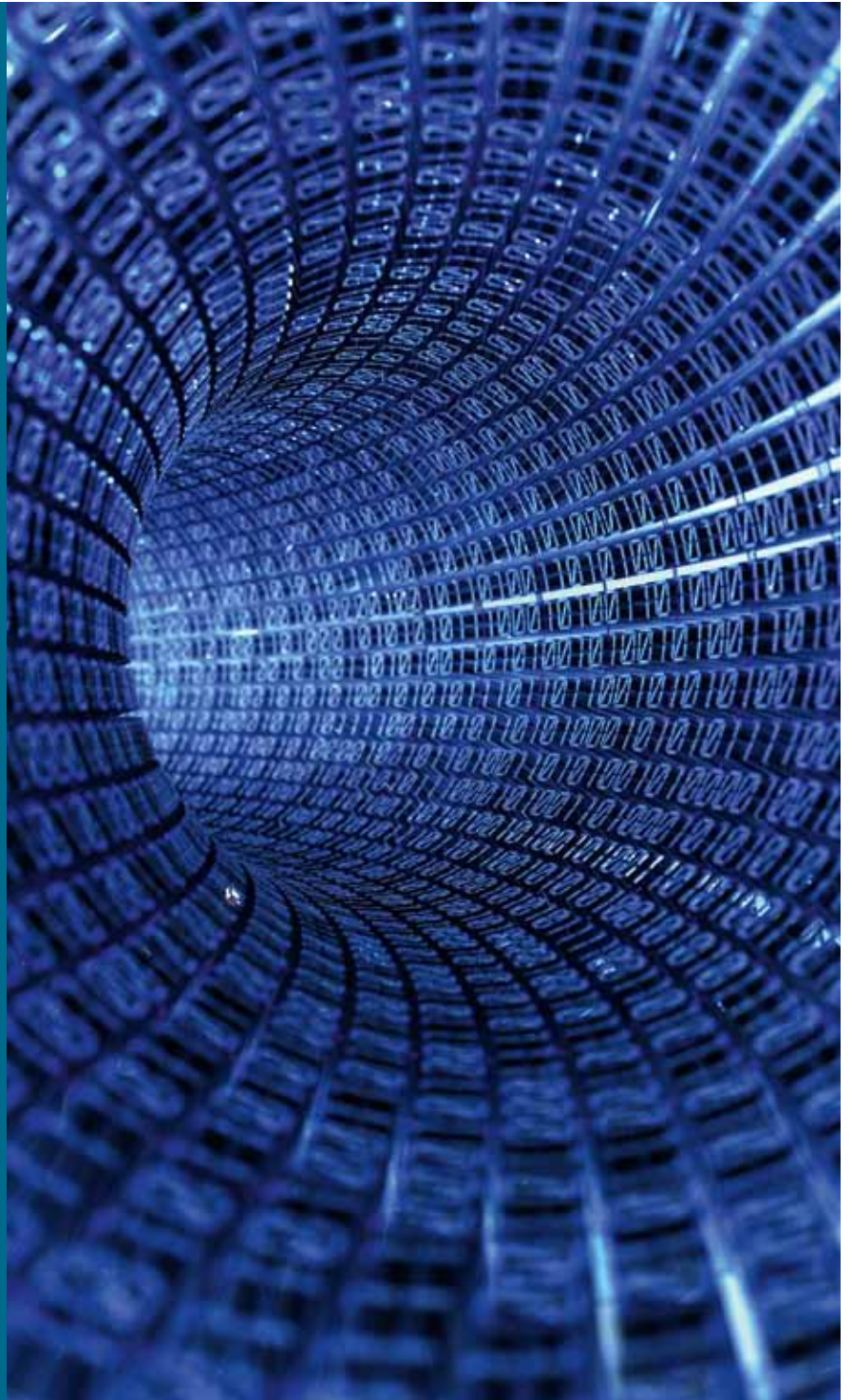


Public Data Transparency

How TSO can help
government organisations
to open up data



July 2010

www.tso.co.uk/opendata



TSO has been at the forefront of working with public sector clients to open up public data. Our established solutions can help you to comply with the requirements of the new Public Sector Transparency Board, ensuring your data is open and re-usable to citizens and businesses.

On 31 May 2010, Prime Minister David Cameron wrote to Government departments outlining the new Government's commitment to opening up government data to improve transparency. The Prime Minister has established a Public Sector Transparency Board which will drive forward the Government's transparency agenda, making it a core part of all government business and ensuring that all Whitehall departments meet the new tight deadlines set for releasing key public datasets. In addition, it is responsible for setting open data standards across the whole public sector, listening to what the public wants and then driving through the opening up of the most needed data sets.

The Public Sector Transparency Board has published draft data principles which include the need to publish public data in machine-readable form, unlocking it from PDFs and other unprocessable formats and making it reusable. The principles recommend that data is re-published in linked data forms which allow the most powerful and easiest re-use of data. Public bodies should also actively encourage the re-use of their public data, working with businesses to grow new, innovative uses of data.

How TSO can help

TSO has already been helping public sector clients to publish data in re-usable linked formats; the publication of data in these new formats is a natural extension of our multi-format publishing solutions. We have worked with clients such as The National Archives to open up data in the Gazettes and legislation and we host data.gov.uk, the website created to provide a way in to the wealth of government data.

TSO's open data solutions will enable you to ensure that all your published data complies with the requirements set out by the Prime Minister. Our experts will: 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 to improve transparency and engagement.

TSO will work with you to open up your published data in three steps:

TSO developed an MS Word based drafting tool for The National Archives and relevant government departments to create secondary legislation. The template contains all elements in the development of a Statutory Instrument, including typesetting styles and metadata, enabling online validation, reducing errors and allowing a quicker publishing process. The template is used to create structured XML and print ready PDFs, with website content generated automatically from the source XML.

TSO has enriched the data in more than 250,000 notices on the London Gazette website, using GATE to apply RDF and create machine readable data. The information in the London Gazette is now available in a range of formats including print, XHTML, XML and RDF, and the data is versatile enough to re-use in combination with other data. An example of how Pollution Notices from the London Gazette have been combined with other information can be found here:
www.london-gazette.co.uk/demo/

TSO's RDF open data platform provides:

- A highly scalable and secure environment for hosting RDF data for re-use
- A number of standards based data access and query options, including SPARQL for easy integration with new and existing applications
- A software as a service pricing model allowing fast deployment at a reasonable cost.

1. Creating and capturing your data in a structured way

To enable your data to be reused it must be created in a structured way. TSO's experts will first work with you to understand the requirements for your data. Once this is established we will create tools and processes that capture content in the most appropriate and efficient way, depending on the nature of your data, the number of users capturing it and their skill set. We have created tools and processes using MS Office templates, XML authoring tools and web portals for content validation and submission.

2. Enriching your data to enhance and transform

To make your data open, linked and reusable it must be unlocked from the usual print and HTML formats, which are only readable by humans, and converted into linked, machine readable formats. TSO's experts will use text analysis frameworks such as GATE and UIMA to automatically enrich your content and extract information from it, enabling it to be converted into open formats including RDF (the recommended format for linked data) XML, XHTML + RDFa and ATOM. The approach can be integrated with templates to improve the automation of your data publishing.

3. Delivering open, reusable data to engage citizens

The ultimate aim of opening up access to public data and creating it as machine readable, linked data is to enable the creation of new, more useful information applications for the public. TSO provides a scalable and secure environment for hosting your RDF data, making it easy for publishers and developers to build the next generation of semantic web applications. TSO's platform is built on 5Store, a highly scalable, clustered, commercial database storage and query engine that holds RDF data, designed and developed by Garlik who are leaders in semantic technology. Several APIs are available to extract the data, including a SPARQL endpoint.

TSO can also host the resulting web applications with our website hosting service, which is certified to ISO 27001 Information Security Standard.

Find out more

To discuss your requirements with one of our experts please email opendata@tso.co.uk

To find out more about how TSO can help to create, enrich and engage citizens with your data, and to read the full case studies visit www.tso.co.uk/opendata

To find out about OpenUp, TSO's open data challenge, which asks data users to let us know what public data they need and how they would like it presented, visit www.tso.co.uk/openup

To keep up-to-date with the latest information from TSO follow us on Twitter at 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