



Department  
for Business  
Innovation & Skills

OPEN DATA STRATEGY

2014-2016

JULY 2014

<b>OPEN DATA STRATEGY .....</b>	<b>1</b>
Introduction .....	3
Wider Government Strategy .....	4
Principle 1: Open Data by Default .....	5
Sector Boards .....	6
Public Data Group .....	6
Breakthrough Fund .....	7
Open Data by Default: Commitments .....	8
Principle 2: Quality and Quantity .....	9
Quality and Quantity: Commitments .....	10
Principle 3: Usable by All .....	11
Usable by All: Commitments .....	11
Principle 4: Releasing Data for Improved Governance .....	13
BIS's Management Information Strategy .....	14
Principle 5: Releasing Data for Innovation .....	15
Open Data Challenge Series .....	15
Intellectual Property (IP) Data .....	15
Technology Strategy Board .....	15
Open Data in Science and Research .....	16
Public Data Group .....	17
Annexes: Content .....	18
Annex 1: Prototype: Map of BIS Datasets .....	18
Annex 2: Case Studies .....	21
Annex 3: List of Datasets Provided on a Statutory Basis .....	28

# Open Data Strategy

The Strategy is structured around the [G8 Open Data Charter](#) principles, published in summer 2013. All G8 members including the UK Government signed up to these principles on the 18th June 2013. There are 5 core principles:

- Principle 1: Open Data by Default
- Principle 2: Quality and Quantity
- Principle 3: Usable by All
- Principle 4: Releasing Data for Improved Governance
- Principle 5: Releasing Data for Innovation

Case Studies covering some of the most significant areas of data holding in BIS and Partner Organisations are included in the Annexes.

Outstanding commitments for BIS and Partner Organisations identified in the [National Information Infrastructure](#) (published October 2013) have been included under the relevant principles.

## Introduction

Alongside other government departments, BIS published its first Open Data Strategy in June 2012 to coincide with the launch of the Open Data White Paper. Since then BIS has done a lot to support the agenda. As a result, we have produced our second Open Data Strategy, which updates and replaces the original Strategy and highlights new commitments and progress to date.

Economic Growth is the priority of the Department and we recognise that open data has an important role to play. Open data contributes directly through the economic opportunities it creates and through its role in delivering the reforms we are introducing. Whilst it is difficult to cite an exact figure because this is an emerging market, there have been a number of reports published on the value of open data to the economy. A [Deloitte report](#) (published May 2013) for example cites £1.8bn of direct benefits, £6.8bn of indirect benefits and a clear potential for significant development.

BIS is fully committed to the open data agenda and is implementing and embedding it in the most appropriate way across all aspects of our work. BIS is playing a major role in supporting and developing the agenda both as a data holder and as a policy lead in areas where BIS is the policy owner. BIS and its Partner Organisations hold a vast amount of data covering a wide variety of subject areas from skills to archaeology to geospatial data to biochemistry to all companies' accounts. This data is frequently produced by a range of world leading centres of excellence - be it the Met Office or the Diamond Synchrotron. Over 20% of the National Information Infrastructure represents data within BIS and even within that forms much of the core reference data for creating linked data across the public sector.

BIS's [Information Economy Strategy](#) recognises the importance of open data, as well as the importance of ensuring businesses have the capability to access and reuse this public sector information to link to economic growth. Published in October 2013, [Seizing the Data Opportunity](#) sets out a range of measures to address key areas of the UK's data capability. As such, open data is benefiting from a range of policy interventions as the Government seeks to make data a

key driver of UK prosperity. From the activities of the Technology Strategy Board to the support for data analytics to the creation of the world leading Open Data Institute, BIS is supervising the spending of £100ms to support this agenda and make the UK a global centre of excellence in applying this data revolution to practical benefit.

Underlying all of this is our overarching principle to engage with stakeholders and end-users to determine which data will be most useful to them. This is shown by the numerous engagement activities and challenges such as the Open Data Challenge Series, recent user engagement events in both the Higher and Further Education Sectors, and the Ordnance Survey's GeoVation series. BIS will continue to work through stakeholder groups to identify data sets which add most value for priority release. From past experience, we know that stakeholders tend to be most interested in data that:

- Provides transparency;
- Informs consumer or end user choice, with the effect of making the market perform efficiently and driving up standards; and
- Can be used onwards to create revenue streams, thus contributing to economic growth.

## Wider Government Strategy

Cabinet Office are the lead Department for the transparency and open data agenda. They have identified the following priorities for 2014:

- Get high quality open data out of government and into the public's hands;
- Bring the power of open data to a wider audience; and
- Maintain Britain's global position as a leader on open data and transparency.

Since the publication of our last Open Data Strategy there have been a number of documents which outline the next steps for the Government's strategy towards open data.

- The [Shakespeare Review Public Sector Information](#) (published May 2013)
- The UK [Government's response](#) to the Shakespeare Review (published 14 June 2013)
- Open Government Partnership UK [National Action Plan](#) (published 31 October 2013)
- National Information Infrastructure [narrative](#) and [datasets](#) (published 31 October 2013)
- G8 Open data Charter [National Action Plan](#) (published 1 November 2013)

## Principle 1: Open Data by Default

BIS and its Partner Organisations hold large amounts of data. Some of the data collected is for statistical purposes, for others collection and dissemination of data is a core statutory function, while other data is collected as part of the day to day running of the organisation. BIS publish a number of National and Official statistics on a regular basis relating to businesses, employment relations, higher and further education which are published on: <https://www.gov.uk/government/organisations/department-for-business-innovation-skills/about/statistics>

Much of BIS' data is held by Partner Organisations or organisations whom we sponsor but for which BIS is not directly responsible. We work closely with our Partner Organisations to instil open data principles and encourage data release and also to develop sustainable data models. For many of our Partner Organisations data is core to their function. For example the Met Office collects more than 10 million weather observations per day which enables it to create tailored forecasts. Without this data it could not achieve its objectives.

In all instances organisations have made major efforts to make their data as accessible as possible. However, as with other parts of Government, there are instances where Partner Organisations are not always able to allow complete and unrestricted access to their data. We support these exceptions case by case, based on the specific needs and sensitivities of each data provider, including:

- World class data organisations like Met Office and Ordnance Survey - that are integral to the UK economy but also have international reach and growth opportunities themselves;
- Working with the needs of the academic community to nurture research while maximising its reach; and
- Protecting individual and consumers personal data as sensitive data sets are combined.

All data published on the BIS website is covered by the Open Government Licence (OGL) encouraging its use and re-use freely and flexibly. We will work with Partner Organisations to ensure that they are also releasing as much of their data as possible via the OGL.

We index all published datasets on data.gov.uk to support a single index of public authority datasets. BIS and its Partner Organisations have published 1,300 datasets on data.gov.uk. BIS is the third highest publishing government department on data.gov.uk.

We recognise that for government data to be better exploited there needs to be clarity about the data we hold. In 2013 as part of the cross-government exercise to develop the National Information Infrastructure BIS and seventeen of our Partner Organisations compiled inventories of our unpublished datasets. These inventories were published on data.gov.uk on the 31st August 2013 to give members of the public the opportunity to highlight the economic and social impacts if these datasets were made openly available. As a result of the exercise information on an additional 45 core BIS datasets and 400 Partner Organisation datasets are now in the public domain.

This exercise led to the first iteration of the National Information Infrastructure which was published by Government on the 31st October 2013. BIS contributed to the development of this list of the 320 most significant datasets held by government. It contains 60+ datasets from BIS

and the following Partner Organisations: British Geological Survey, Companies House, Land Registry, Met Office, Ordnance Survey, Skills Funding Agency and the UK Commission for Employment and Skills. Datasets collated by the Higher Education Statistics Agency for BIS are also included.

We intend to make all data identified during this exercise available as open data unless there is a reason to the contrary. Where it is not possible or appropriate to place data in the public domain – for security reasons, commercial confidentiality or as it contains personal information - we will make this justification public. Where we feel that releasing a dataset will have public value we will proactively consider whether, through aggregation and anonymisation, this can be made possible.

Through the exercise the public now has the opportunity to assess the economic and social value for releasing this data. We will monitor this feedback and use it to help prioritise the datasets that we release.

In addition under the Freedom of Information Act the public have a right to request information held in BIS. Since September 2013 revised internal processes have been developed to support the new datasets provisions in the Act. These new provisions require that when a dataset is released in response to an FOI request it will be done in a way that enables the dataset to be used and re-used. When a dataset is released to a requester the dataset will also be published on the BIS website unless it is “not appropriate” to do so and indexed on [data.gov.uk](http://data.gov.uk). A framework for how BIS responds to such requests is available on our website.

## Sector Boards

The **Social Mobility Transparency Board** was created to match and share a much wider range of data on progression through education and into the labour market, to underpin research, policies and initiatives outside government to improve social mobility. The Board met for the first time in October 2012 and was initially chaired by the Deputy Prime Minister. There have been a further four meetings. The Board is made up of a number of Ministers (David Laws, David Willetts, David Gauke, Matthew Hancock), lead officials and members of the research community.

The [Research Sector Transparency Board](#) chaired by David Willetts. This board includes the Government Chief Scientific Adviser, the Director General Knowledge and Innovation and various senior representatives of research funders and the research community. The Board will advise government on how to increase access to research data, with the aim of fuelling new discovery and innovation, and ultimately economic growth and societal benefit. This will complement existing activities to improve access to research publications. Issues it aims to tackle are: providing leadership and building consensus around open data in science; interdisciplinary standards and metadata; and skills and capability.

## Public Data Group

The Public Data Group (PDG) brings together four government trading funds - Companies House, Land Registry, Met Office and Ordnance Survey – that, alongside their other key functions, excel in the collection, management and distribution of vital data sets.

All four members of the Public Data Group have, however, in addition to providing high quality and reliable data for national and international businesses and the public sector, released a large number of significant data sets as open data and work is on-going for potential future

releases. Further information – including future data releases and activities - is provided in the [Public Data Group 2014 Open Data Statement](#) but highlights of the data released include:

- Companies House has launched a free Accounts Data Product that allows free access to all statutory accounts filed digitally - that's access to approximately 1.7 million sets of accounts since its launch in November 2013. This has made Companies House a world leader in providing iXBRL accounts data free of charge.
- The Land Registry has made available price paid data - dating back to 1995 - in Excel, CSV and linked data formats for records. This means that there are available more than 19 million definitive records of monthly residential property price data.
- Met Office makes available real time observational data including site observations and radar map layers. It also makes available forecast information that includes both daily and 3 hourly site specific forecasts, updated hourly, map layers, text and charts. This data is all accessible in a format suitable for application developers under Open Government Licence.
- Ordnance Survey provides OS Open Data which is formed of over 10 quality assured, continuously updated products, from postcodes and boundaries to digital maps, providing a one-stop shop to start using its products.

## Breakthrough Fund

BIS are supporting wider government and local authorities to make more data available through its Breakthrough Fund. The fund was established to help government departments, government agencies and local authorities to accelerate open data proposals where there are short-term financial barriers to the data release. Funding is available in financial years 2013-14 and 2014-15. Further information can be found at:

<https://www.gov.uk/government/publications/breakthrough-fund-and-release-of-data-fund>

Over £2m was allocated in 2013-14 to support a range of initiatives in central government and local government, including:

- Funding to support the development of a technology platform for accounting taxonomies to be made openly and freely available.
- Funding new software to allow the release of council data as open data, and for this data to be interrogated by citizens.

## Open Data by Default: Commitments

We Commit to:	BIS	Partner Organisations
<b>Working with the remainder of the BIS Partner Organisations to publish their inventories</b>		August 2014
<b>Maintaining the comprehensiveness of the inventories already published by embedding internal processes to identify new datasets.</b>	On-going	On-going
<b>Making public our assessment of the impact of the release of datasets</b>	July 2014	
<b>Highlighting those datasets that we provide on a statutory basis.</b>	See Annex 3 for BIS datasets	September 2014
<b>Highlighting those datasets which we consider falls under our public task</b>	See Annex 3 for BIS datasets	September 2014



## Principle 2: Quality and Quantity

Across all data release BIS is working to the principles of: coherence with wider government strategy; protecting individual data; statistical robustness and validity; easy availability and transparency.

Many of the statistics that BIS publishes are designated as 'National Statistics' and are produced in accordance with the principles set out in the Code of Practice for Official Statistics. By complying with this Code we are ensuring that the statistics we deliver are based on:

- **Knowing our customers** – the content of our resources and products, and the supporting information we provide will be informed by an understanding of our customers' needs.
- **Informing our users** – we will help customers and users to understand our products by publishing or referring to supporting, background or educational information and material.
- When required BIS complies with international standards for data. This includes the Met Office's compliance with the INSPIRE legislation on spatial data.

We are also committed to releasing the data underlying BIS reports and publications to enable and encourage innovative re-use.

BIS's transparency and open data implementation is monitored by the Transparency Board chaired by our transparency 'champion', Director General Finance and Commercial, Howard Orme. It is responsible for overseeing BIS commitments, promoting the transparency agenda, highlighting best practice and challenging the non-release of data for core BIS and Partner Organisations.

BIS's Knowledge and Information Team is responsible for the day-to-day operational running of the transparency publication programme for core BIS and Partner Organisations. As part of the Cabinet Office Quarterly Written Ministerial Statement exercise the team continually monitor releases and assure quality. The team have produced processes and guidance to ensure that data releases meet transparency best practice.

In December 2013 a joint core BIS and Partner Organisation Knowledge and Information Strategy was agreed. This strategy focuses on how we manage information, ensures that the information we collect and hold is fit for purpose, and promotes transparency. The strategy is based on the HM Government's 7 Information Principles. Principle 4 'Information is Standardised and Linkable' and Principle 6 'Public Information is Published' relate specifically to the open data agenda. Other principles including Principle 1 'Information is a Valued Asset' and Principle 2 'Information is Managed' are pre-requisites to enable BIS and Partner Organisations to adhere to this Open Data Strategy.

## Quality and Quantity: Commitments

We commit to:	BIS	Partner Organisations
<b>Reporting through the Quarterly Written Ministerial Statement (QWMS) on the comprehensiveness of our unpublished dataset</b>	July 2014	July 2014
<b>Reporting through the QWMS on the work we are undertaking to engage new users in helping to define further uses for their data.</b>	July 2014	

## Principle 3: Usable by All

We publish information, including datasets, on the BIS website and continually seek to exploit new digital channels to promulgate our messages in as many ways of possible to as far reaching an audience as possible. In addition to our website we make use of flickr, YouTube and Twitter to reach new audiences.

BIS's [Digital Strategy](#) published in December 2012 sets out how we will redesign our information, transactions and digital engagement around the needs of businesses, students, employees and consumers. This strategy will contribute to economic growth through improved policy outcomes, easier to use services, more open policy making and greater efficiencies. We are committed to the ambition, set out in the Government Digital Strategy to deliver digital services that are so straightforward and convenient that those who can use them will choose to do so, and to provide assisted digital support for those who cannot.

We are supporting our Partner Organisations with high volume transactions and those with exemplar projects to have processes in place to provide us with regular transaction data. This data includes cost per transaction and digital take up and is published each quarter on Gov.uk.

Through the BIS Transparency Mailbox ([transparency@bis.gsi.gov.uk](mailto:transparency@bis.gsi.gov.uk)) we listen to our customers' needs and issues. In cases where data has not been released in a suitable format for the customer we endeavour to generate the data in suitable alternative formats.

Currently most of our data is published to 3-star standard, as we have deemed this format to strike the right balance between usability and cost of releasing data for public consumption. However, some data published by Companies House and the Land Registry is to 5-star standard, as are the BIS organograms.

The Open Data Institute launched its certification process at the Open Government Summit on the 31st October 2013. The purpose of the certificate is to convey the quality level of the data, (ranging from 'Raw' to 'Expert'), as well as a human and machine-readable description of the data being released.

### Usable by All: Commitments

We commit to:	BIS	Partner Organisations
<b>Embedding internal processes to ensure that data holders regularly reconsider the use cases for their data</b>	On-going	
<b>BIS will put all their datasets currently available under an Open Government Licence through the ODI's open data certification process, giving priority to the datasets included in the NII, and make the outcome available through data.gov.uk.</b>	December 2014	

We commit to:	BIS	Partner Organisations
<b>Working with the Cabinet Office, the Open Data User Group and other stakeholders to develop a programme of engagement with users to encourage greater assessment of datasets</b>	On-going	On-going
<b>Develop an interactive roadmap to BIS's data for use on our website – prototype is in Annex 1</b>	September 2014	September 2014

## Principle 4: Releasing Data for Improved Governance

BIS is committed to providing the public with information on the performance and productivity of the Department. The public are entitled to understand how their money is being spent and whether they are getting good value. We are committed to the timely and accurate delivery of all Cabinet Office accountability commitments for data release. To enable comparability of data across the public sector we will meet the standards that they determine. Current commitments include:

<b>Title of BIS Data Set</b>	<b>Frequency</b>
<b>Financial Spend Data (all spend)</b>	Monthly
<b>Government Procurement Card Spend</b>	Quarterly
<b>Exceptions to Spending Moratoria covering ICT, Publicity &amp; Marketing, Consultancy Spend, Recruitment and Property</b>	Quarterly
<b>Organogram and staff datasets</b>	Twice Yearly
<b>Workplace Management Information</b>	Monthly
<b>Minister expenses, travel and hospitality</b>	Quarterly
<b>Senior Officials expenses, travel and hospitality</b>	Quarterly
<b>Special Advisers gifts, hospitality and meetings with external media</b>	Quarterly
<b>Contract Finder</b>	On-going
<b>Real Time Energy Use</b>	Daily
<b>Quarterly Data Summary</b>	Quarterly

In these cases we aim to be as transparent as possible, processes exist to ensure that requests from policy officials to redact data are challenged, and if necessary are referred up to the Transparency Board. As an example of our on-going commitment to accountability we will continue to exceed the Cabinet Office requirements and publish all of our spend data.

## BIS's Management Information Strategy

All information relevant to our work, from whatever source, can be described as Management Information (MI) and in BIS good MI should enable staff at all levels to make good decisions. Developed with our Partner Organisations our MI strategy sets the ambition to:

- **Enable delivery of our objectives**, supported by effective, timely and cost efficient MI.
- **Continuously improve the value for money we achieve** through a better understanding of costs and results so that we can make more informed evidence-based, decisions.
- **Empower managers** to manage, and staff to deliver through better MI.

High quality open data relies upon sound systems that produce quality data efficiently. BIS has developed the MI Strategy that is focusing internally on how we collect manage and report our MI to produce high quality MI and promote better decision making by focussing on three improvement themes:

- **Planning and Monitoring** - providing direction and enabling better control through MI
- **Ways of Working** - embedding a MI culture and building skills and capability
- **Systems** - using better technology, improving efficiency and driving better insights

## Principle 5: Releasing Data for Innovation

BIS has been responsible for two strategies this year that include consideration of the impact of data release to the UK economy and how the opportunities can be best exploited.

The [Information Economy Strategy](#), published in June 2013, recognised the growth in the amount of data globally and committed that the Government would publish a data capability strategy, developed in partnership with industry and academia, to set out further actions to take leadership and advance our capability in data for growth.

In October 2013 the Government published [Seizing the Data Opportunity: A Strategy for UK Data Capability](#), which delivered on our commitment in the Information Economy Strategy and sets out plans to ensure that the UK is well-positioned to be at the forefront of extracting knowledge and value from data. The strategy covers the importance of skills, infrastructure and tools, and ensuring that data can be shared and linked appropriately, and includes actions that the Government is taking to develop its own data capability.

BIS has instigated a range of initiatives to fuel innovation in the UK through helping companies develop new commercial opportunities. These include:

### Open Data Challenge Series

BIS is providing over £1.2 million to support companies seeking to develop new products and services using open data. The Open Data Challenge Series, run by Nesta and the Open Data Institute, brings together SMEs and start-ups with data providers, industry experts and business leaders to develop business ideas.

In total, there will be 7 challenges addressing key policy areas –including crime and justice, education and housing – and the winner of each challenge will receive £40,000 – 50,000 to commercialise their idea. On 5th February 2014, ‘Check that Bike Ltd’ were announced as the winner of the Crime and Justice Open Data Challenge – the first of the series.

### Intellectual Property (IP) Data

The Intellectual Property Office (IPO) makes IP data freely available through online journals and online registers. Whilst these sources of data provide quite detailed information on IP cases, they do not allow for easy mass analysis.

In May 2014 the IPO published its first analysable [IP dataset](#) for patent data allowing statistical research to be carried out more easily. Equivalent datasets for trademarks and designs are due to be published during 2014/2015.

### Technology Strategy Board

The Technology Strategy Board (TSB) is active across the open data agenda. The TSB Digital programme addresses the adoption and exploitation of information technologies by businesses, with one strand of investment focusing on the creation and exploitation of a supply of clean, structured, accessible and usable data.

TSB is working with the Natural Environment Research Council (NERC) on investing up to £4m in feasibility studies which use environmental data to develop new solutions to business problems. The fund will support solutions and services that offer specific commercial benefits and/or limit business risks and increase resilience by integrating environmental data with data

from other sources. The focus is on solving problems by finding innovative ways to combine data sources, either multiple environmental datasets or environmental and 'business' data.

The project has aggregated access to datasets from Met Office, Ordnance Survey, NERC, Environment Agency and others, and wants this data to be used to address a challenge felt by businesses in an area such as: Energy, Transport, Built Environment, Agriculture and Food, and Financial Services.

### **Innovation Vouchers**

The Government has made up to £1.5 million available for companies to apply for innovation vouchers from the Technology Strategy Board in 2013/14 and 2014/15. Businesses can apply for a voucher of up to £5,000 to work with a new supplier to develop business ideas using open data.

Over 100 vouchers have already been allocated to date.

### **Connected Digital Economy Catapult**

With over £50 million of funding from the Technology Strategy Board, the Government has created the Connected Digital Economy Catapult. The catapult will bring together innovative businesses with researchers and academia, in order to help UK businesses take advantage of the opportunities offered by the huge growth in the volume and variety of data. The catapult will help develop and promote the UK's most creative and innovative digital businesses.

### **Open Data Institute**

The Open Data Institute, established with £10 million of match funding from the Technology Strategy Board, was officially opened in December 2012 to unlock and stimulate demand for data.

The Institute has already welcomed over 2,000 UK and international visitors. Alongside training packages to support both civil servants and software developers the Institute is incubating start-ups.

### **Open Data in Science and Research**

The essence of the scientific process has long been for scientists to publish their ideas and theories, together with the data on which they are based, so that they are available to other scientists to replicate experiments and measurements, to test associated theories and to re-use data in novel ways.

Expanding access to research outputs, both publications and data, is a key aspect of the Government's transparency agenda. Increased transparency in research has the potential to be a significant driver of economic activity, as easy access to knowledge, information and data can improve the sharing of new ideas, encourage innovation and create new growth opportunities. It also offers significant social benefits by spreading knowledge and raising the prestige of UK research. The Research Sector Transparency Board has been established to consider ways in which transparency in research can be a driver for innovation and discovery while furthering the UK's recognised excellence in science.

The Government committed, in its [response to Ian Hargreaves's](#) review of intellectual property, to facilitate data mining for non-commercial research of copyright material, including academic research papers. A new copyright exception for text and data mining for non-commercial



research came into force on 1 June 2014. This could have substantial benefits, for example in tackling diseases. The [Government accepted all main findings of the Finch Group](#) report on access to research publications, and committed to working with partners, including the publishing industry, to enable free and open access to publicly-funded scientific research to maximise social and economic benefits and aid the development of new research by 2014. A 'one year on' progress report following a final Finch Group meeting held on 24 September 2013, attended by David Willetts, has been published by RIN (see <http://www.researchinfonet.org/wp-content/uploads/2013/02/Final-version.pdf>).

The Research Councils are proposing that the researchers they fund should enable Open Access (OA) by publishing their research papers via the publisher's journal's website immediately at the time of publication, which may require the payment of an 'Article Processing Charge (APC)' to the publisher (so called 'Gold' OA) and requiring researchers to deposit the version of the published paper accepted for publication, that is including all changes resulting from peer review in an open access repository at or around the time of publication. In some cases, where Gold OA has not been undertaken, repository copies may only be accessible after an embargo period (so called 'Green' OA). This OA policy was launched in April 2013 and includes provision for the systematic funding of APCs. The compliance target for the first year 2013/14 is 45% Gold and Green. To date performance has varied across sectors with some excellent examples such as the Biotechnology and Biological Sciences Research Council and Medical Research Council's support for UK [PubMed Central](#) which contains 25 million + abstracts and requires an embargo period of only six months and the Economic and Social Research Council's [Research Catalogue](#).

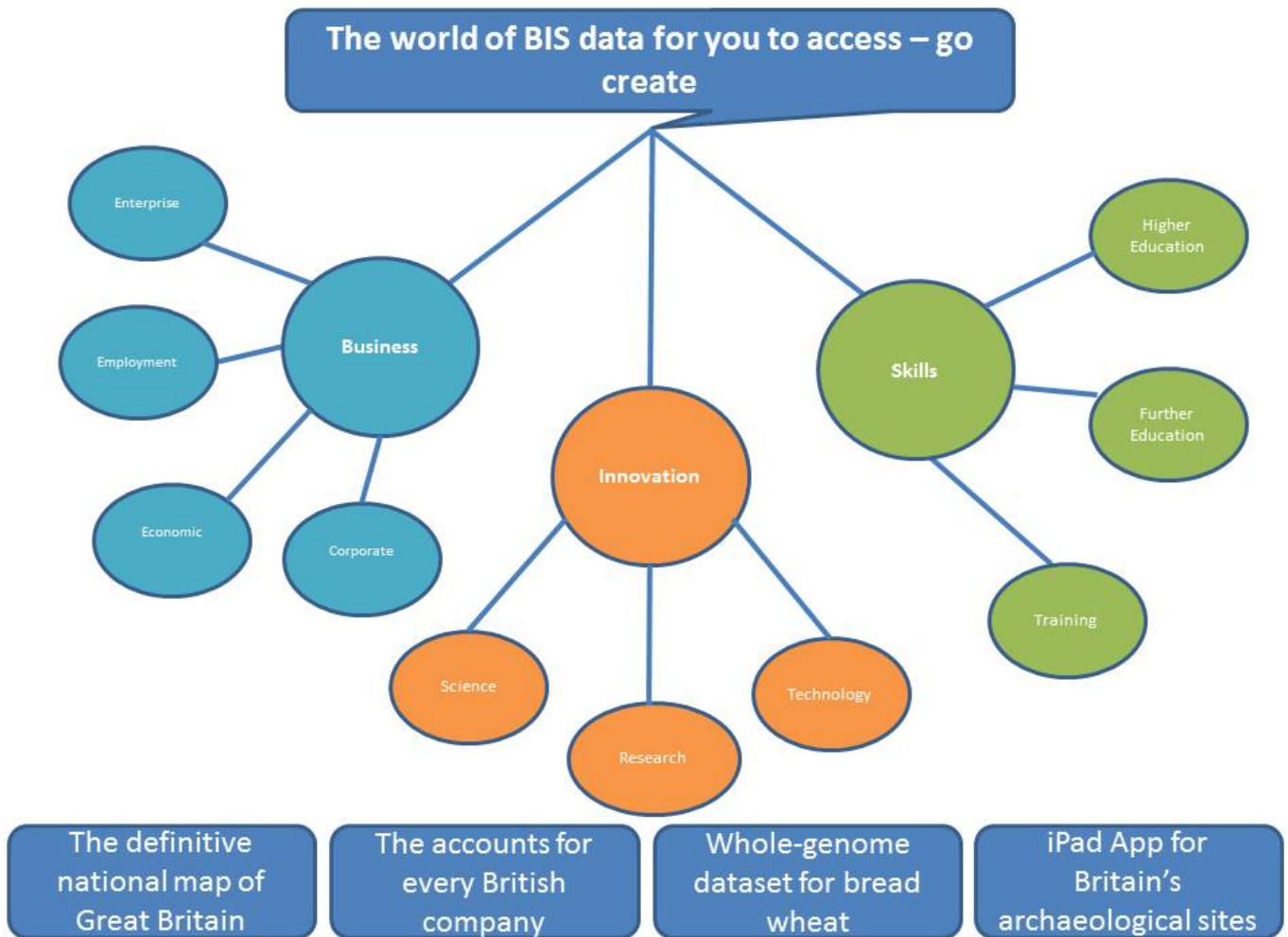
To build on the areas of good practice the Research Councils have invested £2 million in the development of a UK 'Gateway to Research' formally launched on 4 December 2013 - <http://www.rcuk.ac.uk/media/news/42000-research-projects-available-to-business-and-the-public-gateway-launched/>. This allows ready access to Research Council funded research information including published articles and repositories and related data; it has been designed so that it can also include research funded by others in due course. The Research Councils will work with their partners and users to ensure information is presented in a readily reusable form, using common formats and open standards. This Gateway will provide access to the wide body of research undertaken in the UK and its associated data through a single portal, enabling researchers, companies and the public to identify research and results that may be of interest to them. It will facilitate greater networking between the science base and industry and should be of particular benefit to SMEs who lack the resources to monitor developments in the science base closely. It will also demonstrate to potential inward investors the scope and strengths of the UK science base and provide a key web interface for links to the EU's Open data initiatives and elsewhere. A beta version was released in December 2012 and the [alpha version](#) in December 2013.

## Public Data Group

Public Data Group members are also undertaking a range of activities to support the exploitation and development of released data. These include initiatives such as the Ordnance Survey's GeoVation series of themed challenges using geographic data and their OS Open Data Master Classes and Met Office's provision of industry standard open data through Met Office DataPoint and by involvement in hack and open data events such as the NASA Space Application Challenge events and forthcoming "Connected Exeter" event. Further information on the activities of PDG are included in the [Public Data Group 2014 Open Data Statement](#).

## Annexes: Content

### Annex 1: Prototype: Map of BIS Datasets



## Annex 2:

### Case Studies:

#### Further Education

Government wants to ensure that learners and employers have the information necessary to choose the provider that best meets their learning and training needs. By empowering learners and employers to make informed choices through better access to information, we create competition and incentives for providers to be responsive to learners and employers.

FE Learners now have access to their Personal Learning Record (PLR), a web service where learners can store and view their learning achievements. A key aim of the PLR is to help pupils, students, trainees and learners make their own informed decisions about their own future and career aspirations. The PLR also helps to improve the flow of learning-related information about pupils and learners from schools to colleges and other education providers and ultimately employers, leading to more effective information-sharing and better decision-making.

To help learners and employers make choices we publish data on key performance indicators on the [FE Choices comparison website](#) (qualification success rates, employment and education destinations, learner satisfaction ratings and employer satisfaction ratings). Since 2012 this data has been integrated with the National Careers Service and is available, free of charge, to third party organisations from the FE Data Service website.

Learner and employer satisfaction data are collected and published as part of key performance information set on the FE Choices comparative website. FE Choices is the official site for comparing the performance of further education colleges and other organisations that receive Government funding to educate and train people over the age of 16.

Research has established what data learners and employers say they need to make informed choices between FE providers. Most of the information is already available but it needs to be more accessible, comprehensive and consistent. Government and the FE sector are working together to improve the availability of information.

Government already collects and publishes on a consistent basis some of the key information learners and employers say they need to make choices. Government will continue to publish in one place key comparative performance information: qualification success rates, OFSTED inspection results, learner destinations and learner and employer satisfaction. But we want to make this more useful and more accessible for users, so we will:

- Wherever possible, make available quality information at more granular levels, i.e. about the specific subject or course rather than just the provider or broad sector subject area.
- Work with providers and users to improve the efficiency, timeliness and usefulness of information collected by learner and employer satisfaction surveys.
- Make more of the data available in accessible formats so that it can be analysed and re-presented by non-government organisations.

Many learners and employers tend towards 'local' information sources such as college websites rather than national ones. The FE sector is focussing on making information more comprehensive and easily comparable. They will:

- Create a "common information set" of information learners and employers say they need, presented in a standard format on each provider's website to facilitate comparison.
- Highlight (as part of the common information set) a website link to the FE Choices provider quality data to show learners where they can compare performance information for different providers.
- Seek greater completion, by providers, of the course information available to learners on the National Careers Service website.

In early 2013, BIS completed an Open Data Review to look at their approach to making FE and skills data available to the public, and made recommendations for improvement. The review was asked to look at how BIS could shift the focus from presenting data itself to disseminating the raw data on FE and skills and encouraging third parties to use and re-use the data innovatively and imaginatively. As part of an implementation plan to deliver against its recommendations we are working towards producing a unit level file of learner aims that meets the challenge of retaining the economic value of the data while protecting the privacy of those supplying the data. We are increasing transparency through enhanced access of our data such as interactive tools enabling drill down/bespoke table and export functions as well as encouraging greater re-use such as developing APIs for data underpinning FE choice.

## Annex 2:

### Case Studies:

#### Higher Education

BIS is working with its Partner Organisations to roll out the transparency agenda. For Higher Education the funding has been put into the hands of students. This means that it is crucial that they have access to high-quality, easy to use information.

To help prospective students make informed choices about higher education we asked institutions to provide, from September 2012, a standard set of information, at course level and in a comparable format. The “Key Information Set” (see text box below) is based on research which shows that these are the items that students say are the most useful. The data can be compared on the central Unistats website <http://unistats.direct.gov.uk/>

#### **Key Information Sets**

The Key Information Sets will help applicants to find quickly, and compare easily, the headline items which students consider most important. The content includes:

#### **Course information**

- Student satisfaction
- Proportion of time spent in different learning and teaching activities
- Different assessment methods used
- Professional bodies that recognise the course

#### **Costs**

- Accommodation costs
- Tuition charges
- Bursaries, scholarships and other financial support

#### **Employment**

- Destinations of students six months after completing their course
- Proportion of students employed in a full-time ‘graduate’ job six months after completing course
- Salary for course six months after graduating
- Salary for that subject across all institutions six months after graduating
- Salary for that subject across all institutions forty months after graduating

#### **The students’ union**

- Impact students’ union has had on time as a student

The Key Information Set draws some information (around half the data items) from the National Student Survey (NSS). This survey provides an opportunity for final year students to give opinions on their time at an institution/course as well as things that could have been improved.

Improvements were made to the Unistats website in September 2013, to make it easier to search (e.g. by location) and easier to access via mobile devices. HEFCE will begin a major review of the Key Information Set and the Unistats website in 2014.

The four UK higher education funding bodies are currently working with representatives from the sector to address the information needs of postgraduate students. A new online decision making tool called 'PGT choices' will be launched in 2015 designed to help prospective postgraduate taught students find the right higher education courses.

The full dataset of the Key Information Sets has been made available, free of charge, to external organisations such as bestcourse4me, Push, Student Room, Monster, Which?, and other social enterprise and commercial organisations that provide information to HE students.

The benefits stemming from this include:

- Free provision of KIS data helps 3rd parties to provide better analysis.
- New information providers coming into the market, encouraged by wider availability of data, will provide increased choice and wider access to data and information.
- A standard set of comparable information items for students, for each course, directly via HE Institution websites, driving consumer choice and raising standards.

The 'Students at the Heart of the System' White Paper highlighted that the Department's reforms of the English higher education system to make it more responsive to students and employers are dependent on high quality information about different courses and institutions.

Since the White Paper, we have set up a Ministerial Data Group to explore how we can make data more available, and explore the potential of linking, whilst ensuring that sound statistical principles are maintained and that individuals' personal data is not identified or used inappropriately. The group includes representatives from:

- BIS and DfE.
- Major holders of student data – HESA, UCAS, HEFCE and the SLC.
- Current and emerging student information providers - bestcourse4me, Push, Student Room, Monster and Which?
- Several leading academics in this field.

BIS held a higher education statistics user engagement event on the 30 January 2014 to gather feedback on the range of statistical publications produced by BIS, Department for Education, Higher Education Statistics Authority and the Student Loan Company. A summary of the feedback collected and actions being taken forward is available at the following link: <https://www.gov.uk/government/publications/higher-education-statistics-user-engagement-event-findings>

## Challenges

Whilst we are embedding the principle of open data in our policies, we have come against a number of challenges that have affected our ability to link or release data sets that we would have liked to, including: low sample sizes meaning that it can be difficult to provide information at very specific (e.g., course) levels for all students; and legal restrictions to publication of data – especially for Data Protection reasons – which are amplified when linking datasets.

We are working to minimise the demands made upon HE and FE providers – including only collecting data which is essential. To achieve an appropriate balance between the burden of data collection and providing increased transparency, we are using our stakeholders to guide us as to where to focus our effort.

The Higher Education Data and Information Improvement Programme (HEDIIP) has been established to enhance the arrangements for the collection, sharing and dissemination of data and information about the HE system.

The establishment of HEDIIP has been overseen by the Regulatory Partnership Group and HEDIIP is now carrying forward a programme of changes to build a more coherent, responsive and less burdensome information landscape. HEDIIP is hosted by HESA and the HEDIIP Programme Board is chaired by Sir Tim Wilson.

Further details about the work of HEDIIP can be found at <http://www.hediip.ac.uk/>

## Annex 2:

### Case Studies:

#### midata

The Government set out its plans for midata in the Consumer Empowerment Strategy [Better Choices: Better Deals](#) published in April 2011.

The programme is a voluntary programme that encourages suppliers in the energy, personal current account, credit card and telecoms sectors to release consumption and transaction data back to their customers in a reusable format, with the aim of empowering consumers with their own data.

In 2012 a power was taken through the Enterprise and Regulatory Reform Act that would allow the Government to bring forward regulations to mandate the release of data back to consumers with a particular focus on the sectors mentioned above. The Government has recently reviewed progress under the midata voluntary programme and has concluded that at this stage there is not a strong, objective case to bring forward regulations at this time. This is because of the encouraging progress that has been made, particularly in the energy sector where all of the large energy companies now provide access to a data file in a consistent format to their customers. Building on this progress in June 2014 the Secretary of State for Energy and Climate Change and Jenny Willott, Consumer Affairs Minister at BIS, held a roundtable with suppliers and agreed with them and third party intermediaries that they will work together towards providing third parties with automated access to consumer's data in a safe and secure way. Further progress on this initiative will be available by the end of the year.

In addition a recent commitment from the big banks, where they agreed to provide data back to their customers in a consistent way by spring 2015, should ease switching in the personal current accounts market.

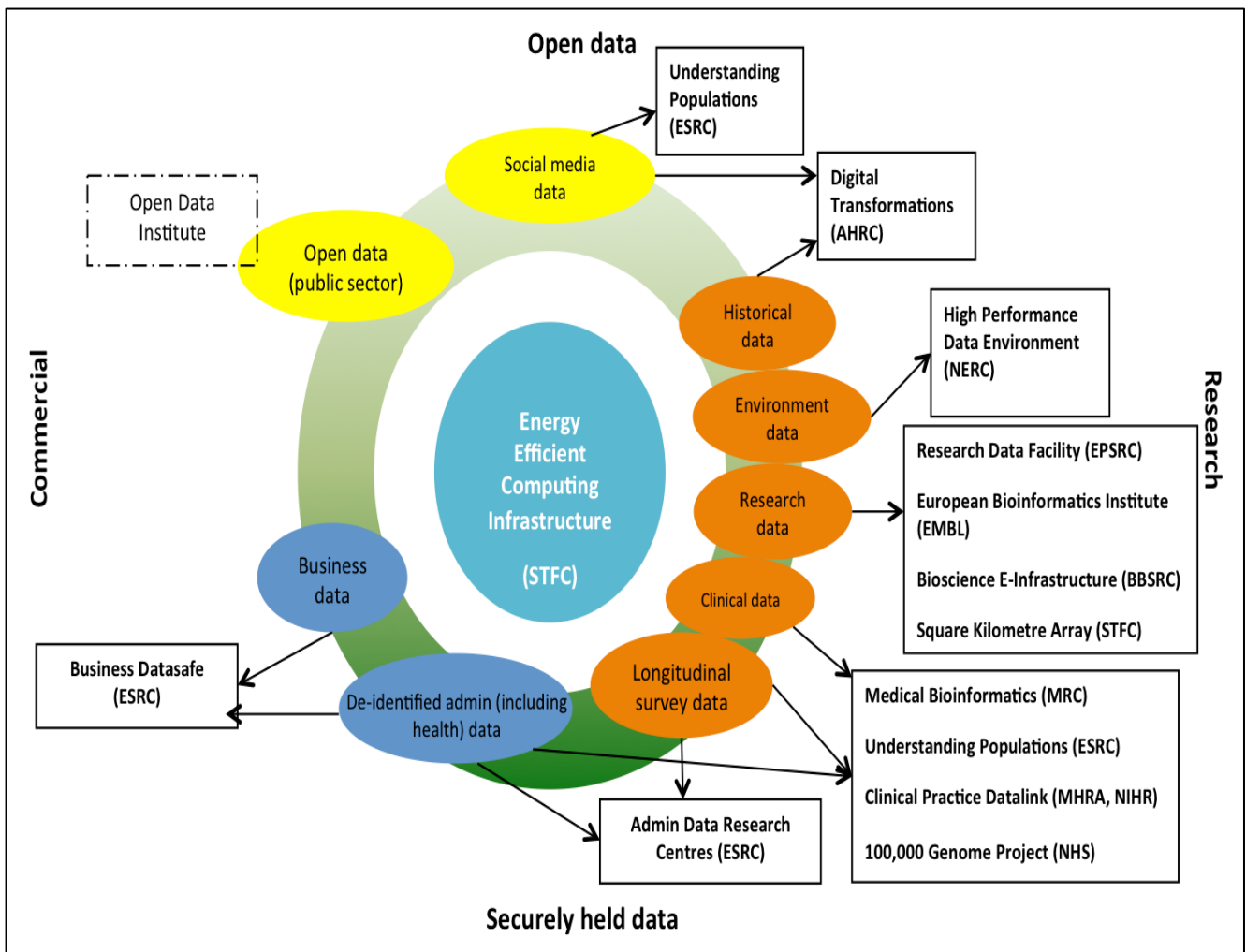
Another important consideration of the midata programme is considering the issues around protecting consumers' data and generating trust in third party intermediaries as this will be key to empowering consumers and unlocking more innovation. To help address this the Connected Digital Economy Catapult, funded by the Technology Strategy Board, will work with BIS, industry and key stakeholders to develop a code of practice for intermediaries for ethical use of personal data that is centred on consumer protection and benefit.



**Annex 2:**  
**Case Studies:**  
**Research Councils**

The Research Councils have a published data policy (<http://www.rcuk.ac.uk/research/datapolicy/>) related to the data that results from the research they fund.

**Diagram Showing Research Council Recent Investments into UK Research and Public Data Infrastructure**



## **Big Data and Energy Efficient Computing Capital Investment**

A suite of investments is being made by the seven Research Councils in priority research areas as identified by RCUK in the Investing for growth: Capital Infrastructure for the 21st Century framework. The proposal also has resonance with a range of Technology Strategy Board (TSB) priorities in enabling technologies such as information and communication technologies, but also those within the challenge-led areas such as connected digital economy, future cities, transport and health. Close working with the TSB will enable new knowledge generated to translate into maximum business benefits.

There is a package of £189 million of investment that will enable substantial new underpinning infrastructure. Projects are managed by their respective Research Councils with some overall co-ordination being provided by ESRC:

### **Square Kilometre Platforms (STFC lead)**

#### **Energy Efficient Computing for Big Data (STFC lead)**

The Hartree Centre, the world's largest centre dedicated to software development, is home to the most powerful supercomputer in the UK. The investment will support the progress of power efficient computing technologies designed for a range of industrial and scientific applications, and particularly in the development of software that can handle the huge amounts of data created by large experimental research initiatives, such as the Square Kilometre Array (SKA) and CERN, the largest generators of scientific data in existence. Energy efficient supercomputer software development to meet big data challenges. Economy-boosting partnerships between research and industry are just some of the benefits poised to come from the investment.

### **Research Data Facility (EPSRC lead)**

#### **The Alan Turing Institute (EPSRC lead)**

The UK Research Data Facility (UK-RDF), funded by EPSRC and NERC, is co-located with ARCHER and is housed at the University of Edinburgh's Advanced Compute Facility. It manages and safeguards the massive amounts of data being generated every day by publicly funded research. The Alan Turing Institute will catalyse a major collaboration between research and business benefiting from a Government investment of £42 million. It will be a national institute of global significance bringing together advanced mathematics and computing for human benefit across a wide range of sectors - from health to finance.

### **Environmental Virtual Observatory (NERC lead)**

Funds will be used to develop the cyber-infrastructure that is needed to give open access to big data, to provide computers capable of running complex environmental models and to capture real time data from sensors. It will also support the synthesis of this information so as to inform a variety of stakeholders, including researchers, policy-makers and business.

### **e-Infrastructure for Biosciences (BBSRC lead)**

An open access infrastructure to drive knowledge generation from genotype (simple big data) to phenotype (massively complex big data).

**Farr Institute for Health Informatics Research (MRC lead)****Four Health Informatics Research Centres (HIRCs)**

Funded at a total budget of £17.5m over 5 years as Centres of Excellence in research linking electronic health data. Based at UCL London, Manchester, Swansea and Dundee they involve a total of 19 UK universities and 2 MRC Units. The objectives of the centres are to undertake high quality, cutting-edge research linking electronic health records and health-related data with other forms of research and routinely collected data in safe data environments; and build capacity in health informatics research.

**UK-wide Administrative Data Research Centres (ESRC lead)****Business Datasafe (ESRC lead)****Understanding Populations (ESRC lead)**

In Phase 1 of the Big Data Network the ESRC has invested in the development of the Administrative Data Research Network (ADRN), which will provide access to de-identified administrative data collected by government departments for research use. Phase 2 will focus primarily on business data and local government data, and Phase 3, will focus on Third Sector data and social media data.

**Digital Transformations in the Arts and Humanities (AHRC lead)**

The call for proposals for projects to explore Big Data from an Arts and Humanities perspective closed in September 2013.

## Annex 3: List of Datasets Provided on a Statutory Basis

Statutory requirements:

- UK Innovation Survey - EC Regulation No. 995/2012 <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2012:299:0018:0030:EN:PDF> (previously EC No. 1450/2004)
- Science, Engineering and Technology Indicators (responsibility for these statistics has now been passed to ONS) - some elements covered by EC regulation 995/2012 Statistics on Science and Technology  
  
<http://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2012:299:0018:0030:EN:PDF>
- Strategic Export Controls - Export Control Act 2002 requires publication of a report relating to operation of the Act, which could be interpreted as including some statistics. <http://www.legislation.gov.uk/ukpga/2002/28/contents>

For Partner Organisations:

- Scientific Advice on Matters Related to the Management of Seal Populations (Natural Environment Research Council) - Conservation of Seals Act 1970 <http://www.legislation.gov.uk/ukpga/1970/30/contents>
- Certification Officer Annual Report (Certification Officer) - Trade Union and Labour Relations (Consolidation) Act 1992 <http://www.legislation.gov.uk/ukpga/1992/52/contents>

Those are the releases which are explicitly required by legislation. There are others for which there is a legislative requirement for data to be provided but for which there is no specific requirement to publish a particular set of statistics. For example:

- The Further and Higher Education Act 1992 includes a duty for FE and HE institutions to supply information to the Funding Councils
- Statistics of Trade Act 1947 provides a general power for all departments to collect statistical information from business on production, distribution and services. Largely used by ONS these days it originally empowered the Board of Trade and BIS is still a "competent authority". BIS contracts ONS to carry out surveys of building materials and some of these are statutory surveys. This allows legal action ("enforcement") to be taken if companies do not complete questionnaires.

General principles around the production and publication of official statistics are set out in:

- Statistics and Registration Service Act 2007 <http://www.legislation.gov.uk/ukpga/2007/18/contents>
- Pre-release Access to Official Statistics Order 2008 <http://www.legislation.gov.uk/uksi/2008/2998/contents/made>
- Official Statistics Order 2013 <http://www.legislation.gov.uk/uksi/2013/1163/contents/made>

- Code of Practice for Official Statistics (section 13 of the Statistics and Registration Service Act includes a requirement for all National Statistics to comply with the Code)

Building Price and Cost Indices is not required by statute, but there are statutory instruments which require the use of the indices. These are:

- The Community Infrastructure Levy Regulations 2010 (2010 No 948) and annual updates <http://www.legislation.gov.uk/all?title=community%20infrastructure>
- The Housing (Right to Buy) (Service Charges) Order 1986 (1986 No 2195) <http://www.legislation.gov.uk/uksi/1986/2195/contents/made>
- The Housing (Right to Buy) (Service Charges) (Amendment) (England) Order 2010 (2010 No 2769) <http://www.legislation.gov.uk/uksi/2010/2769/contents/made>

© Crown copyright 2014

You may re-use this information (not including logos) free of charge in any format or medium, under the terms of the Open Government Licence. Visit [www.nationalarchives.gov.uk/doc/open-government-licence](http://www.nationalarchives.gov.uk/doc/open-government-licence), write to the Information Policy Team, The National Archives, Kew, London TW9 4DU, or email: [psi@nationalarchives.gsi.gov.uk](mailto:psi@nationalarchives.gsi.gov.uk).

This publication available from [www.gov.uk/bis](http://www.gov.uk/bis)

Any enquiries regarding this publication should be sent to:

Department for Business, Innovation and Skills  
1 Victoria Street  
London SW1H 0ET  
Tel: 020 7215 5000

If you require this publication in an alternative format, email [enquiries@bis.gsi.gov.uk](mailto:enquiries@bis.gsi.gov.uk), or call 020 7215 5000.