



Department
for International
Development

Open Data Strategy

April 2012 – March 2014

Contents

Summary	2
Introduction	3
Part 1: Vision and Stocktake	4
Vision.....	4
Where we are now.....	6
Part 2: Strategy	8
1. Capture and release of Big Data	8
2. Capture and release of My Data	11
3. Satisfaction and Experience Data	12
4. Creation of Dynamic Information Markets.....	14
5. Continuous improvement of data quality.....	16
Annexes	23
Annex 1: DFID Published Datasets	23
Annex 2: List of data to be released	25
Annex 3: Information Principles	26

Summary

Increasing the transparency and traceability of aid has the potential to transform the way aid is delivered and to improve its impact, enabling both taxpayers and beneficiaries to hold governments and other providers to account for delivering effective development. It has the potential to support the drive to improve national transparency and accountability in the countries where we work – further empowering their citizens to hold their own governments to account for public expenditure.

This Open Data Strategy sets out how the Department for International Development (DFID) will drive reform and improvement through greater transparency and citizen participation as part of the way we work. The Strategy develops and refreshes the Information Strategy previously published as part of the DFID Business Plan and is supported by a schedule of new data release commitments for the period from April 2012 to March 2014.

Big Data

To support traceability of aid we will redevelop our aid information platform to incorporate data from DFID and partners. Through a new aid transparency initiative we will drive increased transparency by our partner organisations (other government departments, NGOs, multilaterals, private sector) which will help citizens trace the use and impact of aid funding. We will publish further information on DFID's aid projects, including first results against DFID's results framework and detailed location information. We will provide greater access to data from DFID-funded research. We will explore ways of linking DFID aid data with partner county budget data and to promote national budget transparency.

My Data

DFID does not provide any significant public services directly to UK citizens, and therefore does not hold significant volumes of data on such individuals.

Satisfaction and Experience Data

We will run pilot projects in a number of developing countries to test ways of seeking feedback from the people directly affected by development projects, using mobile technology where appropriate. We will use the opportunities provided by social media to further engage with members of the UK public on DFID activities.

Dynamic Information Markets

Building on our success in bringing the providers of 75% of global ODA within the International Aid Transparency Initiative in the last 12 months, we will actively encourage adoption and use of IATI standards through an Aid Information Fund.

Data Quality

Building on lessons learned and good practice across Whitehall we will develop a Transparency Data Quality Improvement Plan to improve the quality of data we publish.

Introduction

The Department for International Development (DFID) Open Data Strategy demonstrates how we will take forward improvements to the data we already publish; contribute to the drive for reform and improvement through transparency and citizen participation; and work with those in receipt of UK aid to drive transparency and accountability right through the delivery chain.

Increasing the extent and depth of aid transparency will open up the whole aid delivery chain to scrutiny by taxpayers and beneficiaries alike. This has the potential to transform the way aid is delivered and to drive greater impact and effective use of aid resources. It also has the potential to help drive greater transparency and accountability for development resources overall. This is articulated further in Part 1.

Part 2 of the document outlines the strategy for the next two years. The main focus of this strategy is on developing the extent, quality and usability of the Big Data that we make available under the DFID transparency, results and research agendas, as well as publishing more of our management information to reflect activities across DFID. This is covered in Chapter 1.

Direct transactions involving citizens are limited compared to those of other UK Departments; but these are reviewed in Chapter 2 looking at My Data.

How we aim to gather and assess feedback from customers for our data is included in Chapter 3, and ways of facilitating its greater use are covered in Chapter 4. Our plans to improve the quality of the data that we publish are set out in Chapter 5.

Part 1: Vision and Stocktake

Vision

DFID has a bold and ambitious vision on transparency, both for itself and for its partners. DFID recognises that increasing transparency and traceability of aid has the potential to transform the way aid is delivered and to improve its impact, and to help increase transparency and accountability between governments and citizens in the countries where we work.

More open information on aid inputs and results will provide evidence of what works and what doesn't work. Showing the geographical areas where aid projects are concentrated will allow governments, agencies and people to direct aid better. Opening up the aid delivery chain will enable taxpayers and beneficiaries to check that the aid promised has been delivered.

Principles

DFID's vision of transparency is underpinned by the following principles:

- 1) We believe that **aid spending should be fully transparent**. This means that anyone anywhere should be able to track every pound of aid spending right through the system, from DFID all the way to the activity on the ground.
- 2) We believe that people living in the countries in which we operate should be able to see exactly what our aid spending is intended to achieve so they can **hold us and others to account for delivery**.
- 3) We believe that transparency can significantly **reduce the administrative burden** on recipient governments and agencies by ensuring that information is only released once – widely and to the entire world – in place of the current bureaucratic reporting requirements often imposed by multiple and separate aid bodies.

Where do we want to be?

We want to be able to **trace the flow of aid** from donor to implementation – so that we are able to see the impact and results of aid spending. This will provide clear visibility to taxpayers in the UK and to citizens of countries where DFID is providing development assistance.

We want to make comprehensive information **easily available**, including on our website, and to incorporate detailed data from our development and research partners to provide a full picture of aid delivery.

We want to provide our data as **downloadable open data** complying with international standards, so that others can freely reuse the information.

We want to build up the **public good infrastructure** and **encourage innovation** in the use of open data for aid by freely sharing the tools and approaches that we use, and by learning lessons from others.

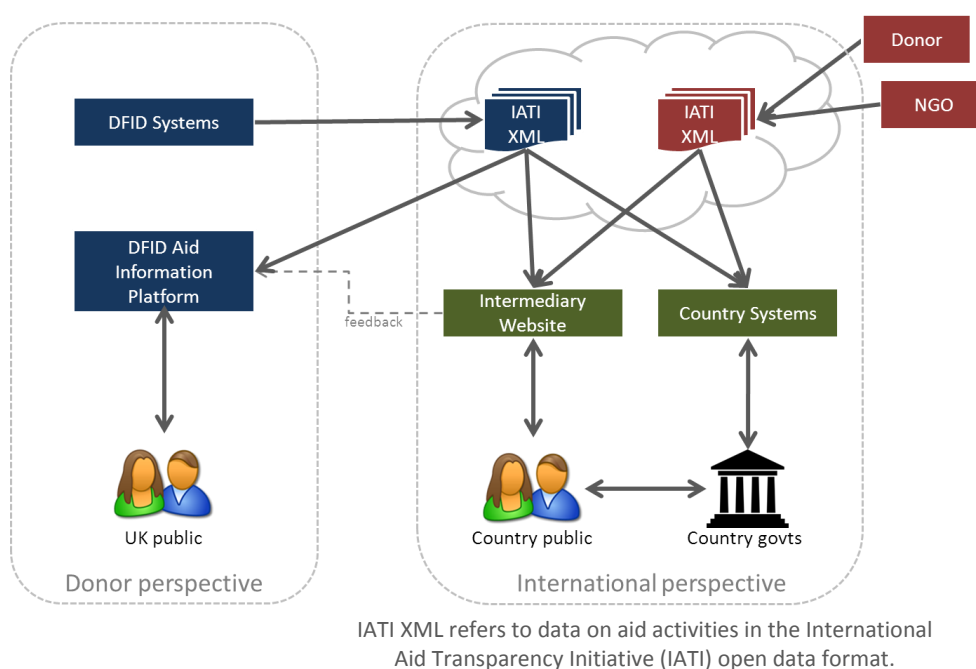
We want to **encourage others to be transparent** about their activities, so that there is a complete picture of global aid and resources for development.

We want to **encourage partner governments to be transparent** so that they can be held to account by citizens of those countries.

We want to be a **reflective and learning organisation**, able to use open data from a variety of sources to continually improve our capability and effectiveness.

We want to **listen to others**, whether experts in open data, aid delivery, partner governments or citizens in the UK and across the world, to continually improve our aid impact.

DFID Open Data Model



DFID's Open Data Model describes two key perspectives:

The **Donor Perspective** means that we publish data on DFID activities so that the UK public can understand how taxpayer's money has been used.

The **International Perspective** means that as we publish data in an open, internationally agreed standard, others can reuse and repurpose this data from different perspectives. Examples include a civil society organisation in a developing country looking at aid flows to that country, a developing country government managing aid budgets, or an international NGO collating information on activities in the health sector.

Where we are now

The **Prime Minister's letter**¹ of May 2010 set out the core transparency requirements for all government departments, including transparency on spend, contracts and staffing. The letter also committed the government to publish full information on all DFID international development projects over £500 including financial information and project documentation.

DFID launched the **UK Aid Transparency Guarantee**² in June 2010. This set out our ambition to publish detailed information about all new DFID projects in a common standard, and to push for full transparency across the international aid system. A review was published in April 2012³.

DFID published an Information Strategy as part of the Transparency section of its **Business Plan for 2011-2015**⁴. This set out the wider commitment of DFID to becoming a global leader on aid transparency, on the basis that this is a fundamental requirement if we are to be more accountable to UK citizens and to citizens in the countries where we work. Greater transparency and citizen participation are key elements of how we aim to demonstrate good value for money and greater effectiveness in reducing poverty through our aid programme and projects.

In this previous Information Strategy we undertook to publish the following categories of new information online:

Type of information	Progress
Full information on all new DFID projects over £500 (from Jan 2011) including those funded through the tri-departmental Conflict Pool and managed by DFID	Now published and updated monthly.
Expand the project data included in our new project database to be in line with international best practice, providing more detail and more frequent updates	Project information is now published in line with the International Aid Transparency Initiative (IATI) and is refreshed monthly, although data quality can be improved.
Make our data and information more accessible by changing the format of the information we publish, including through revised project documentation and country business plan formats	Project documents (Business Case, Annual and Project Completion Reviews, Logframes) are published from 2011.
Online details for all new tender documents for contracts over £10,000 (from Sep 2010)	Tenders released centrally from the UK are published on the ContractsFinder website.
Names, job titles and annual pay rates for Directors and above (from Oct 2010)	Now published and updated in line with Cabinet Office timetable.
Organisational structures covering all of the Senior Civil Service (from Oct 2010)	Now published and updated in line with Cabinet Office timetable.

Annex 1 list details of the full range of datasets we make available.

¹ <http://www.number10.gov.uk/news/letter-to-government-departments-on-opening-up-data/>

² <http://www.dfid.gov.uk/ukaid-guarantee>

³ <http://govuk.info/90q5>

⁴ <http://www.dfid.gov.uk/About-us/How-we-measure-progress/DFID-Business-plan-2011-2015/Transparency/>

All aid projects with spend of over £500 are published online in both a user-friendly website ([DFID Projects Database](#)) and in the International Aid Transparency Initiative (IATI)⁵ open data format. The IATI format data is supported by a simple Application Programming Interface (API) allowing others to readily reuse the data. There are currently approximately 3250 projects available online, each with a unique identification code that allows linking to specific projects.

The published data contains descriptions of each project, the countries and sectors impacted, budgets, spend (including detailed transactions), and is supported by comprehensive document sets including new Business Cases, Annual Reviews and Project Completion Reviews from 2011 onwards.

All information is published on the DFID website and on the [data.gov.uk](#) website in open reusable formats under the Open Government Licence. DFID data in IATI format is available from the IATI Registry at <http://iatiregistry.org>.

The **International Aid Transparency Initiative (IATI)** is a major international open data initiative involving donor countries, developing country governments, non-governmental organisations and experts in aid information.

In February 2011 the participants agreed a common open international standard for publishing information on aid. There are now 29 donor and 22 partner country signatories, covering 75% of global aid. Twenty-four organisations are currently publishing data to the standard.

DFID had a significant role in leading IATI and was the first organisation to publish information to this international standard.

Following feedback from the public, we redeveloped our Projects Database to make the information clearer and easier to understand. The redesign was strongly influenced by a structured usability study carried out with members of the public and aid specialists. Since the relaunch in December 2011 the number of visitors to the Projects Database has grown by 32%⁶. The database is limited to DFID information and does not support traceability.

Following the Bilateral Aid Review in 2011 DFID introduced a **strengthened approach to budget support**⁷ that placed more emphasis on domestic accountability of partner countries. We made the commitment to spend an amount equivalent to 5% of budget support to enable people in budget support countries to hold their governments to account for the use of public resources, including aid. This involves governments making budget information fully available and accessible, providing opportunities for citizens to engage with government, and ensuring that officials can be called upon to justify their actions. Publishing documents related to budget support is one way we are encouraging governments to be more transparent.

The R4D resource on the DFID website⁸ holds information from DFID funded research and links to supporting data sets. Currently this comprises 5200 project records and 30,000 outputs. R4D increasingly links to other research sources and knowledge. The current platform shares R4D data in formats that make it easier to reuse elsewhere.

⁵ The International Aid Transparency Initiative standard can be found at <http://iatistandard.org>.

⁶ Unique visits from a 3 week period in Oct 2011 compared with a 3 week period in Jan 2012.

⁷ <http://www.dfid.gov.uk/What-we-do/How-UK-aid-is-spent/how-we-decide-where-aid-is-spent/>

⁸ www.dfid.gov.uk/r4d

Part 2: Strategy

1. Capture and release of Big Data

DFID currently publish a wide range of information on DFID activities, covering both aid delivery and other business functions, and this is summarised in Annex 1.

Data on UK aid projects

In order to support traceability of aid from donor to implementation, we will redevelop our **aid information platform** (<http://projects.dfid.gov.uk>) to incorporate data from DFID and our implementing partners. The platform will allow full drill-down from country or sector to individual projects and expenditure, and will also support more detailed location and results information. The platform will be based on IATI format data using open source tools and we will release the source code to enable others (for example NGOs, private sector partners, developing country CSOs) to present their aid data. We will relaunch the platform in late 2012.

We will launch a further **aid transparency initiative** under the UK Aid Transparency Guarantee aimed at driving much greater transparency of aid data by our **implementing partners**, including other government departments, NGOs, multilateral organisations, private sector partners. We will require those directly managing our funds to release open data on how this money has been spent, and require sub-contractors or sub-agencies to be similarly transparent, so that money can be followed right through the supply chain. This will also allow us to investigate ways of using open data to allow people to better **track use of UK aid**, thereby reshaping our understanding of global aid flows.

As committed under the UK Government's National Action Plan for the Open Government Partnership, **other UK government departments and devolved administrations** will put in place plans to publish their aid funding data in IATI format. We will work with them to achieve this.

In order to best meet the demands of people using our data, we will consult with data users and, where possible, adapt and provide datasets. We will also establish an external **Sector Transparency Board** to inform and advise on future direction of aid transparency within the UK, and to provide a real challenge function. The Board will be launched in September 2012.

We will further develop the published data on DFID aid projects to include **Results** so that the public can understand what is achieved from specific aid projects. The first results delivered against the 25 centrally monitored indicators in DFID's Results Framework will be published as reusable data by June 2012. Other results are published within project documents and we will include project-level results in IATI data. Project performance scores using DFID's new scoring model will be released as reusable electronic data by February 2013.

We will provide more detailed location information (also known as **geocoding**) for individual projects using recognised international methodology, with a view to applying this methodology to all new aid projects by March 2014, based on a pilot to be carried out in 2012.

We will publish all **project evaluations** from May 2012 onwards.

Transparency, accountability and partner countries

We will establish a methodology for linking DFID aid budget data with **partner country budget data** by July 2013, following trials during 2012.

DFID works with governments and other donors to make the process around **budget support** more transparent and inclusive. We will extend the impact of our commitment to spend up to 5% of budget support funds on strengthening local transparency, by reviewing work to date and updating guidance to staff by October 2012, and we will continue to promote:

- The publication of Performance Assessment Frameworks, annual review reports, results, and conditionality in government/ donor budget support agreements;
- The provision of information on the national budget and budget support in ways the public can access and understand;
- The involvement of parliaments and civil society in the identification and monitoring of budget support conditions and benchmarks.

DFID funded research

We will further enable access to our **DFID funded research and supporting data sets** through our new Open and Enhanced Access Policy to be published on the DFID website by the end of June 2012. Under this policy, DFID undertakes to address the Right to Data principle in another area of its operations. Users worldwide will be provided with unrestricted, irrevocable and free online access to full-text/full version scientific and scholarly material. R4D will hold simple data sets as well as unstructured data and provide links to data sets held on other repositories across the world.

The policy will apply to contracts awarded after 30 June 2012, and those holding grants awarded prior to that date will be encouraged to follow the same process. The target % for open access to data sets has not yet been agreed but a target of 80% has been set for availability of peer reviewed articles from DFID funded research, on open access, by end 2013.

Where applicable, we will link research and aid projects using the IATI Standard as a common reference. We will explore the opportunities provided by an extension of the IATI Standard to include research data.

Corporate information

We will work with Cabinet Office in the publication of corporate information on staffing, spend, contracts and other performance indicators⁹, so that this information can be compared across Government.

We will expand our publication of tenders and contracts on the Contracts Finder website to include tender opportunities released and contracts awarded by our offices overseas. We will look for enhancements to the IATI Standard to allow for clearer publication of contracts.

Capability to use open data

While DFID will publish a large amount of data about our own activities, we will also be a **user of open data produced by others**. In order to be effective at using data from others, we will increase our capability and will promote the use of DFID open data by our own staff.

We will continually **review our reporting requirements**, in order to strip out any administrative requirements that could be fulfilled by an open data approach. We want to ensure that our transparency agenda fulfils its objective of reducing administrative burdens, rather than adding to them.

In countries where we work, we will identify opportunities **to support citizens, legislatures and media to utilise data** that we publish to scrutinise aid delivery and to inform citizens.

Benefits and Information Principles

The publication of more information on aid projects and corporate activities has benefits in:

- Increasing accountability of DFID to the public.
- Encouraging better decision making, leading to better use of public money.
- Linking aid spend to outcomes.

Having ready access to aid data is also of great value to developing country governments as it helps inform the way they manage their budgets. It is also of benefit to development organisations such as local civil society organisations and international NGOs who are looking to make sure they direct their funds effectively. Making datasets and research papers available enables secondary analysis of data and contextualising of results to different circumstances. It can also avoid duplication of work.

Information Principles 1 – 6 are addressed in these commitments as outlined in Annex 3, and demonstrate the high priority that DFID gives to better management of its data as a valuable corporate asset, to the development of the necessary skills and corporate processes, and to ensuring that data is suitable for publication, easy access and re-use.

⁹ This includes the Quarterly Data Summary and the Equality Information Report.

2. Capture and release of My Data

Unlike most other government departments, DFID does not provide any significant public services directly to UK citizens, and therefore does not hold significant volumes of data on such individuals.

Access to information by citizens (for example, pensioners, current, former and potential employees) is managed through existing channels, including our Human Resources and Overseas Pensions departments and the Openness Unit.

We do not propose any further initiatives to deliver personal data online.

Benefits and Information Principles

The services under which we provide this data to pensioners and to other individuals provide benefits as follows:

- Increasing accountability of DFID to the public
- Supporting public service productivity

Annex 3 details how Information Principle 7 is addressed through the services of the Openness Unit as the first port of call for enquiries to DFID by email, letter or telephone from citizens wishing to access their data. We have established from feedback that online access to transactional information is not likely to benefit the community of overseas pensioners that we serve.

3. Satisfaction and Experience Data

Satisfaction and experience data are important to ensure that we learn lessons from what worked well and crucially from what did not work well. Satisfaction and experience data applies to the following groups of people.

People who benefit directly from UK aid

We will run pilot projects in a number of developing countries and through the Global Poverty Action Fund to test ways of seeking **feedback from the people directly affected by aid projects**. These pilots will test various mechanisms to receive feedback, including the use of web and mobile technology.

Results of the pilots will be used to influence future project design in country programmes to ensure that the voices of those who receive UK aid are considered. We will also establish recommendations for how feedback could be incorporated into the IATI Standard.

UK citizens and taxpayers

We will explore how to engage members of the UK public in order to improve aid delivery, using new digital tools including social media.

Consultations on the usability of the Projects Database showed that both public and specialist users wanted more descriptive information to explain the nature, aims, progress and outcomes of projects and programmes. We will use this feedback in developing new ways of communicating aid project information.

Surveys asking for customer feedback from public enquiries have demonstrated a high degree of satisfaction from users of the public enquiry service. We will ensure that information on issues of frequent public enquiries continues to be available on our website.

Data on our performance in dealing with Freedom of Information requests is published centrally on the Ministry of Justice Public website. We will publish some further information on common requests made to DFID under the Freedom of Information Act on our website.

Users of DFID funded research data

User satisfaction with the R4D service was surveyed early in 2011 and through usability testing with key users during the integration of R4D with the DFID website. The survey showed that over 60% of respondents were either very satisfied or satisfied with their visit to the website.

User experience will continue to be assessed under new arrangements for the service to be agreed in 2012.

Overseas pensioners

Customer satisfaction data is an inherent part of overseas pensions administration and data is gathered in a systematic way from the pensioner community, and is published online as part of the Annual Report.

We will consider publishing customer satisfaction data online in a data format, as at present this information is contained within a Word document.

Benefits and Information Principles

Collecting and responding to feedback on satisfaction and experience relating to our activities has benefits in the following areas:

- Improving the quality of the data service that we offer e.g. improved usability
- Informing our decision making leading to better use of public money
- Increasing accountability of DFID to the public, both UK citizens and those in developing countries

Citizens in developing countries want to know that aid contributes to service delivery, economic development and local infrastructure. We therefore need to make sure that the information we publish on the work that we do is accessible and understandable to the people for whom aid is intended. We also need to go further and promote feedback. Our plans include a pilot to test innovative ways of doing this e.g. through new technologies, local think tanks and grass roots organisations.

The Information Principles are addressed in these commitments through inviting feedback and taking steps to respond e.g. by publishing more usable data or by adapting our programmes and activities based on the views of aid project recipients.

4. Creation of Dynamic Information Markets

Aid data

Within the last 12 months, the providers of 75% of global ODA have committed to publishing data to the International Aid Transparency Initiative (IATI) Standard. Recent signatories include the USA, Canada and UNICEF. The IATI Standard is now the foremost technical standard for reporting aid project information, and will continue to be supported and extended by the international community.

DFID have been working closely with partner organisations to encourage the development of tools¹⁰ that make use of aid transparency information published to the IATI Standard.

We will continue to encourage adoption and use of IATI standard, and will support the **establishment of a fund** to support developers directly to produce innovative and useful tools to make use of IATI format data in key areas such as:

- Traceability of aid through the aid delivery chain
- Partner country use of IATI data
- Converting IATI format data into Linked Data
- Visualisation of aid information

DFID country offices will be encouraged to participate in the challenge, for example by developing initiatives linked to their own particular development context, or by using local developers.

We will establish a governance board for this fund including external experts in aid and data management to assist us in identifying and prioritising opportunities. One of the governing principles is that these tools must show benefits to the end recipients of aid.

As part of our commitment to invest in **public good infrastructure**, we will require that tools developed through this fund are made open source and that there is support for others to use and further develop these tools. We will release the source code for the aid data platform referred to in Chapter 1.

In addition, we will continue to encourage the growth of a **developer ecosystem** by supporting developer awareness and “hack day” events, and by working with Civil Society, NGO and private sector (SME) partners.

¹⁰ Examples of the tools now available can be found at aidinfo.org.

Research data

Publication of more DFID funded research data under the new Open and Enhanced Access Policy will be of interest to a range of stakeholders who can take the data and add value, as shown in the side panel.

DFID funded research on R4D is available through shared, open formats to encourage reuse of both the published research documents and the data about those publications (metadata; for example, the standard descriptions of research). Information can be searched and downloaded from the website in Excel format. Themed subsets of metadata relating to research in particular countries or subject areas are also available through the Open Archives Initiative Protocol for Metadata Harvesting¹¹ – a global standard for sharing research information between archives.

Using linked open data, R4D have developed a prototype widget¹² that links R4D publications with DFID projects to provide information on scientific research linked to any aid project. The tool will be made available to external websites providing a flexible mechanism for sharing up to date data to help inform project research teams and other interested people.

Young Lives is a DFID funded project which is an international study of childhood poverty involving 12,000 children in Ethiopia, India, Peru and Vietnam over a 15-year period. It is a good example of a project which is making its data easily available through open archiving, data visualization, linked data and documentation.

Young Lives deposits all of their survey data, plus supporting documentation, in the UK Public Data Archive, hosted by the Economic and Social Data Service (ESDS). Some of the survey data held by ESDS is complex so they have been experimenting with some data visualization tools such as interactive graphs.

These initiatives have led to a steady increase in users accessing the data via ESDS. Young Lives have also run some in-country workshops to encourage researchers to use the data.

For further details, see <http://www.younglives.org.uk>

Benefits and Information Principles

Creation of these information markets offer benefits in relation to opening up and enhancing access to the outputs of the aid projects and the research that we fund. Lessons from these outputs help to ensure that future projects are more effective. Scientific and social science knowledge can have great benefits in providing practical ways of addressing poverty. The information that is made available also has an impact on wider society and the economy through the development of new tools and methodologies, new policies and regulatory controls, and new products or services.

The Information Principles are addressed by proactively offering re-use opportunities and trying to overcome constraints on re-use.

¹¹ <http://www.openarchives.org/pmh/>

¹² <http://r4d.herokuapp.com/>

5. Continuous improvement of data quality

We recognise that the feedback loop provided by open data publication has the potential to drive internal improvements in data quality and therefore improve standards.

Building on lessons learned and good practice across Whitehall we developed a Transparency Data Quality Improvement Plan in April 2012 that sets out steps to improve the quality of data we publish.

What is good data quality?

Good quality data has the following characteristics according to Audit Commission guidelines¹³ and as referenced in the Information Principles.

Data must be...		Relevance to DFID open data
Accurate	Accurate enough for the intended purposes	Data is drawn from internal systems where it is used to support decision making. Actions listed below will help improve the accuracy of data in our internal systems.
Valid	Recorded and used in compliance with the relevant requirements	Data is recorded to support internal processes and decision making, and is subject to review and approval.
Reliable	Reflects stable and consistent data collection processes across collection points and over time	The system controls enforce data collection of both project and financial information to ensure consistency.
Timely	Captured as quickly as possible after the event or activity and made available for the intended use within a reasonable period of time	DFID data is published at the end of each month, reflecting the position at the end of the previous month. The emerging international standard for aid transparency is to publish at least quarterly for the previous quarter.
Relevant	Relevant to the purposes for which it is to be used	Information is captured to support DFID's business processes and can be used when published for a variety of purposes.
Complete	Data requirements should be clearly specified based on the information needs of the body and data collection processes matched to these requirements.	DFID publish the information that we hold in our internal systems. A limited number of exclusions are made based on security considerations.

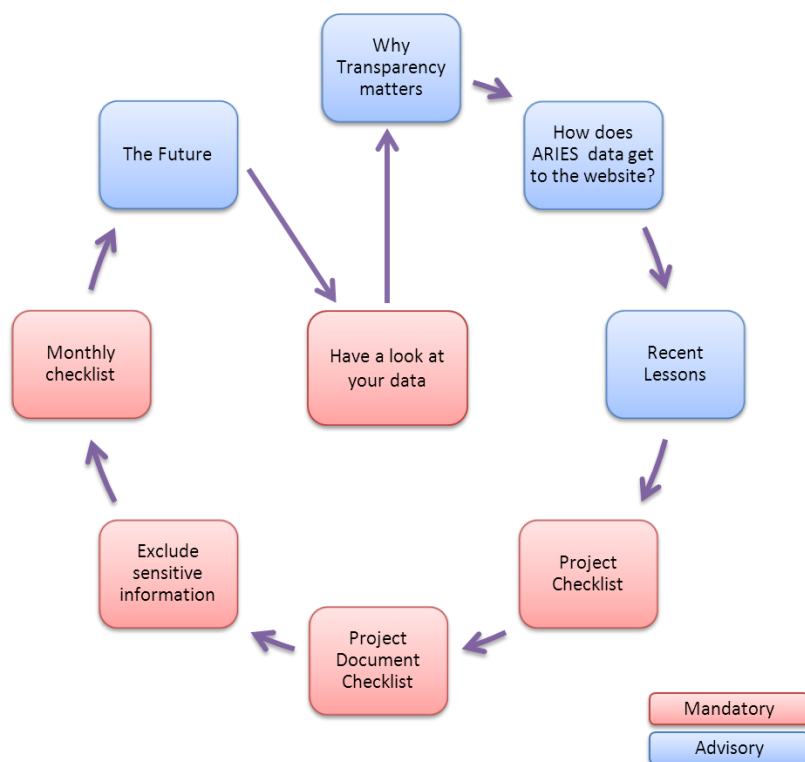
Steps towards improved data quality

Transparency is helping to drive improvements in data quality. Where information is only available within an organisation, decision makers make allowance for a lower level of

¹³ Nothing but the Truth: a discussion paper. Audit Commission, 2009. ISBN: 1-86240-583-2.
<http://www.audit-commission.gov.uk/nationalstudies/localgov/Pages/nothingbutthetruth.aspx>

quality. Sharing information openly enables other users of the information to highlight areas of improvement, for example where information could be explained more clearly or expressed more precisely.

To support staff in improving the quality of the project data that they have responsibility for, we have developed a **data quality continuous improvement cycle**, as shown below.



We have identified key steps to help improve data quality, and we will:

- Ensure that transparency and data quality are noted in all our guidance on systems and procedures.
- Increase the visibility of published and pre-published information to internal DFID users, encouraging them to regularly check the quality of data for their areas of responsibility.
- Circulate regular monthly "lessons learned" to all staff to provide direct feedback from open data users.
- Introduce easy to use check lists to assist in data quality improvement.
- Provide detailed league tables and statistics on divisional performance for senior management to drive awareness of data quality.
- Regularly review the data exclusions policy and existing data exclusions to ensure that as much data as possible is published.
- Encourage our research partners to collect and store data from the outset in a manner that enables data archiving, and to deposit datasets in existing, specialist repositories with high quality standards.

Continually improving the IATI Standard

The IATI Standard is a good foundation for delivering comparable data on aid activities, but there are a number of areas where further developing the standard could be explored:

- Enabling publication of budgets in accordance with **partner country budget classifications**.
- Establishing a clear framework for **exclusions**, including explaining where and why information has been excluded.
- Providing a **research** extension to IATI, to enable research activities to be captured in a comparable format.
- Enabling **contracts** to be published as part of IATI, to support traceability.
- Supporting the recording and publication of **feedback** from those affected by the delivery of aid.

We will work with the IATI Steering Committee to agree priorities for further development of the standard. We will also work with partners in developing and testing IATI compliance tools.

Moving up the 5 star rating system

Sir Tim Berners-Lee has introduced a 5-star rating system for open data¹⁴.

★	Data is accessible on the web, readable by humans but in a closed document format (e.g. PDF file).
★ ★	Data is accessible on the web in a structured machine-readable but proprietary format (e.g. Excel file).
★ ★ ★	Data is available in an open format (e.g. CSV file, XML data).
★ ★ ★ ★	Data elements are referenced using Uniformed Resource Identifiers (URIs), enabling linking (e.g. every Supplier has a unique organisation identifier; every Country has a unique country identifier, every Project has a unique identifier).
★ ★ ★ ★ ★	Data is linked to other data enabling discovery of related content (e.g. the organisation identifier is linked to further information about that organisation, supporting linkage of information not contained in the original dataset).

DFID data is generally published at 3-star level, either as CSV files (for example, monthly Spend data) or as XML (IATI format aid project information). Organograms are released as RDFa data.

We will investigate how to define a Linked Data approach for aid data, based on the IATI standard, and will conduct a study by October 2013, exploring implementation by March 2014. This will enable aid project data to be converted into 4- and 5-star data, and will support efforts to provide greater traceability of aid spend.

In addition, we will investigate publication of our Spend dataset as RDFa data by March 2013.

The experimental R4D tool for research data described in the example in Chapter 4 above is already moving towards delivering a 5-star open data rating.

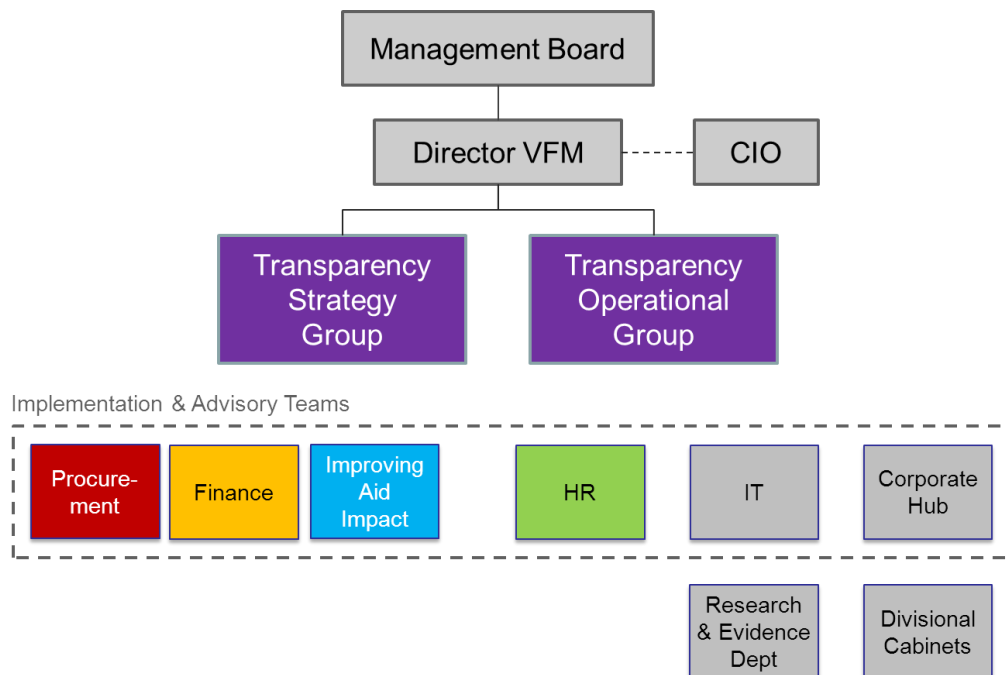
CSV: Comma Separated Values. A text file with information fields separated by commas.

XML: eXtensible Markup Language. A structured format for data transfer that supports hierarchical information.

RDFa: Resource Description Framework. A technique for adding further meaning to the information in a data file.

¹⁴ For further information on the 5 Star Rating System for Open Data, see <http://www.scribd.com/doc/68607846/What-are-the-5-Stars-of-Open-Data>

Transparency governance



DFID's transparency activities are managed by three groups working together:

The **Corporate Hub** is responsible for day-day operational running of the transparency publication process, identifying opportunities to improve data quality and to streamline the process. They work closely with Business Solutions Division to ensure that the IT infrastructure and tools support the production of open data.

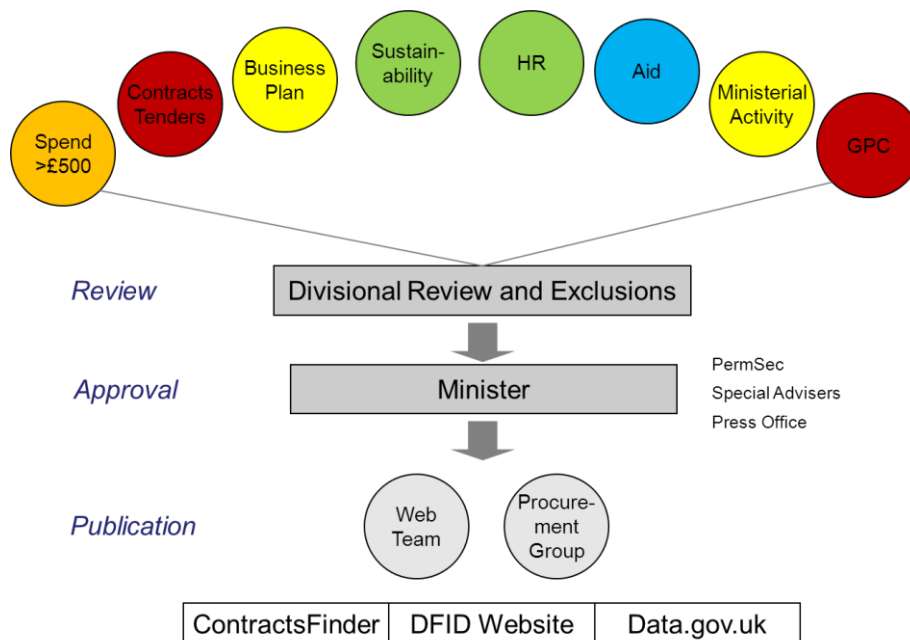
The **Transparency Operational Group** provides advice and guidance to the Hub on operational matters. This group contains experts from all transparency areas: Finance, Aid Management, Procurement, HR, IT.

The **Transparency Strategy Group** sets and monitors DFID transparency strategy, and contains senior leaders from Finance and Aid Effectiveness, Procurement, IT and Corporate Hub.

The **Director, Value for Money** is the Senior Responsible Owner for Transparency activities.

Transparency Publication Model

Routine data publication



The publication model is co-ordinated by the Corporate Hub and involves the following steps each month:

- Extraction of data from the various source systems
- Review of the data, identification of exclusions.
- Approval by a Minister that a robust process has been used, and that Press Office is aware of any potential issues of that may arise from the data release.
- Publication of data to the DFID website, Contracts Finder website and data.gov.uk.

Research publication

The publication model for the DFID Research Open and Enhanced Access Policy involves the following steps:

- Research partners complete an Accessibility and Data Management Plan which outlines the strategy for maximising the opportunities to make research outputs open access
- Researchers implement their Plan through open access publishing, depositing outputs and metadata on R4D, archiving data in open access data repositories and seeking further opportunities to make data available online for free
- Compliance with the policy and with the agreed Accessibility and Data Management Plan is monitored as part of project annual and completion reviews
- Data on open access publishing is recorded and fed into the RED Results Framework

DFID RED works with other teams and departments that commission research within DFID to encourage adoption of the policy and similar working arrangements.

Information Principles

Information Principles 1 – 6 are addressed in these commitments as outlined in Annex 3. The Transparency Data Quality Improvement Plan will build on our existing monitoring framework. We will consider ways of encouraging good data quality as part of the design of existing and future systems as well as through improved guidelines.

Annexes

Annex 1: DFID Published Datasets

Area	Dataset	Contains	Where published	Format	Frequency
Cross-government datasets					
Finance	COINS	Spending data in HM Treasury format	data.gov.uk	CSV	annual
Finance	Spend > £500	Transactions for DFID >£500	DFID website	CSV	monthly
Procurement	Tenders and Contracts	Tender opportunities and contract documents	ContractsFinder website	documents	ad hoc
Procurement	Government Procurement Card spend over £500		DFID website, data.gov.uk	CSV	monthly
HR	Job titles, grades, names and salary bands for Directors and above		DFID website, data.gov.uk	CSV	6 monthly
HR	Junior staff salary data	Summary details of staff below SCS	DFID website, data.gov.uk	CSV	6 monthly
HR	Organogram	Of SCS posts and junior staff numbers	data.gov.uk	RDFa	6 monthly
HR	Workforce management information	Headcount and FTEs	DFID website	CSV	monthly
HR	SCS staff who received a bonus	Summary figures	DFID website	CSV	annual
Ministerial	Ministers' activities	Meetings, visits, gifts, hospitality, travel, media for ministers and Special Advisers	DFID website, data.gov.uk	CSV	quarterly

HR	Allowances and expenses incurred by senior DFID employees	Business expenses, meetings and hospitality for Management Board members.	DFID website	ZIP/CSV	annual
Corporate	Quarterly Data Summaries	Progress on the Business Plan	DFID website	XLS	quarterly
Environmental	Water and energy consumption in HQ buildings		ecodriver website		30 mins
Aid					
Aid	IATI	Information on DFID projects >£500	DFID website, IATI Registry, data.gov.uk	IATI XML, XML streaming of documents	monthly
Aid	Publishing Project Information website	Information on DFID projects >£500	DFID website	HTML + XLS download, documents	monthly
Aid	Pakistan Floods Monitor Data 2010	Detailed breakdowns of how UK aid is being allocated and spent	DFID website	XLS	one-off dataset
Aid	Statistics on International Development	Official financial resources allocated by the UK to support international development, broken down by destination country or organisation, type of assistance and purpose.	DFID website	ZIP/XLS	annual
Research supported by DFID					
Research	R4D	Full text research papers, datasets and links to those hosted elsewhere	DFID Website	Documents and metadata in Excel, XML, OAI-PMH and RDF linked open data	ongoing

Annex 2: List of data to be released

For the first time during period 1 April 2012 to 31 March 2014

Dataset Name	Description of data (including fields)	To be first published on (i.e. date, including month and year)	Frequency of update (e.g. monthly, quarterly, annually)	Issued under the Open Government Licence (OGL) Yes or No* ¹⁵ (*If no, please provide details of alternative publication arrangement)
Results of aid projects	Results of aid interventions in DFID's Results Framework	June-12	Annual	Yes
Project scoring	Project scores in the new scoring model (as part of IATI dataset)	Feb-13	Monthly	Yes
Geocoded aid projects	Detailed mapping of aid projects (as part of IATI dataset)	Apr-14	Monthly	Yes

¹⁵ The OGL allows the public and business to, for free: copy, publish, distribute and transmit the Information; adapt the Information; exploit the Information commercially for example, by combining it with other Information, or by including it in their own product or application. The full licence can be found at <http://www.nationalarchives.gov.uk/doc/open-government-licence>

Annex 3: Information Principles

This chapter describes how the **Information Principles for the UK Public Sector**¹⁶ apply to the DFID Open Data Strategy.



Information is a Valued Asset

The delivery of our currently published datasets and the list of new data to be released under our future plans (Annexes 1 and 2) demonstrate the high priority that we give to managing data from our aid projects and other corporate information. The lists provide evidence of how we are developing the capabilities we need to manage data as one of our most valuable assets. We have an Information Asset Register that assigns responsibilities for each asset to an Information Asset Owner. All owners must have the required training and understand the nature and purpose of the asset. They are assigned responsibility for monitoring the usage, addressing the risks and ensuring the asset is fully used for the public good.

Information is Managed

We have a governance structure and information management and assurance roles, policies and guidelines that ensure we address legal and regulatory requirements, review the risks and manage and protect our data throughout its lifecycle. We have an electronic documents and records repository (Quest) with associated records management processes that are monitored both by local Information Managers and by a central team. Our financial and project management data links to this repository. Our processes for retention and disposal are being reviewed in the light of digital continuity guidelines from The National Archives and DFID's business requirements. We have incorporated digital continuity into the

¹⁶

http://www.cabinetoffice.gov.uk/sites/default/files/resources/Information_Principles_UK_Public_Sector_final.pdf

corporate process for the adoption of proposed new projects. We have an Information Sharing Policy that emphasises how sharing with external partners is central to DFID's transparency commitments but also points staff to the security and policy considerations that they need to take into account. Transparency is addressed in DFID's Risk Register and this is monitored regularly. Guidance on the requirements for Data Protection is provided to all staff and training is mandatory for all staff routinely handling personal data. The new Civil Service Training module on core information management skills is to be used by staff in DFID. A review of DFID's current Information Management Maturity Model is planned for the summer of 2012. This follows the Knowledge Council model.

Information is Fit for Purpose

Initiatives are under way to further train and guide staff in what makes our data fit for purpose and related data quality issues. A central team has responsibility for monitoring how well guidelines are being met and addressing any problem areas. Information Managers train staff in good information management and the use of Quest and they support a central team in monitoring performance of staff in following the guidelines. Quality of aid project data is a priority for improvement. Ways of encouraging better data quality are considered when systems are being developed. The technical design process aims to ensure that we provide platforms and formats that are aligned with the way the information is to be used.

Information is Standardised and Linkable

We are committed to adopting Open Standards wherever possible and to promoting these in the development field. Under the International Perspective of DFID's Open Data Model we publish data in an open, internationally agreed standard that others can reuse and repurpose. The plans we have set out above show how we aim to further develop our ability to provide data in standardised, linkable formats, to enable re-use and publication in other formats. They also show that we are planning to pilot a methodology for linking our aid budget data with partner country budget data. We link our funded research R4D data to our aid project data. We apply common indexing standards wherever possible and can produce data to open standards as required. A Management Information Project is looking at better standardisation of concepts such as roles and the use of consistent identifiers across our systems.

Information is Re-used

We are already offering re-use opportunities for our aid project, research and corporate data through the DFID website and data.gov.uk and we aim to actively assess further opportunities that offer real benefit as part of reviews of the Open Data Strategy and as new assets are brought into use. We are making our aid project data available in machine readable format through the IATI Registry. We aim to promote re-use through initiatives such as the planned geocoding pilot. We aim to investigate better tracking of the use of UK aid with data from partner organisations being made available. The new Open and Enhanced Access Policy for our DFID funded research will provide access to more of our data and documentation for re-use as is occurring with the Young Lives project. Staff are provided with clear guidelines on the importance of data quality to allow effective re-use of

data and searching across systems. Our Information Sharing policy seeks to balance the need to exploit and to protect our data and the risks are monitored under the Transparency Section of our Risk Register. Our approach to managing reference data is under investigation.

Public Information is Published

We have a Framework in place for responding to legal obligations around access to our information under the Freedom of Information Act, Data Protection Act and Environmental Information Regulations. A central team deal with enquiries and our performance in meeting the targets is recorded on the Ministry of Justice website along with that of other Government Departments. We publish both unstructured and structured information e.g. we provide access to project documents as well as project data as part of our Projects Database. Those who do not have access to the Internet can access data through the Public Enquiry Point, by letter or telephone. We have carried out surveys to inform our plans for future data releases and these plans are reviewed and updated on a regular basis. Transparency is taken into account in the development of new systems and legacy information is considered for inclusion when upgrades are being planned and where clear benefit outweighs additional effort. Staff are provided with clear guidelines on the need for informative titles and the need to spell out acronyms etc. to provide the context needed for information in the public domain.

Citizens and Business can Access Information about Themselves

Where we hold information on members of the public (for example overseas pensioners, correspondence with members of the public, suppliers) we use existing channels to allow people to access this information. Feedback from pensioners suggests that online access to their own data would not meet their needs due to the nature of the community and their being based overseas. Data showing a high rate of satisfaction with the existing service is available in the Annual Report and we are considering publishing this as a data set. The Public Enquiry Point is managed in the Openness Unit along with staff responsible for Freedom of Information and Data Protection. The Openness Unit is the first port of call for enquiries from citizens wishing to access their data.