

# Open Data Maturity Model

The open data maturity model is a way to assess how well an organisation publishes and consumes open data, and identifies actions for improvement.

The model is based around five themes and five progress levels. Each theme represents a broad area of operations within an organisation. Each theme is then broken into areas of activity, which can then be used to assess progress.

## Themes

1. Data management processes — identifies the key business processes that underpin data management and publication including quality control, publication workflows, and adoption of technical standards.
2. Knowledge & skills — highlights the steps required to create a culture of open data within an organisation by identifying the knowledge sharing, training and learning required to embed an understanding of the benefits of open data.
3. Customer support & engagement — addresses the need for an organisation to engage with both their data sources and their data re-users to provide sufficient support and feedback to make open data successful.
4. Investment & financial performance — covers the need for organisations to have insight into the value of their datasets and the appropriate budgetary and financial oversight required to support their publication. In terms of data consumption, organisations will need to understand the costs and value associated with their re-use of third-party datasets.
5. Strategic oversight — highlights the need for an organisation to have a clear strategy around data sharing and re-use, and an identified leadership with responsibility and capacity to deliver that strategy.

For information about the design of the model and its application, consult the model documentation

<b>Authors</b>	Leigh Dodds Andrew Newman	leigh.dodds@theodi.org andrew.newman@defra.gsi.
<b>Status</b>	Edition 1.0	
<b>Published</b>	31 March 2015	
<b>Licence</b>	<a href="#">Creative Commons Attribution Licence</a>	

## 1. Data management processes

The key business processes that underpin data management and publication including quality control, publication workflows, and adoption of technical standards.

Activity	Aspect	Level 1 - Initial	Level 2 - Repeatable	Level 3 - Defined	Level 4 - Managed	Level 5 - Optimising	Beneficial effects
Data release process	Publication	<p>Little or no published open data.</p> <p>Datasets that are published are done so using ad hoc, possibly manual processes.</p> <p>Released datasets are rarely, if ever, updated.</p>	<p>Specific projects or products may have defined a repeatable process for releasing an individual dataset or series, but there are no common standards.</p> <p>Some datasets are being released and updated to a regular schedule.</p>	<p>There is a repeatable organisation-wide standard release process for publishing datasets.</p> <p>The release process may be tailored for individual projects.</p> <p>The release process has been adopted for some published datasets.</p>	<p>All datasets are released according to the standard organisational process.</p> <p>Updates to datasets are released to a scheduled, published timetable.</p> <p>Archived versions of key datasets are provided to support historical analysis.</p> <p>Internal reports are available to highlight stale datasets.</p>	<p>The organisation collects and monitors metrics on its release process, e.g. time taken to release and refresh datasets.</p> <p>The organisation acts to optimise the release process to reduce time between changes and releases.</p>	<p>Reduce overheads associated with data releases.</p> <p>Become more efficient at delivering data to re-users.</p> <p>Data is more likely to be up to date and be more useful to re-users.</p>
Standards development & adoption	Publication & re-use	<p>Datasets are not released using common standards.</p>	<p>Specific projects may have agreed on specific technical standards for formatting and packaging specific types of data, but these are not universally applied.</p> <p>Standards will be driven by internal priorities and needs.</p>	<p>The organisation has defined a set of key technical standards that will be used when publishing datasets. This list of standards is published internally. Standards will address data formats (e.g. CSV) as well as structure of the data (e.g. schemas).</p> <p>The organisation has created standard code-lists and identifiers which will be used in its datasets.</p> <p>The standards, codes and identifiers have been adopted for some, but not all, datasets.</p>	<p>All datasets are released to conform with the organisation's technical standards.</p> <p>The organisation begins using open standards in preference to closed or bespoke alternatives.</p> <p>Datasets across the organisation are using standard code lists and identifiers. Some datasets may be using open identifiers &amp; codelists.</p>	<p>The organisation has defined a strategic approach for managing its data standards.</p> <p>The organisation monitors technology trends to ensure that its datasets are released according to evolving technical standards and best practices.</p> <p>The organisation is routinely using third-party code-lists and identifiers in its datasets to help align them with other data sources.</p> <p>The organisation is routinely adopting applicable open standards in preference to closed or bespoke alternatives.</p>	<p>Reduce/remove costs of maintaining bespoke systems, technical standards and code lists.</p> <p>More agile adoption of new technologies.</p> <p>Data becomes easier to use.</p> <p>Data from different departments, or different organisations, is easier to combine.</p>

Data governance	Publication & re-use	<p>Datasets are not managed in any consistent way.</p> <p>There is no clear ownership around internal datasets.</p>	<p>Specific teams or products have begun defining their own lightweight governance processes.</p>	<p>The organisation has defined a standard data governance process that has been applied to high value datasets.</p> <p>The process addresses key organisational issues and considers the complete dataset lifecycle.</p> <p>Datasets have a well-defined owner through their lifecycle.</p>	<p>The standard data governance process is applied to all datasets. The process may be tailored to the needs of specific projects.</p> <p>The organisation routinely monitors data quality for its high-value datasets.</p> <p>Feedback from re-users is used to help drive improvements in data quality where appropriate.</p>	<p>The organisation is working to improve its data governance processes, e.g. by monitoring the process to address circumstances where data governance may lapse.</p> <p>The organisation is working in partnership with re-users to help improve &amp; maintain data quality. This may include accepting external contributions.</p>	<p>Higher quality data.</p> <p>Data assets have clear owner.</p>
Managing sensitive data	Publication	<p>The organisation has not defined any approaches for aggregating or anonymising data.</p> <p>Datasets that contain sensitive information are not considered for release.</p>	<p>Some open datasets based on sensitive information are released, and efforts are made to remove or aggregate data. Where personal information is released, appropriate consent is obtained.</p> <p>The processes for desensitising the data and carrying out risk assessments are ad hoc.</p>	<p>The organisation has identified standard approaches for anonymising and aggregating data and begins applying these techniques on new projects.</p> <p>Risk management is a component of the data governance process.</p>	<p>Standard anonymisation and aggregation approaches as well as risk assessments, are applied to all datasets prior to publication.</p> <p>The organisation seeks independent validation prior to release of high risk datasets.</p>	<p>The organisation routinely assesses all datasets, during their life-cycle to highlight sensitive information.</p> <p>The organisation monitors the changing data landscape, e.g. to recognise and address emerging issues. This may involve changing the approach to anonymisation, or even stopping a data release.</p>	<p>More data can be released.</p> <p>Reduced risks and liabilities associated with data release.</p>

## 2. Knowledge & skills

The steps required to create a culture of open data within an organisation by identifying the knowledge sharing, training and learning required to embed an understanding of the benefits of open data.

Activity	Aspect	Level 1 - Initial	Level 2 - Repeatable	Level 3 - Defined	Level 4 - Managed	Level 5 - Optimising	Beneficial effects
Open data expertise	Publication & re-use	<p>The organisation does not provide any direct training or support for open data skills.</p> <p>Early adoption is driven by enthusiasm of self-taught internal evangelists.</p>	<p>No shared understanding around open data in the organisation, although some areas of expertise are developing.</p> <p>Training &amp; support for individuals and/or teams is driven by the needs of specific projects.</p> <p>The organisation relies on external expertise for key skills.</p>	<p>The organisation is aware of where further support and understanding is required, identifying key skills and expertise for development.</p> <p>Training &amp; support around open data topics is offered to some teams. This is largely focused on building operational skills (e.g. licensing, technology).</p> <p>The organisation has identified internal experts that may help support &amp; mentor others.</p>	<p>The organisation is actively building a shared understanding around the practice of open data. This covers common knowledge, operational requirements and strategic insight.</p> <p>Open data awareness &amp; training becomes a standard part of staff induction and development.</p> <p>Staff are clear on how open data impacts their personal roles and responsibilities.</p>	<p>Knowledge and understanding of open data exists at all levels in the organisation.</p> <p>Staff are clear how open data guides and informs organisational strategy.</p> <p>The organisation acts to build &amp; foster networks of expertise both within the organisation and the wider community.</p>	<p>Reduce reliance and cost of using external expertise.</p> <p>More efficient use of internal skills/resources.</p> <p>Greater ability of the organisation to take advantage of open data as a tool to achieve its goals.</p> <p>Greater ability of individuals to use open data to help them be more informed as they do their jobs.</p>
Knowledge management	Publication & re-use	<p>No standard approach for capturing important learning from open data projects, or for ensuring knowledge transfer between staff at end of projects.</p>	<p>Individual teams may be implementing processes to capture useful information relating to their open data activities.</p> <p>These processes may focus more on knowledge transfer at the end of a project, or when team members change, rather than across the full lifecycle.</p>	<p>Organisation defines policies relating to knowledge capture, transfer and sharing.</p> <p>Knowledge capture starts to be applied throughout a project or dataset lifecycle rather than at the end.</p>	<p>All staff have access to relevant documentation and policies that relate to the support and use of data assets.</p> <p>The location of documentation and other outputs relating to datasets is clearly defined.</p>	<p>The organisation actively encourages sharing of data and knowledge internally. For example, through internal groups, presentations, etc.</p> <p>The organisation is engaging with others in its sector to share experiences.</p>	<p>Better internal use and awareness of how data can be applied.</p> <p>Knowledge of key datasets not lost with staff turnover.</p>

### 3. Customer support & engagement

The requirements for an organisation to engage with both their data sources and their data re-users to provide sufficient support and feedback to make open data successful.

Activity	Aspect	Level 1 - Initial	Level 2 - Repeatable	Level 3 - Defined	Level 4 - Managed	Level 5 - Optimising	Beneficial effects
Engagement process	Publication	<p>Little or no attempt made to identify potential re-users for released open data.</p> <p>Data releases are either driven by internal priorities, or based on demands from re-users.</p> <p>No attempts are made to measure levels of engagement.</p>	<p>Some teams attempt to identify and engage with potential re-users of the datasets they are publishing.</p> <p>The organisation has identified a means by which potential re-users can request data.</p> <p>Some ad hoc attempts are made to assess the impact of data publishing efforts.</p>	<p>The organisation identifies a repeatable approach for engaging with external users which is tailored, as appropriate, to support individual data releases.</p> <p>The engagement process addresses outreach to the community both before and after the release of data.</p> <p>The organisation begins to use demand for data from re-users to guide its engagement strategy and inform release of data.</p>	<p>The organisation begins tracking the effectiveness of its interventions as an open data publisher, identifying standard ways to measure and monitor impact of key data releases.</p>	<p>The organisation is routinely tracking metrics relating to all its open data publishing efforts.</p> <p>The organisation is working to increase engagement within identified, high-value stakeholder communities and for high-value datasets.</p>	<p>Greater insight into re-user needs &amp; activities.</p> <p>Better relationships with partners in sector.</p> <p>Better prioritisation of releases and development of new features.</p> <p>Greater ability to justify data releases internally and externally in business cases.</p> <p>Measurable ROI.</p>
Open data documentation	Publication	<p>Datasets are released with little or no documentation.</p> <p>The level of documentation released with datasets is based on the enthusiasm/drive of individuals involved in the release process.</p>	<p>Individual teams begin to identify the essential documentation required to support some of data they are releasing.</p> <p>Some common templates or approaches are defined to creating and managing documentation.</p>	<p>The organisation has identified a standard set of documentation and metadata that should be released with each dataset.</p> <p>Only some, e.g. high-value datasets, are released with this standard set of documentation and metadata.</p>	<p>All datasets are released with supporting documentation and metadata using a standard template.</p> <p>Creation and maintenance of supporting documentation is part of the data publishing process.</p> <p>The organisation invests in additional documentation and supporting materials to help re-users.</p>	<p>The organisation routinely releases high quality metadata and documentation as appropriate for all of its datasets.</p> <p>Externally published documentation is reviewed by key internal stakeholders before release.</p> <p>Re-users are invited to provide feedback and contribute improvements to published resources. The organisation promotes the availability of third-party learning resources &amp; tools.</p>	<p>Easier for re-users to use data.</p> <p>Increase use of published data.</p> <p>Lowers risk of misuse of the data.</p>

Re-user support processes	Publication	<p>No user support process.</p> <p>Re-users need to seek out and engage with individuals in the organisation in order to ask for support.</p>	<p>Individual teams provide ad hoc support for re-users of their data.</p> <p>Some data releases may provide contact points for re-users to ask for support.</p>	<p>The organisation creates user forums as a means to collect external feedback and discussion around datasets and to communicate with its users.</p> <p>The type and level of support available to re-users is clearly indicated on some datasets.</p>	<p>The organisation has an identified support team that helps re-users. This team is the primary means for re-users to seek help.</p> <p>The organisation identifies level of support available to re-users for all datasets. For example, expected response times.</p> <p>The organisation engages with re-users via multiple channels, e.g. social media, discussion forums.</p>	<p>The organisation acts to measure the cost and efficiency of its support operations, creating additional learning materials and community support groups.</p> <p>The organisation ensures that it is delivering on its support arrangement for all datasets. For example, that response times are within goals.</p>	<p>More efficient and effective support processes.</p> <p>Easier re-use of published data.</p> <p>Lower risk of misuse of the data.</p>
Community norms	Re-use	<p>The organisation may not be identifying and acknowledging all of its data sources as required by licensing.</p>	<p>Individual teams work to ensure that they attribute use of third-party open data sources as required by licensing.</p> <p>Individual product teams or business units may share their experiences in using open datasets to solve specific problems or build applications.</p>	<p>The organisation defines process to review use of third-party datasets to ensure use of the data. For example, attribution requirements, conforms to licensing terms.</p> <p>This process is applied across the business.</p>	<p>The organisation is transparent about all open datasets it uses, regardless of licensing issues.</p> <p>The organisation actively engages with its open data suppliers to share experiences, provide feedback.</p>	<p>The organisation shares its experiences with using open data more widely, creating case studies to highlight value to its sector.</p> <p>The organisation engages with its sector to help promote further open data activities.</p>	<p>Visibility as an innovator.</p> <p>Key player in sector.</p> <p>Cost savings and efficiencies deriving from increased open data adoption in the sector.</p> <p>Reduced risk of key re-used datasets becoming unmaintained.</p>

#### 4. Investment & financial performance

The requirements for organisations to have insight into the value of their datasets and the appropriate budgetary and financial oversight required to support their publication. In terms of data consumption, organisations will need to understand the costs and value associated with their re-use of third-party datasets.

Activity	Aspect	Level 1 - Initial	Level 2 - Repeatable	Level 3 - Defined	Level 4 - Managed	Level 5 - Optimising	Beneficial effects
Financial oversight	Publication & re-use	Data releases are unfunded and done as exceptional expenditure.	Individual projects may include open data publication costs as part of their budget - typically not separately addressed.	Project funding and operational costs routinely include long terms costs for open data publication.	<p>The organisation actively monitors the financial costs and benefits of open data publication and re-use.</p> <p>The organisation has clarity around cost/benefits of open data versus ongoing data governance.</p>	<p>The organisation looks for efficiency savings around open data publication and re-use.</p> <p>For example, by looking for ways to reduce data management and licensing overheads through adoption of open data.</p>	<p>Costs of publishing and using data clearly understand.</p> <p>Potential cost savings by more effective use of internal and external open data.</p>
Dataset valuation process	Publication & re-use	No attempt made to quantify the value of either published or re-used datasets.	<p>The organisation carries out some reporting of the value of some datasets, eg. to a specific community. The valuation is not quantified and is largely couched in general benefits.</p> <p>The approach to valuation is ad hoc. For example, it may be done retrospectively to justify ongoing release of a dataset but is not always carried out prior to release.</p>	<p>The organisation begins defining more quantitative ways to measure the value and ROI associated with both the publication and re-use of datasets.</p> <p>Valuation decisions begin to inform the development of business cases in support of specific releases.</p>	<p>The organisation adopts one or more standard approaches for quantifying the value of its data assets.</p> <p>This approach is applied to all key datasets, including those its re-uses from third-parties.</p>	<p>Organisation routinely values all data assets. The valuation is used to drive all investment decisions relating to both data releases and use of external data.</p> <p>The metrics and approaches to measuring value are regularly assessed to ensure they align with organisational goals.</p> <p>Organisation is transparent about its approach to valuing datasets.</p>	<p>Investment in data strategy and governance can be driven by value of assets.</p>
Open data in procurement	Publication & re-use	<p>Procurement processes and contracts do not address data supply or re-use.</p> <p>Organisation is unable to release any data that results from, or is derived from existing contracts.</p>	The organisation begins seeking retrospective clarity around data rights. This is driven by the needs of specific projects that seek to open data.	<p>The organisation begins tailoring individual contracts and procurement activities to address data licensing, as required by the needs of specific projects/products.</p> <p>The organisation begins to use open data support as a differentiator between suppliers.</p>	<p>The organisation develops standard clauses in contracts and ITTs to ensure there is clarity around rights clearance and data re-use.</p> <p>The organisation begins to review and revise existing contracts where necessary and appropriate, to ensure that rights are clear.</p>	<p>All ITTs and tenders from the organisation include reference to open data where appropriate for the contract. Requirements are proportionate to the type of contract.</p> <p>The organisation begins tracking whole-life cost of contracts.</p>	<p>Clarity around rights associated with all data assets.</p> <p>Reduced risks associated with misuse of data.</p> <p>Reduce whole-life cost of contracts.</p>

## 5. Strategic oversight

The requirements for a clear strategy around data sharing and re-use, and an identified leadership with responsibility and capacity to deliver that strategy.

Activity	Aspect	Level 1 - Initial	Level 2 - Repeatable	Level 3 - Defined	Level 4 - Managed	Level 5 - Optimising	Beneficial effects
Open data strategy	Publication & re-use	<p>The organisation has no strategy or policy with regards to open data.</p> <p>Initial data releases are seen as R&amp;D exercises or are driven by external influence.</p>	<p>Individual business units identify benefits to open data for their individual activities.</p> <p>No coherent strategy across the organisation with regards to open data.</p>	<p>The organisation defines an open data strategy. This will include elements addressed in other activities, including data valuation, governance, standards, etc.</p> <p>There is clear responsibility and budget support for delivery of the open data strategy, at a senior management level. Ownership for key processes, such as data governance, is also well understood.</p>	<p>The organisation has aligned delivery on open data policy and strategy with organisational objectives.</p> <p>The organisation has set measurable targets for implementation of the strategy.</p> <p>Performance assessment of key executives is tied to delivery on the objectives laid out in the strategy.</p>	<p>The organisation is using open data as a key element of its overall organisational strategy.</p> <p>Metrics and goals are reviewed and adjusted over time, to continue to stretch the organisation and to ensure that the open data goals are consistent with those of the organisation.</p>	<p>Clarity within the organisation and externally about open data strategy.</p> <p>Coherent data activity across the organisation.</p>
Asset catalogue	Publication & re-use	<p>No approach to managing data sources as assets.</p>	<p>Individual teams or projects maintain directory of data assets and data resources that they use.</p>	<p>An organisation wide asset catalogue has been developed to identify key datasets being published and used.</p> <p>All of the organisation's high-value datasets are in the catalogue.</p> <p>Elements of the catalogue may be publicly and openly shared. For example, to identify datasets that are being released, planned for release, or will not be released.</p>	<p>All datasets published from or used by the organisation are entered into the asset catalogue.</p> <p>New products/projects attempt to re-use datasets referenced in the catalogue, rather than creating new assets.</p>	<p>The organisation tries to identify cost and efficiency savings by identifying overlaps and commonalities between datasets, and either aligning or replacing them as necessary.</p>	<p>Data is managed as an asset.</p> <p>Cost savings through use of external data, removal of duplicate data.</p>