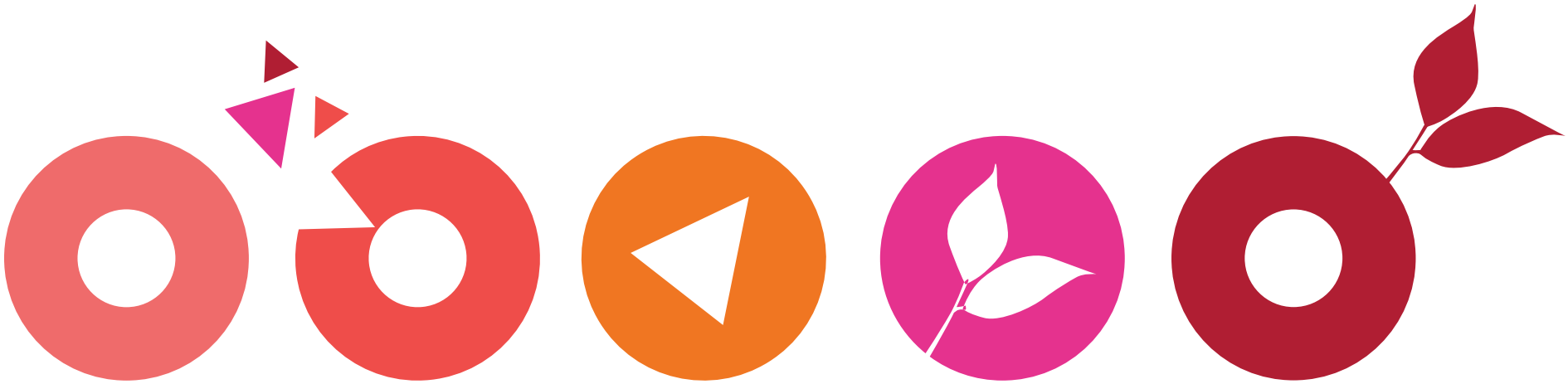
DATAORCHARD

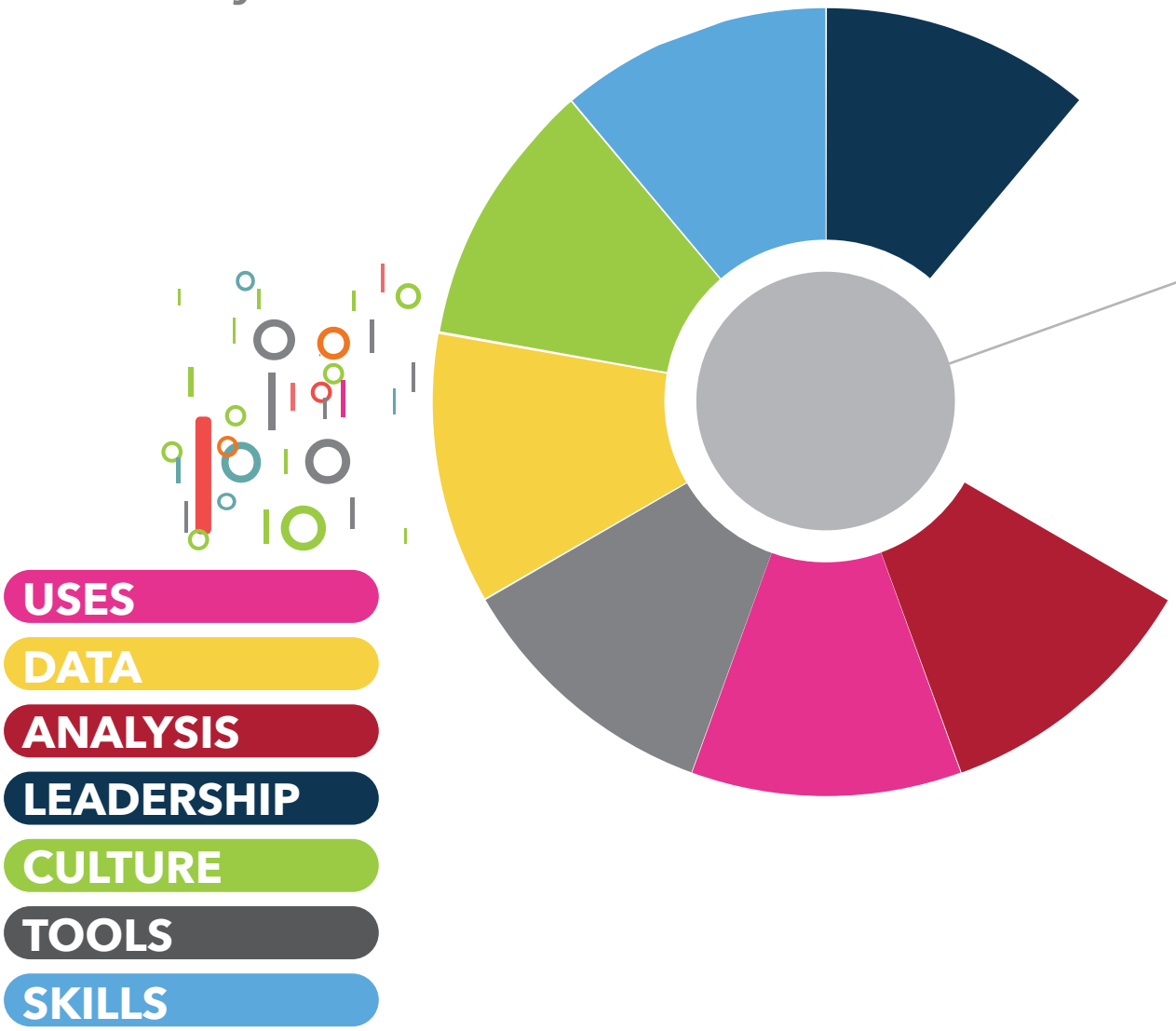
Data Maturity Framework for the Not-For-Profit Sector

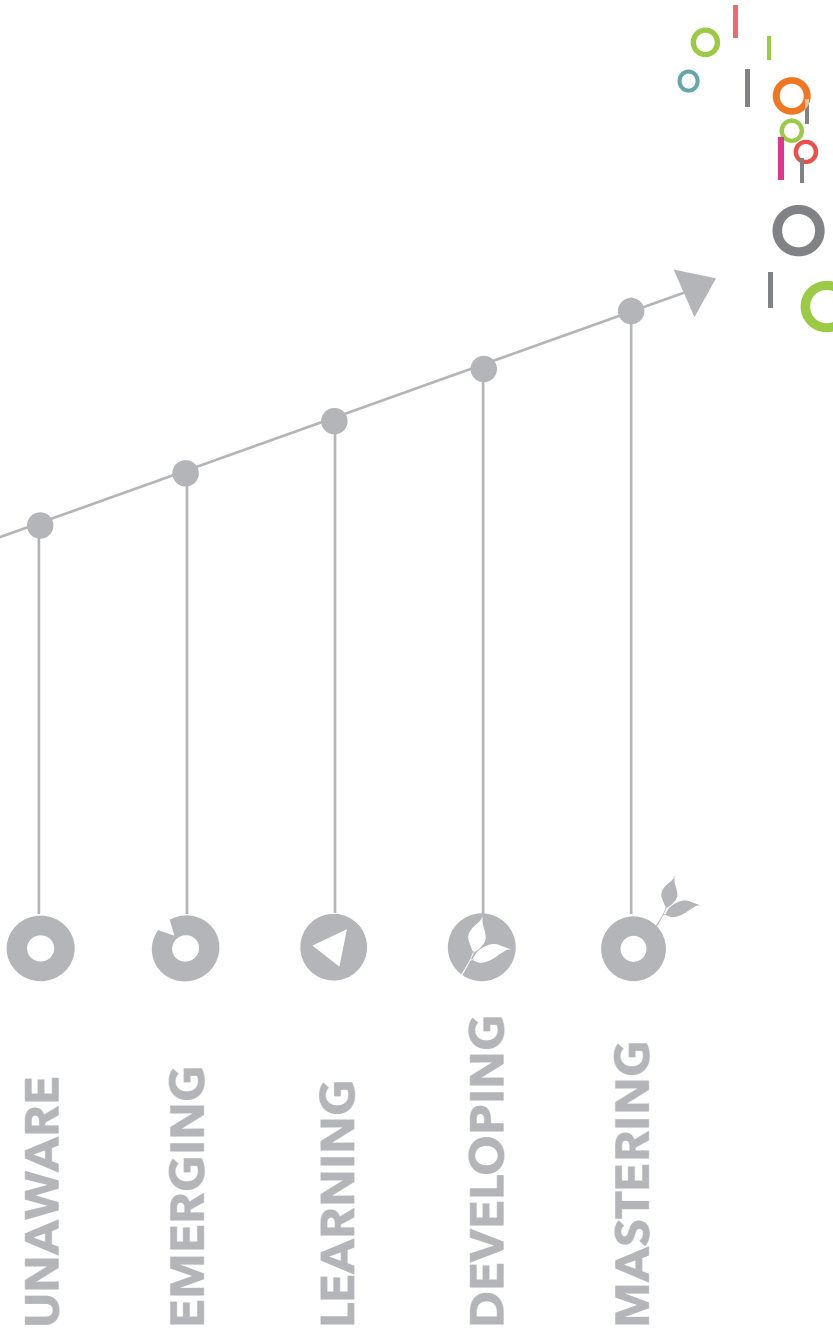


Data Orchard | info@dataorchard.org.uk | www.dataorchard.org.uk

Version 2 created by Data Orchard CIC September 2019, replacing version 1 published in January 2017 in partnership with DataKind UK.

The Data Maturity Framework





7 KEY THEMES

5 STAGE JOURNEY

Data Maturity Themes

USES

- Purposes for collecting and analysing
- Benefits and rewards

DATA

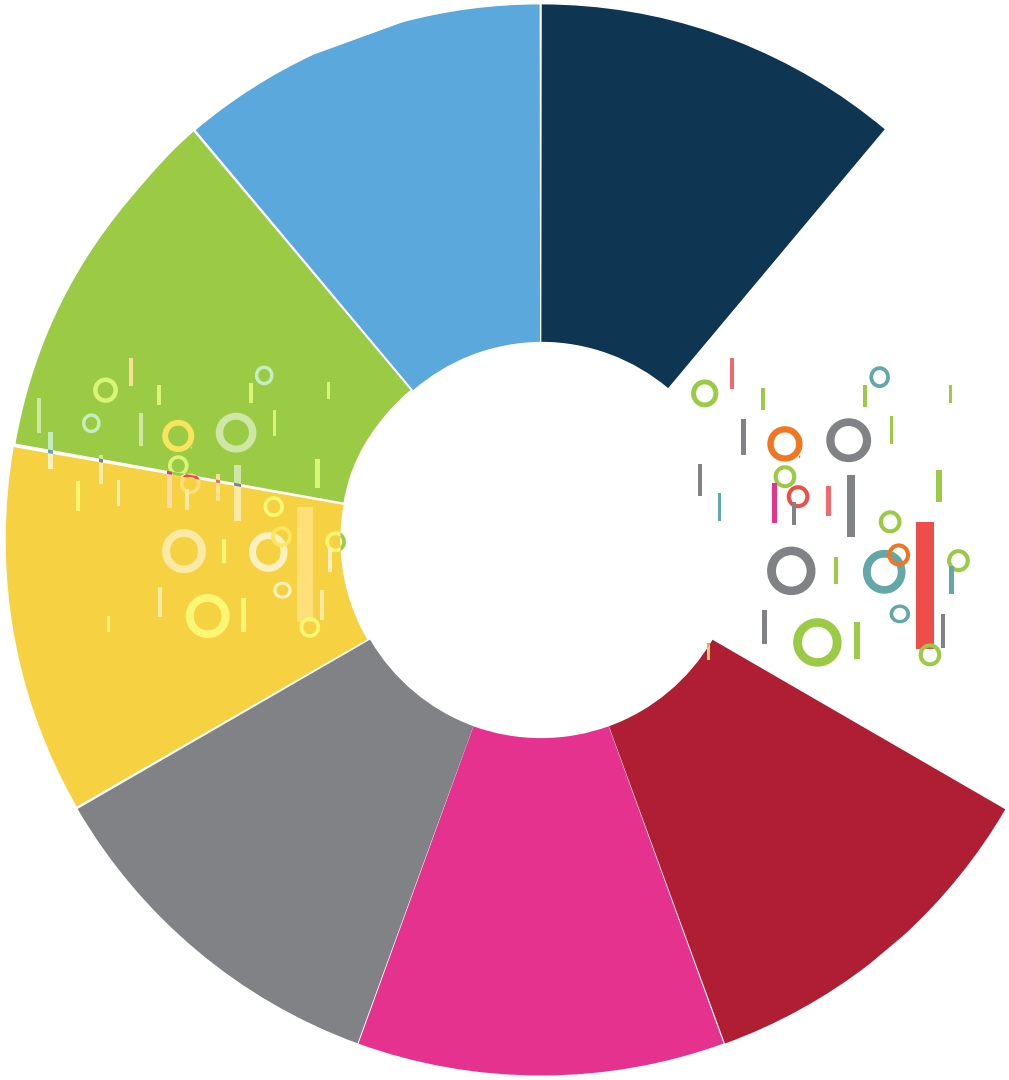
- Collection
- Quality
- Sources
- Assets

ANALYSIS

- Type
- Technique
- Joining
- Presenting

LEADERSHIP

- Attitudes
- Plans
- Capability
- Investment



CULTURE

- Team approach
- Self-questioning
- Openness
- Protection

TOOLS

- Storage
- Type
- Quality
- Sharing
- Integration

SKILLS

- Capacity
- Skills
- Training
- Access to knowledge and expertise

USES

Purposes for collecting and analysing data | Benefits and rewards

UNAWARE
SCORE 0-1

- Records basic client information and activities/work delivered in order to operate at a basic level.
- Little or no benefits or rewards in relation to services.
- Collect and use data for requisite purposes e.g. basic financial management and legal/funder/ contract compliance reporting.
- Continued funding may be seen as the only reason for collecting data.

EMERGING
SCORE 1-2

- Collect more data than required by legal/funders/contracts.
- Most data collection and analysis relates to capturing activities, measuring outputs, and basic financial analysis and forecasts.
- Rewards mostly around improved understanding of beneficiaries and ability to articulate the scale of activities delivered.
- Raising income and understanding fundraising performance likely to be key focus for additional data collection e.g. fundraising events, donors, grants, contracts, and sales.
- Able to feedback information to funders, partners and networks around specific projects/services and the scale of activities. May struggle with multiple and or/repeat service use data.
- Starting to explore the difference between outputs and outcomes.

LEARNING
SCORE 2-3

- Collect data to be able to understand and evidence the types of clients/ needs and problems the organisation addresses. Uses both internal and reliable external data sources to do so.
- Starting to use data to understand different ways beneficiaries initially contact and engage with services over time (or not).
- Able to demonstrate activities being delivered for specific types of users across a range of projects and services.
- Capturing some outcomes data and learning to measure outcomes consistently.
- Starting to capture feedback, monitoring and evaluation data on service/product quality and performance to inform improvements.
- Use data for income generation and some forecasting of sales and donations leading to more effective fundraising and commercial income.
- Starting to lead conversations with funders and partners using data.
- Building internal knowledge and expertise based on the analysis of data and dialogue on how to act on this.
- Strategic planning, particularly around efficiency and service development, is becoming more data informed.

DEVELOPING
SCORE 3-4

- All client, activity, output, and outcomes data is routinely collected.
- Services/products/campaigns are monitored to show performance on how, when and where these are used by whom.
- Services/products/campaigns are starting to be targeted at specific demographics and/or geographic locations in response to better understanding of needs/problems.
- Services/products/campaigns are regularly reviewed and adapted in response to data to optimise outcomes.
- Beginning to test assumptions on how the organisation has an impact and for whom.
- Starting to differentiate between approaches and understand what's working and what's not.
- Able to differentiate and explain attribution and contribution in evidencing and communicating impact.
- Exploring and learning how to measure long term impact.
- Data is starting to be used to inform efficiency savings (resources, processes and service/product design).
- Can coherently make the case to funders/investors/clients for existing and new services/products/campaigns.
- Learning and evaluation is embedded internally and informs both service and staff/volunteer performance.
- Strategic planning and decision making is becoming considerably data informed.

MASTERING
SCORE 4-5

- Data is used extensively, and in inter-related strategic ways, for a wide range of purposes.
- Sophisticated use of analysis delivers insights and predictions to influence service and organisational success.
- Evidencing and improving outcomes and impact is a primary focus. Experiments to identify differentiated impact and how to optimise this.
- Predict user needs and service/product options based on understanding client behaviours and how to influence these for the best outcomes.
- Design and delivery of services/ products/campaigns is optimised at an individual/personal level.
- Services and interventions are highly targeted possibly in collaboration with other partners/service providers.
- The organisation is embedded in networks of knowledge and research in the context of its work. Partnerships and networks are strengthened through collaborative data sharing.
- Use data to increase efficiencies (resources, processes, services/product delivery).
- Robust evidence ensures credibility and is used to influence external policy and decision makers.
- Learn, evaluate and build knowledge - harness data for continuous improvement of products/services/ campaigns.
- Strategic planning and decision making is highly informed by data and based on past, present and future analyses.

DATA

Collection | Quality | Sources | Assets

UNAWARE
SCORE 0-1

- Limited data (if any) collected.
- Collected manually, mostly on paper, only when needed for specific purpose.
- Data isn't meaningful or useful to the organisation.
- Infrequently updated (if ever).
- Not checked for validity or accuracy.
- Data isn't shared internally.
- No external data sources used.
- Nobody is aware or interested in the data assets in the organisation.

EMERGING
SCORE 1-2

- Data collection is both on paper and in digital forms though there may be inconsistencies and inefficiencies in approach.
- Staff and volunteers may not be trained in data collection.
- Data is regarded as meaningful or useful primarily for meeting external legal/funder/contract requirements.
- It may be patchy and inconsistent with gaps and duplications.
- Data is rarely updated and cleaned.
- Mixed levels of confidence and trust in data.
- Starting to find out what data is available internally. Know where most data assets are but there may be more squirrelled away in parts of the organisation.
- Occasional use of external information sources relating to the wider context of the organisation's work.

LEARNING
SCORE 2-3

- Though errors remain, data collection methods and processes are being improved.
- Staff and volunteers are trained in data collection.
- Data is reviewed to assess how relevant, meaningful, and necessary it is.
- The organisation knows how good or bad its different data sets are; and knows which data sources can/can't be trusted.
- Gaps, overlaps, and mismatches in the available data have been identified.
- Data becoming richer, more relational and therefore versatile.
- Starting to explore how internal data (perhaps in multiple, fragmented or isolated locations) can be shared and utilised.
- Additional internal and external data is sourced.
- All data assets are known but may not be formally recorded.

DEVELOPING
SCORE 3-4

- Data requirements defined and consistently collected via a range of methods.
- Investment in staff and volunteer training to ensure consistency and quality of data collection.
- Tests and refines how relevant, meaningful, and necessary its data is.
- Focuses on collecting the right data for clearly specified purposes.
- Data is monitored for quality including completeness, accuracy, and validity. Tools and systems exist for cleaning and maintenance.
- Richer data collection with more integration/alignment between systems reduces duplication, inefficiency and error.
- Data sharing internally is becoming the norm.
- Occasional commissioning of independent research and evaluation.
- Extensive use of externally available research and evidence around needs and successful interventions.
- Open datasets are occasionally used.
- Exploring shared measures and benchmarks with other organisations.
- Recorded list of all data assets, whether they include personal/sensitive data, and assigned responsibility.

MASTERING
SCORE 4-5

- Focuses effort on collecting the minimal amount of the most important and valuable data.
- Staff and volunteers are trained in data collection and collection is automated where possible.
- Knows all its data is relevant, meaningful and necessary.
- Monitors and fully understands the quality of the data it holds and hence has high levels of confidence and trust in its data.
- Data is versatile and re-usable for multiple purposes and audiences.
- Invests in resources to collect, clean, maintain, and manage data well across the organisation.
- Shares data internally from different teams, departments and services.
- Compares its data with other organisations through shared measures and benchmarks.
- Uses publicly available external research (e.g. government/academic) and may contribute to this.
- Regular use of valuable open/public data sets.
- Commissions external independent research and evaluation.
- Maintains full inventory of data assets across the whole organisation with clearly defined variables, ownership, review periods, and development plans for each.

ANALYSIS

Type | Technique | Joining | Presenting

UNAWARE
SCORE 0-1

- Limited analysis of financial, contract-related, and legal compliance data.
- Mainly simple manual counts from paper forms or digital records are used.
- Strategic discussions are not informed at all by analysis of data.
- Data is not used in reports. Anecdotes and observations are preferred.

EMERGING
SCORE 1-2

- Analyses starting to explore service users/clients/customers and target audiences.
- May include use of external research relating to the context e.g. to evidence scale of need/problems.
- Basic analysis, using counts and spreadsheets.
- People verbally report on data as part of strategic discussions.
- Analysis and report creation skills are being explored though results may be variable.
- Bar charts and pie charts may be the only types of analytical presentations.

LEARNING
SCORE 2-3

- Whole-organisation analyses are beginning to be performed on an ad-hoc basis.
- Beginning to focus on what analysis is meaningful and useful.
- Starting to identify what data should be routinely analysed and potentially automated.
- Analysis is mostly descriptive about what happened e.g. summarising the overview, averages, variation, range.
- Comparative trend analysis is starting to be conducted over time (perhaps on an annual basis).
- Exploration and use of filters and cross tabs are being used to delve further into data.
- Data is manually collated in reports using data from different sources.
- Data is manually reworked for presentation in written reports for different internal/external audiences.
- Learning how to create more sophisticated graphs and presentations of the data (though audiences may find them difficult to interpret and understand).

DEVELOPING
SCORE 3-4

- Embedding systems for analysing meaningful and useful data.
- More consistent and regular approach to data reporting and trends analysis on users/needs, activity, outcomes and impact.
- Monitors what's happening in present as well as what's happened in the past. Some forward looking analysis/ forecasts may challenge views of future performance.
- Analysis is more diagnostic about where/why things happened e.g. exploring root causes, clustering patterns, anomalies, discovering differences and trends. Some attempts at A/B testing. Occasional use of predictive analytics in some areas e.g timeseries/forecasting.
- Aware of difference between correlation and causality.
- Routine data analysis is partially automated and partially manually collated from different sources.
- Presentation and communication of data is honed to ensure its meaning is understood.
- Some use of dashboards and/or businesses intelligence systems
- Beginning to explore interactive data visualisation.
- Both static and real-time dynamic reporting conducted for different audiences, some may be available for non-specialists to independently access.

MASTERING
SCORE 4-5

- Analysis extends beyond the organisation to its wider context with cooperative analyses performed with partners/other agencies.
- Analysis extends through descriptive, diagnostic, predictive, and prescriptive techniques.
- May conduct Randomised Control Trials.
- More focus on what will happen in the future with forecasting and predictive models to plan for the future needs of beneficiaries, target services, increase income, and maximise impact.
- Advanced prescriptive analytics support decision-making on how do things in the best way e.g. optimisation of location, staff/volunteer capacity, recommending decisions for effective intervention, experimental design, simulation, artificial intelligence.
- Data from different sources is brought together and analysed in an automated way to provide an organisation-wide analysis.
- Dashboards, business intelligence systems, visualisation tools draw from data warehouse.
- Data visualisation delivers meaningful analysis to different internal and external audiences.
- Non data specialists are able to interactively explore, analyse and report on the organisation's data.

LEADERSHIP

Attitude | Plans | Capability | Investment

UNAWARE
SCORE 0-1

- Not interested in data at all and not seen as a priority.
- No business plan or plans around data.
- Don't use data for decision making, instead rely on gut feeling, experience and what seems to work.
- No data or analytics expertise or understanding among leadership.
- Don't invest in data and analytics.

EMERGING
SCORE 1-2

- Some recognition of the importance of data but don't see the value of it.
- Little awareness of the potential uses of data so not seen as a priority.
- There may be a business plan though perhaps with no defined or measurable organisational goals.
- Limited or very basic data and analytics knowledge and experience among leadership.
- Typically use data about what happened in the recent past and verbal accounts of what's happening for decision-making.
- Very little investment though some may occur under the radar within specific projects/parts of the organisation.
- People are expected to learn 'on the job' and there's no investment in specific roles, tools, or training.

LEARNING
SCORE 2-3

- Know data is important, curious to learn about its potential uses and benefits.
- Leadership occasionally ask questions about the data they are given but are not entirely convinced about its value.
- Data is an interest of the organisation but not a priority.
- There is a business plan with some defined and measurable goals, though data collection/analysis may not align.
- Learning about impact e.g. exploring theory of change.
- Learning through experience, building 'adequate' skills.
- Assessing data and analytics skills, knowledge, and roles across the organisation.
- Might use past and current data for decision making with some simple trend analysis.
- Invest small amounts in some basic/ existing tools and staff training on an ad hoc basis.
- Exploring ideas and needs for the future though no formal plan related to data.

DEVELOPING
SCORE 3-4

- Becoming engaged, supportive, ask the right questions of the data, and active in harnessing its value.
- Data is becoming more of a priority for the organisation as a whole (and considerably so in some projects/ teams).
- Data is becoming aligned to overarching business plan and desired impact.
- Monitors what's happening in the present as well as past trends. Some exploratory forward-looking research and predictions.
- Data champion within senior management.
- Addressing skills gap in leadership as a whole including understanding around impact measurement and management.
- Starting to plan and prioritise data organisation-wide.
- Beginning to commit significant investment in people (job roles), skills, learning, and tools.

MASTERING
SCORE 4-5

- Value, plan and prioritise data and analytics as a vital organisational resource.
- Fully understand how to use data to improve what the organisation does. Data drives questions and the organisation is influenced by what data tells them.
- Viewed as a major organisational priority.
- Overarching business plan with clearly defined goals based on outcomes and differentiated impact, forecasting, and prediction of need.
- Use past, present and forward looking data for business planning and decision making (including forecasting, modelling, prediction and optimisation).
- Range of people with data and analytics expertise in leadership including at Board level.
- Invest substantially in continuously in improving data collection and analysis aligned to business plan.

CULTURE

Team approach | Self-questioning | Openness | Protection

UNAWARE
SCORE 0-1

- Nobody is interested in data.
- Data only accessible to a single person or team, usually junior staff.
- Opinion, observation, passion, and belief are used for decision making.
- Data is seen as a chore with questions/ requirements mostly externally driven.
- Data is never shared internally or externally.
- There are no policies related to data.
- Minimal, if any, security and protection of data on paper, computers, or devices.

EMERGING
SCORE 1-2

- Data is seen as the responsibility of 'someone else', usually in an administrative role.
- Recognition that data should be collected but it is not seen as a 'whole team' activity.
- Data mostly sought out and used to support and evidence what the organisation already believes or knows.
- Organisation's culture doesn't encourage data sharing across teams, though this may occur occasionally verbally or via reports.
- Data never shared externally except for legal/contractual reporting purposes.
- Basic policies for data protection and security may be in place but not monitored or enforced.

LEARNING
SCORE 2-3

- Data is starting to be recognised as important at a more senior level.
- People across the organisation are starting to talk about how they can work together to deliver better data for decision making.
- Beginning to ask challenging questions of the data: why? what's changing? what difference are we making? (mostly about the past.)
- Access to data may be limited by default rather than design.
- People would like to share more but may be constricted by access/ permissions/cultural barriers.
- Some data insights are shared with partners, networks, and in the public domain.
- Data protection and security policies are in place.
- Senior management have some limited understanding of legislation and the organisation's responsibilities.

DEVELOPING
SCORE 3-4

- The whole organisation is starting to use and share data. People from different teams/levels of seniority regularly discuss data and how to act on it.
- Specialist staff in some teams are starting to use data to ask difficult questions, challenging assumptions, practices and impact.
- Concepts of right and wrong (ethics) are being considered, particularly in relation to personal data.
- Data and analysis is becoming more available and accessible to staff though may require some intervention by specialists to provide this.
- External data sharing is done on an aggregated basis and insights are shared including some shared measures and benchmarks.
- Exploring how data could be shared with beneficiaries on an individual basis as part of service delivery.
- Policies and practices are well established to ensure data is safeguarded (e.g. rules on passwords, how data is stored, rights/privileges to access organisational and beneficiary data).
- Risks have been identified though may not have been tested.
- Systems have been created to ensure data about identifiable individuals is deleted when no longer necessary and respond to subject access requests.
- Trustees and senior management keep abreast of current legislation and best practice.

MASTERING
SCORE 4-5

- Data seen as a team effort and critical asset for every part of the organisation.
- Everyone in the organisation is committed to ensuring quality data is available to support decision-making.
- Very comfortable using data to ask difficult and complex questions, to challenge practices and preconceived notions about the past and future.
- Aware of the practical difference between: correlation and causality; attribution and contribution; 'known and unknown' unknowns.
- Explores potential negative impacts of interventions as well as data ethics.
- Internal openness and data sharing is fundamental to the culture, subject to data protection/security.
- Everyone can access analysis they need when they need it.
- Data insights/evidence publicly available.
- Extensive data sharing, with protocols in place with partners, networks, stakeholders to address shared problems and solutions.
- Data may be shared with beneficiaries as part of service/support.
- High levels of confidence about the security of data held in the organisation.
- Systems, automated if possible, in place to delete personal data no longer necessary and respond to subject access requests.
- Risks monitored and tested to improve data security and protection.
- Widespread knowledge/skills sharing.
- Trustees and senior management keep abreast of future changes in legislation and best practice, and regularly check Data Protection compliance.

TOOLS

Storage | Type | Quality | Sharing | Integration | Investment

UNAWARE
SCORE 0-1

- Data is disorganised and unmanaged and stored in a range of places: on desks, in filing cabinets, individual peoples' email inboxes, computers, phones, or other devices.
- Nobody has any formal responsibility for managing any data in their job. Any organising, archiving, updating or data cleaning is performed in an ad hoc way by individuals.
- Data is mostly collected via paper, email, SMS messages. This may not get transferred onto spreadsheets or structured document filing systems.
- Typically uses a spreadsheet/s or basic finance software package, though not in any analytical way.
- No data sharing happens in the organisation.
- No planned investment of time or money in any tools, systems, or infrastructure.

EMERGING
SCORE 1-2

- Data is stored in designated physical locations, in paper and digital formats, in organised ways.
- Some data is managed and controlled by people with clear responsibility for maintenance and cleaning (e.g. administration, finance, HR, managers).
- Data is mostly collected on paper/ phone/in person and then entered into a database or spreadsheet for basic analytical and reporting tasks.
- Unstructured data is disparately held on different individuals' computers and devices.
- Data may be shared mostly by emailing spreadsheets and documents as attachments with duplication, version control, and security issues.
- Some people or teams may use cloud-based document storage to share some data (e.g. OneDrive, Google Drive, Dropbox, Box). Note these may be personally or organisationally owned.
- Tools are limited. May not be up-to-date, don't meet current needs, and are not documented or supported.
- Tools are acquired on a 'needs-must' basis e.g. for a specific purpose/project or when a crisis/system failure occurs.

LEARNING
SCORE 2-3

- Digital data is mostly centrally stored on a (secured, backed-up) cloud-based system or local server with managed access. Some may remain inaccessible on computers, central shared drives or devices.
- Unstructured data is becoming better organised and searchable (e.g. folder structures/file naming conventions).
- Data held in a range of systems all separately managed. Tools likely to include at least one database or CRM system. There are also likely to be multiple spreadsheets in use.
- Tools mostly used operationally rather than analytically. May allow some basic inbuilt analysis and reporting but most often data has to be exported for analysis in another tool. Possible advanced analytical tool used for basic data processing or descriptive statistical analysis.
- Joining data or analysis across teams/ services/functions requires manual exporting and re-stitching.
- Tools likely to be purchased or built as 'one-offs' for specific purposes with limited flexibility for change or improvement. Replacements and upgrades are being researched and costed. Possibly some regular expenditure on cloud-based software tools and ad-hoc hardware replacement.

DEVELOPING
SCORE 3-4

- Data fully centrally stored in secure digital systems with managed access.
- Data held in appropriate databases/ CRMs or other technologies accessible by expert users and some non-experts.
- People are formally responsible for managing the storage, cleaning and maintenance, security, and backup of all data. Where possible this is becoming routine and/or automated.
- Data is collected and automatically stored digitally wherever possible e.g. online forms/apps directly into databases.
- An inventory of tools and systems (including hardware, software, licence, passwords and access) is managed and maintained.
- Most tools are up to date with support available. Work-arounds are understood and replacements planned for poorer tools.
- Advanced tools being used for sophisticated analytics in some parts of the organisation e.g. R, SAS, SPSS, Python.
- Some integration is beginning to occur between systems with automated/ aligned reporting e.g. basic use of business intelligence tools.
- Significant investment is being committed to new tools/integrations across the organisation. Staff time, upgrades, and replacement are being built into annual budgets.

MASTERING
SCORE 4-5

- Data is drawn from a minimal number of sub-databases and systems and held in singly accessible database e.g. data warehouse.
- Capacity is available to store, manage, and analyse increasingly large volumes of data from multiple sources.
- The organisation has a robust infrastructure, integrating tools wherever possible.
- Analytical infrastructure to support effective decision making is a priority.
- All tools are documented and supported (internally and/or externally).
- Tools for delivering 'batch analytics' and real-time 'streamed' data are used.
- Advanced analytics and data science tools present throughout the organisation.
- Tools able to access and utilise internal and external data directly, for both experts and non-experts.
- Automated reporting e.g. through dashboards.
- Self-service analytics available both inside the organisation and to some external partners/stakeholders.
- Ongoing investment of time and money in developing and improving tools, systems and infrastructure to support data and analytics.

SKILLS

Internal capacity | Skills and training | Access to knowledge and expertise

UNAWARE
SCORE 0-1

- No staff commitment beyond basic administrative level and finance roles.
- Little or no internal skills, training, or expertise.
- No access to external knowledge or expertise around data or analytics.

EMERGING
SCORE 1-2

- Responsibility for data collection and control is at administrator and finance officer level.
- Different staff collect, manage and use data within other roles e.g. fundraising, projects.
- No real understanding of the needs and skills required for building data capabilities.
- Data literacy is patchy, mostly low, amongst staff.
- Basic/adequate skills and training in using data for operational and administrative purposes.
- Little or no staff/volunteer awareness or training in data protection and security, though perhaps at least one person has completed a course.
- Occasional support from trustee/ volunteers perhaps mostly relating to database/finance or reporting.

LEARNING
SCORE 2-3

- Beginning to understand needs around data skills and capabilities.
- Dedicated person/team in charge of data, perhaps a data manager or senior administrator. Some skilled data people in other roles, though perhaps with limited capacity to fulfil the task.
- Adequate data analysis/reporting skills with some investment in more advanced skills e.g. Database/CRM administrator.
- Commitment to improving data literacy.
- Exploring up-skilling and recruitment to fill skills gaps.
- In house or externally provided training for using data systems.
- Staff and volunteers have basic data protection and security training though might not be very confident.
- Establishing relationships with external support and advice, mostly around specific tools, systems or projects with some skills development.

DEVELOPING
SCORE 3-4

- Understand different skill sets within data and analytics.
- Dedicated skilled analytics roles established with several people responsible for data in different roles/ teams.
- Possibly a senior person/team bringing organisation-wide data together.
- Increased data literacy/responsibility across the organisation.
- Individuals responsible for data have advanced training and skills and regularly engage in learning to develop and improve systems and embed these across the organisation.
- Staff know how to respond to subject access requests (where individuals request a copy of their data) or changes in preferences on personal data.
- Staff know how to respond to a data breach, potential breach, or near miss.
- Regular use of advanced external expertise and specialist suppliers.

MASTERING
SCORE 4-5

- High levels of staff commitment at senior, specialist, technical, and administrative levels.
- Senior data strategist embedded at heart of leadership decision making.
- Able to independently manage/drive and maximise data analytics to an advanced level.
- All staff trained with ongoing investment in developing data skills with high levels of data literacy across the organisation.
- Becoming experts that other partners/ peers use as a resource.
- Specialist staff regularly update skills and knowledge through training and conferences.
- Active in online learning networks and data and analytics communities of practice exploring new tools and skills.
- Awareness about openness and protection of data are embedded throughout the organisation with regular training, and in-house learning network to develop and maintain best practice.
- Ongoing relationships with a range of trusted suppliers providing advanced support and specialist expertise, periodically reviewed.

Glossary

A/B Testing
A randomized experiment with two variants, A and B. It includes application of statistical hypothesis testing or "two-sample hypothesis testing". A/B testing is a way to compare two versions of a single variable, typically by testing a subject's response to variant A against variant B and determining which variant is more effective.

Anonymised
Removal of identifying details from data before sharing for statistical or other purposes, to preserve privacy.

Artificial Intelligence
An area of computer science that emphasises the creation of intelligent machines that work and react like humans e.g. speech recognition.

Batch Processing
The processing of previously collected jobs in a single batch.

Business Intelligence System
Technology for and practice of the collection, integration, analysis and presentation of business information to support better decision-making.

CRM
Customer Relationship Management system - technology for managing all of an organisation's relationships and interactions with customers and potential customers. Note this could be cloud based.

Cross Tabulations
Data tables that present the results of an entire group of data e.g. survey respondents as well as results from subgroups. To examine relationships within the data that might not be apparent when analysing the whole group e.g. all survey respondents.

Dashboard
An information management tool that visually tracks, analyses and displays key performance indicators and measures to monitor the health of an organisation or department. They can be cloud based and customised to meet the specific needs of a department or organisation.

Data Analysis
The process of cleaning, analysing and summarising data to discover useful information, inform conclusions and support decision making.

Data Analytics
The process of data analysis (compiling and analysing data) and the tools and techniques to do so, to support decision making. Could be basic counts and/or charts; descriptive (about what happened); diagnostic (about why it happened); predictive (about what will happen in future) or prescriptive analytics (about how you can do it in the best way).

Data Assets
A collection of data that holds valuable information or knowledge. This can include databases, CRM systems, spreadsheets, mailing lists, records of transactions or bookings, collections/libraries of documents or images.

Data Collection Methods
Various ways in which data is gathered and measured to answer relevant questions in an accurate and systematic way. Methods vary according to the field of research but some examples are: observations; interviews; questionnaires and surveys; focus groups; ethnographies; oral history; case studies; experiments; randomised control trials.

Data Infrastructure
A digital infrastructure promoting data sharing and consumption. It includes hardware (computers, phones, devices, storage and backup) and software tools which might be cloud-based.

Data Mining
The process of discovering new information from large data sets, involving methods at the intersection of machine learning, statistics and database systems. Also referred to as knowledge discovering in databases or KDD. Often used to find anomalies, patterns and correlations to predict outcomes.

Data Protection
Legal control over access to and use of data held by an organisation.

Data Security
A set of standards and technologies that protect data from intentional or accidental destruction, modification or disclosure.

Data Warehouse
A system that pulls together data from many different sources within an organisation for reporting and analysis. The reports created from complex queries within a data warehouse are used to make business decisions.

Database
A structured collection of data, generally stored and accessed electronically (including cloud-based) that is organised to be easily accessed, managed and updated.

Forecasting
The process of making predictions of the future based on past and present data and most commonly by analysis of trends. To help management cope with the uncertainty of the future, starting with certain assumptions based on experience, knowledge and judgement.

Impact
The overall difference you make to those you're trying to help; the long-term, big picture general change you make as an organisation.

Glossary

Machine Learning
An application of artificial intelligence that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programmes that can access data and use it to learn for themselves, relying on patterns and inferences.

Modelling
The process of creating an abstraction of the real world in order to understand it. A model organises elements of data and standardises how they relate to one another and to properties of the real world entities.

Monitoring Performance
Measuring activities in process to identify operational health of an organisation.

Open Data
Data that anyone can access, use and share. Used to bring about social, economic and environmental benefits.

Optimisation
The process of making something as good or effective as possible.

Outcome
The difference an activity makes to those you're trying to help; the short-term, specific change you make.

Profiles
Customer insights – characteristics of people you connect with e.g. demographic data like age, gender, ethnicity, nationality or geographic data e.g. where people live or what the geographic reach/ need is by area.

Qualitative
Relating to, measuring, or measured by the quality of something rather than its quantity, using non-numerical data. For example, describing an emotion or feeling.

Quantitative
Is information about quantities; that is, information that can be measured and written down with numbers. For example, how many people respond to a survey.

Randomised Controlled Trials
A study where subjects (usually people) are randomly assigned to one of two groups: one (the experimental group) receive the intervention that is being tested, and the other (comparison or control group) receive an alternative (conventional) treatment.

Raw Data
Data that has not been processed for meaningful use, either manually or using computer software.

Relational Data
Usually in the form of a set of tables or database, where data points can be related to one other using a unique identifier, so that the data can be accessed or reassembled in many different ways.

Rich Data
Usually qualitative data which reveals the complexities and richness of what is being studied, particularly in understanding human behaviour. For example, combining customer data with social media and market demographics to understand consumer behaviour.

Service Quality
An assessment of how well a delivered service conforms to the client's expectations, in order to identify problems, improve the service and client satisfaction.

Stream Processing
Processing data in real time, analysing it as soon as it is produced and received. For example, instant analytics of website activity.

Structured data
A standardised format for providing information in a fixed field within a record or file. This includes data contained in relational databases and spreadsheets.

Transactional Data
For example when somebody registers for a service, signs an attendance register, books to attend an event, or perhaps when they purchase or donate.

Unstructured Data
Information that does not have a pre-defined data model or is not organised in a pre-defined manner (e.g. not in a database), and often includes text and multimedia content. Examples include e-mail messages, word processing documents, videos, photos, audio files, presentations, webpages and many other kinds of business documents. Experts estimate that 80 to 90 percent of the data in any organisation is unstructured.

Versatile Data
Data that can be used or adapted for more than one purpose e.g. you might report on one set of beneficiary characteristics over one time-period for a funder and use the same dataset to report on different characteristics for a manager or board of trustees.