The UAE Digital Data Interoperability Framework

Part 2: Digital Data Interoperability Implementation Guide

For Government Entities

Version 3.0

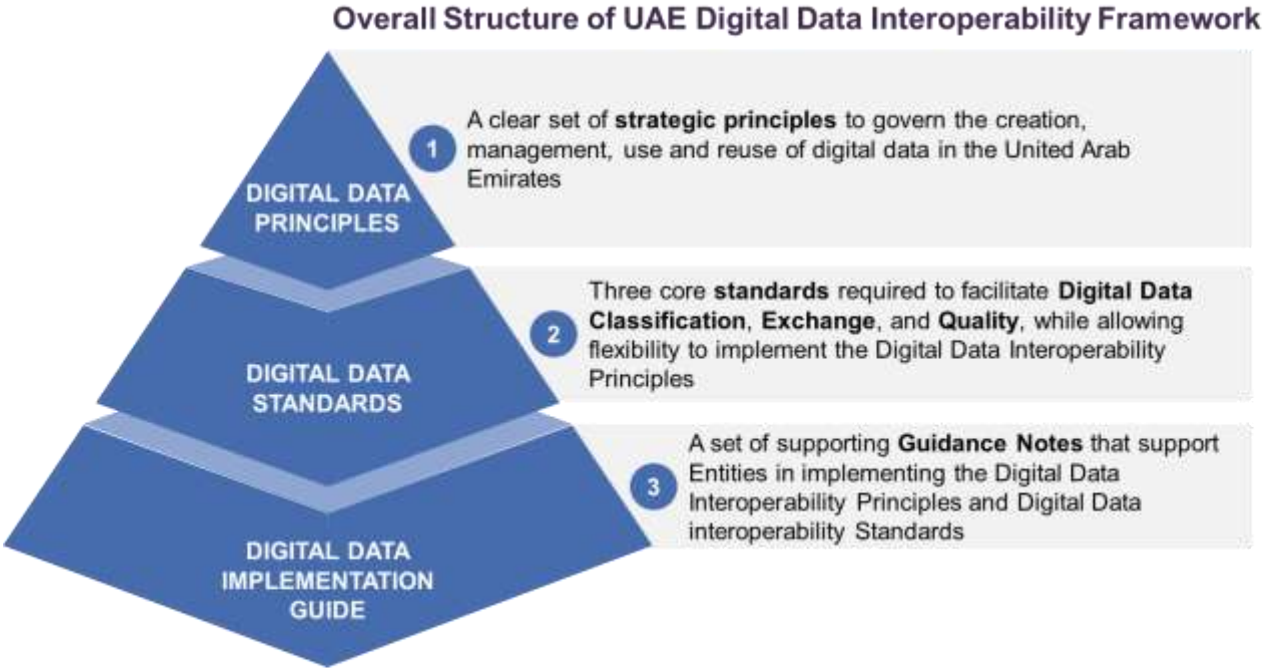
Document Date: 2021

TABLE OF CONTENTS

INTRODUCTION

Context

This Digital Data Interoperability Implementation Guide forms part of the UAE's Digial Data Framework, as illustrated below.


Overall Structure of UAE Digital Data Interoperability Framework
DIGITAL DATA PRINCIPLES
1
A clear set of strategic principles to govern the creation, management, use and reuse of digital data in the United Arab Emirates
DIGITAL DATA STANDARDS
2
Three core standards required to facilitate Digital Data Classification, Exchange, and Quality, while allowing flexibility to implement the Digital Data Interoperability Principles
DIGITAL DATA IMPLEMENTATION GUIDE
3
A set of supporting Guidance Notes that support Entities in implementing the Digital Data Interoperability Principles and Digital Data interoperability Standards

The Digital Data Interoperability Framework outlines a common basis for each UAE Government Entity to develop its own approach for managing digital data, in ways that provide maximum flexibility for the Entity to respond to their own business needs yet which also enable a common approach to digital data classification, exchange of digital data, and digital data quality.

This Digital Data Interoperability Implementation Guide, structured in a series of five Guidance Notes, provides guidance, best practice and recommended processes for Government Entities to follow to ensure they meet the requirements set out in the Principles and Standards of the Framework.

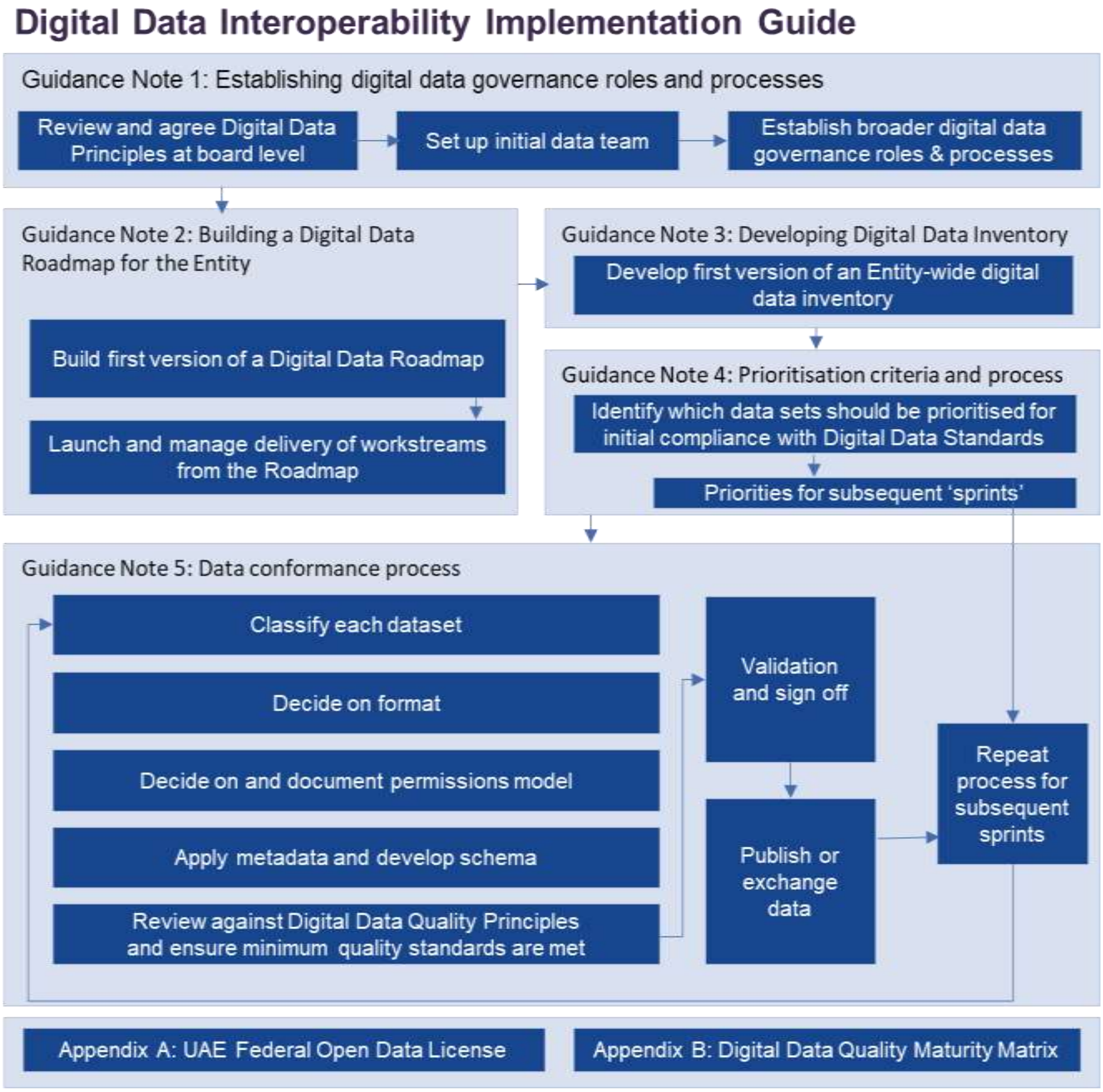
Overview

The diagram on the following page illustrates a typical process that an Entity might go through, supported by the five Guidance Notes, to implement the Digital Data Interoperability Framework in a phased process over time:

1. Establish the Entity's **digital data Interoperability governance roles and processes**
2. Build a **roadmap** to set out the key digital data Interoperability management and change management actions that will need to be taken across the Entity
3. Map out key datasets within an Entity-wide **Digital Data Inventory** (if that does not already exist)
4. **Prioritize** which datasets need action first in terms of applying the core Digital Data Standards

5. Implement the **Digital Data Standards conformance process** through a series of 'sprints' through which digital datasets are aligned with the Standards in a phased and prioritized process over time.

This process is a recommended not mandatory one. An individual Entity may decide to follow a different approach in some areas, provided that this still results in alignment with the UAE Digial Data Interoperability Principles and conformance with the UAE Digital Data Interoperability Standards.


Digital Data Interoperability Implementation Guide
Guidance Note 1: Establishing digital data governance roles and processes
Review and agree Digital Data Principles at board level
Set up initial data team
Establish broader digital data governance roles & processes
Guidance Note 2: Building a Digital Data Roadmap for the Entity
Build first version of a Digital Data Roadmap
Launch and manage delivery of workstreams from the Roadmap
Guidance Note 3: Developing Digital Data Inventory
Develop first version of an Entity-wide digital data inventory
Guidance Note 4: Prioritisation criteria and process
Identify which data sets should be prioritised for initial compliance with Digital Data Standards
Priorities for subsequent 'sprints'
Guidance Note 5: Data conformance process
Classify each dataset
Decide on format
Decide on and document permissions model
Apply metadata and develop schema
Review against Digital Data Quality Principles and ensure minimum quality standards are met
Validation and sign off
Publish or exchange data
Repeat process for subsequent sprints
Appendix A: UAE Federal Open Data License
Appendix B: Digital Data Quality Maturity Matrix

GUIDANCE NOTE 1: ESTABLISHING DIGITAL DATA INTEROPERABILITY GOVERNANCE ROLES AND PROCESSES

Purpose	This Guidance Note provides Entities with guidance on governance roles and processes to support implementation of the UAE Digital Data Interoperability Framework.
When to use	At the outset of each Entity's Digital Data Interoperability program
Responsibility	Entity Management Board

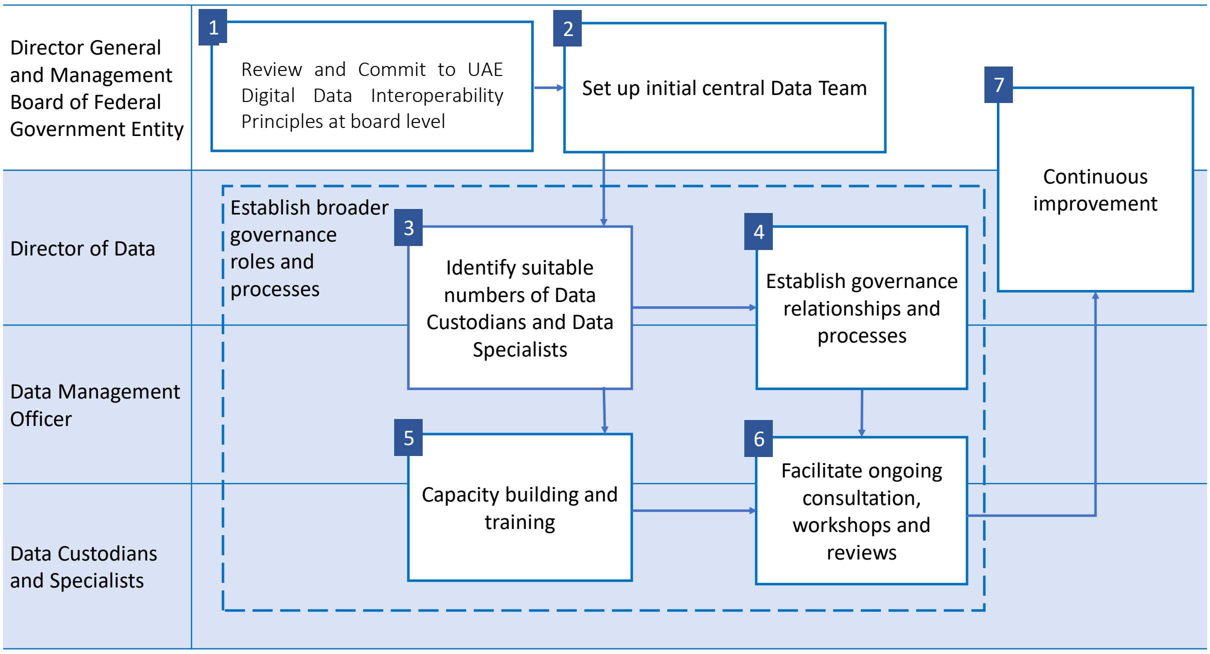
Overview

This Guidance Note recommends good practice on digital data interoperability governance roles and processes, as a guide for Entities then to tailor to their specific needs. It provides guidance in turn on:

- A recommended process for establishing key digital data governance roles, identifying suitable candidates and growing expertise over time
- Sample job descriptions with responsibilities and skills recommended the key roles.

Recommended process for establishing digital data governance

The diagram below summarises the process that Government Entities are recommended to follow when establishing governance for their Digital Data interoperability program. This is followed by further detail on each step.


Director General and Management Board of Federal Government Entity
Director of Data
Data Management Officer
Data Custodians and Specialists
1
Review and Commit to UAE Digital Data Interoperability Principles at board level
2
Set up initial central Data Team
7
Continuous improvement
Establish broader governance roles and processes
3
Identify suitable numbers of Data Custodians and Data Specialists
4
Establish governance relationships and processes
5
Capacity building and training
6
Facilitate ongoing consultation, workshops and reviews

1. Review and commit to the UAE Digital Data Interoperability Principles at board level

The UAE Digital Data Interoperability Framework is rooted in a set of guiding principles, which are summarized below and described in more detail in Part 1 of the Digital Data Interoperability Framework: Principles and Standards. The principles for Digital Data Interoperability that every Entity should embed within its own governance systems and business processes cover the following topics.

Digital Data Data Interoperability principles: summary

1. **Digital Data as an asset:** In order to enable service-oriented government, support evidence-based decision-making, and promote transparency and citizen engagement, Entities should manage all their digital data as a collective national asset, acting as custodians of that digital data on behalf of the United Arab Emirates.
2. **Sharing and re-use of digital data:** In order to enhance the quality of government services, Entities should collaborate closely and efficiently to maximize the sharing and of re-use United Arab Emirates digital data.
3. **Duplication of digital data:** In order to improve customer-centric government services, Entities should collaborate to avoid duplication and inconsistencies in their digital data, employing the concept of a 'single source of truth.'
4. **Open Data publication**: In order to provide greater access to information for all users across the United Arab Emirates, Entities should publish non-personal digital data openly whenever possible.
5. **Privacy, Confidentiality, and Intellectual Property Rights:** In order to secure the broad social benefits of data exchange while respecting the rights of individuals and organizations, Entities should protect the privacy of individuals, the confidentiality of organizations, and the legal rights of intellectual property holders at all times.
6. **Open standards:** In order to empower government service automation through the sharing and re-use of digital data, Entities should utilize open standards to make it easy for others to discover, interoperate with, and consume their digital data as a service. This applies to all data, not just Open Data – because the most efficient way of sharing confidential and sensitive data between Entities is to make it publishable per open standards.
7. **Digital Data quality:** In order to enable the efficient and effective delivery of customer-centric services, improve the accuracy of evidence-based decision-making, and build confidence in both, Entities should manage and improve data quality over time.
8. **Digital Data insights:** In order to improve the effectiveness of services and policy as close to moment of decision and action as possible, Entities should maximize the insights derived from data by facilitating the collection, analysis, and use of real time or near real time data – both their own and that collected by others.
9. **Collaborative governance:** In order to promote greater cross-organizational collaboration and efficiency, Entities should participate in UAE-wide shared services and collaborative governance mechanisms for digital data.
10. **Continuous improvement:** In order to ensure full implementation of the Digital Data Principles and support standardization of processes, Entities should continually adopt improvements and manage change over a sustained period of time, focused on creating an open, data-driven and data-sharing culture.

A key initial starting point should be for the top management team of each Entity to review and sign up to these principles at Board level, and to identify a senior, empowered member of the Board to be accountable for leading the Entity's work to implement these principles.

2. Set up initial Digital Data team

It is the responsibility of each Government Entity to decide how best it will operationalize the principles and standards of the UAE Digital Data Interoperability Framework within the Entity, and this includes the choice of data governance roles. The right approach to staffing will vary from Entity to Entity, dependent on current levels of maturity of data management within the Entity, on the scale of the Entity's operations and on how important digital data is to the functions of the Entity.

However, it is recommended that each Entity establish at least the following roles or their equivalent:

- **Director of Data (DD):** a senior and empowered staff member, who will lead the Entity's Data program, champion and promote data management processes and effective data publication and exchange and ensure strategic goals are realised. Ideally, the Director of Data should be a member of the Entity's management board; as a minimum, they should be a senior and empowered individual with an ability to rapidly escalate key risks and issues for resolution at the highest levels in the Entity. For smaller Entities this role might be performed on a part-time basis, for example by an existing member of staff but with additional assigned responsibilities.
- **Data Management Officer (DMO):** to report and deputy for the DD and lead on the operational work managing and coordinating required change management, processes and coordination to ensure conformance with Digital Data Interoperability Framework standards. This is an important and full time role and requires the person responsible to spend a significant part of their time on digital data interoperability standard conformance.
- **A suitable number of Data Custodians and Data Specialists: to** act as business and technical owners of key datasets and digital data sources within the Entity. They will understand the contents and business value of the data, how it was collected and processes and the accuracy and quality of the digital data. These could be existing data owners and IT staff with new responsibilities.

Given that each Entity is different and will have varying existing data infrastructure and processes – the specific setup must be at the discretion and judgement of the Entity itself. The Entity can choose to create new positions for each of these or add the titles and responsibilities to existing roles.

The important thing is that the business functions and responsibilities of these roles are carried out by suitably informed and skilled staff.

We recommend that the appointment of a Director of Data and Data Management Officer are the first positions filled. The people who fill these roles require a background in data management, government operations and digital technology – they will lead the Entity's digital data interoperability program and ensure the Entity'digital data is published and exchanged in line with the requirements of the Digital Data Interoperability Framework. A background or knowledge of open data is an advantage but not essential. Detailed job descriptions are given at the end of this Guidance Note.

3. Identify suitable numbers of Data Custodians and Specialists

The DD and DMO should identify or hire Data Custodians and Data Specialists for each core data source and/or data business unit. The DMO should set up a reporting and communication system for ensuring that work and processes for meeting the Digital Data Interoperability Framework requirements can be met across the business units by suitable and qualified personnel (individuals who know about parts of the Entity's data systems). Depending on the size of the Entity, there could be many Custodians that report to a specific department or business unit Data Custodian who in turn reports to the Data Management Officer.

4. Establish governance relationships and processes

Entities should establish clear governance processes in which all relevant people and teams are clear on who is responsible, accountable, informed and consulted on all work the Entity carries out to achieve Digital Data Interoperability objectives. This should be agreed between the Director of Data and Data Management Officer.

Entities may wish to consider developing a formal RACI matrix to summarise these processes, setting out for each one:

- Who is **R**esponsible for managing the process
- Who is **A**ccountable for the business results delivered by the process
- Who should be **C**onsulted on the process
- Who should be kept **I**nformed about it.

5. Capacity Building

Custodians and Specialists should receive training and guidance to help them understand their role and responsibilities. They should read the Standards and Implementation Guide of the Digital Data Interoperability Framework. We also recommend discussing how best to implement the Digital Data Interoperability Principles with key representatives of each of the Entity's business functions. Facilitating such internal discussion is a key role for the Data Management Officer.

6. Organise and facilitate regular consultations, workshops and reviews

In order to facilitate learning and change management it may be useful to have regular workshops for digital data-related roles in the Entity to report on progress and highlight learnings, challenges and tactics so these can be acted upon and shared across the Entity. The DMO could organize consultations on the Implementation Guide so that the Data Custodian teams could decide which parts to use and which to amend and how the various processes should be adopted across the Entity consistently.

7. Continuous improvement

As the Entity's digital data maturity and business processes develop and as the Custodians, Specialists and Data Management Officer run through multiple sprints of formatting and cataloguing data to ensure their conformance with the Digital Data Interoperability Standards, the Entity should continue to refine the roles and responsibilities of its data staff. The Director of Data should keep the effectiveness of governance arrangements under review, agreeing changes with the Entity's Management Board as required.

Role Descriptions

The tables below give more detail on the key recommended data management roles for a typical Government Entity, setting out in turn:

- An overview of the role profile
- The key responsibilities that the role needs to manage
- The skills and competencies needed to do this effectively

Title	**Director of Data - Reports to Director General of Entity**
Role profile	The Director of Data is the senior champion and leader for digital data within the Entity. They are responsible both for communicating the social, economic and business benefits of open data and data exchange as well as ensuring the Entity's conformance with the Digital DataFramework standards. They should be responsible for the development of the data strategy and policies that are relevant to the government entity and supervise the execution and the implementation of initiatives that contribute to the management of data efficiently and work on the data exchange between government entities in a safe, secure and reliable way to develop methods for service delivery and utilization and make it available as open data to induce innovation. The role should be fulfilled by a senior employee, with the necessary influence and authority within the Entity to be effective in the role. This person will also be outward facing, collaborating and communicating with external stakeholders. The Director of Data will be the senior point of contact between the Entity and the Entity Overseeing Digital Data, responsible for communications, coordination and escalation.
Key Responsibilities	**Leadership** • Overseeing the development of the Entity's implementation plan and roadmap for meeting the Digital Data Interoperability Standard requirements, and directing delivery of that plan • Leading a program of cultural change within the Entity aimed at embedding the Digital Data Interoperability Principles within the Entity, promoting the new ways of working and championing the benefits of higher data quality and data exchange • Performing public outreach and presentations to increase the strategic use of the Entity's datasets • Leading the Entity's open data initiative • Leading the Entity's work on benefit realization: ensuring that the benefits of open and shared digital data are maximized, through high levels of adoption and utilization of the data to improve services and decision-making. **Governance** • Putting in place the necessary roles (with the appropriate skills) within their Entity, as outlined in this document • Providing regular reports and conformance information as requested to the Entity Overseeing Digital Data • Improve collection, usage and exchange of digital data. **Conformance** • Ensuring that the Entity's digital data is consistent with the applicable laws and policies of digital data in the UAE, and meets the mandatory requirements of the Digital Data Interoperability Framework

	• Reviewing classified and catalogued digital datasets to check they are conformant with the Digital Data Interoperability Framework standards and approving them for publication and exchange with other Entities • Reviewing data quality reports and statistics • Ensuring timely and effective response to queries and feedback from the public in relation to the Entity's Open Digital Datasets • Investigating any complaints made in relation to the Entity's Digital Datasets by the Entity Overseeing Digital Data, a digital data user or a member of the public.
Skills and competencies	• Ability to communicate effectively, including ability to explain technical content to non-technical audiences • Ability to collaborate and network with subject matter experts, organizations, and individuals to provide effective enterprise data management • Experience in technology and digital data, developing digital data strategies and overseeing the improvement of data quality and data exchange • Ability to formulate and set goals • Proven ability to lead cross-functional teams at all organizational levels in dealing with complex issues

Title	**Data Management Officer - Reports to Director of Data**
Role Profile	This role is the delivery and operational lead for the Entity's data program, reporting to the Director of Data. They will need to deliver and manage much of the work to ensure the Entity is conformant with Digital Data Interoperability Framework standards. Their role involves ensuring the right staff are selected as Data Custodians and Data Specialists and directing, supporting and reviewing their work on Inventories, Prioritization, Classification and Digital Data Conformance. They are responsible for ensuring the readiness, reliability and security of the Entity's digital data and accuracy of the metadata, its availability, accessibility and use in a timely manner to support the operations of the Entity and guide the analysis of digital data for decision-making.
Key Responsibilities	**Leadership** • Effective implementation and oversight of all the digital data management initiatives and processes needed to deliver on the requirements of the Digital Data Interoperability Framework in the Entity • Providing support and advice to the Data Custodians within their business unit when classifying data within the scope of their management and assessment of risks associated with disclosure or exchange • Supporting the Director of Data to ensure that the benefits of open and shared data are maximised and participating in the development of the Entity's data roadmap • Provide mentorship and professional development of staff **Governance** • Determining priorities with respect to Open Data publication or Shared Duigital Data exchange. • Co-ordinating the work of the business unit as it prepares its open and shared digital data for publication and exchange

	• Support resolution of any issues and problems in digital data or conformance to the digital data interoperability standards **Conformance** • Preparing regular reports and conformance information as requested to the DD and Entity Overseeing Digital Data • Administering the process of inventorying, prioritising, cataloguing and classifying the Entity’s digital data • Cascading knowledge about classification principles and procedures to required roles within their Entity • Responsible for consistency and quality of digital data is fit-for-purpose across Entity • Reduce digital data duplication across the Entity
Skills and competencies	• Ability to coordinate and manage the work of a large and diverse team • Ability to collaborate with senior management of Business Units, functional organizations and individuals to provide effective enterprise digital data management • Ability to provide data object domain insight and direction to Data Custodians • Experience in data management and data processes in all its aspects • Displays understanding of all business processes dependent on data in their object domain • Proven ability to create presentations and effectively present to management

Title	**Data Custodian - Reports to Data Management Officer or Senior Data Custodian**
Role Profile	There should be a Data Custodian per dataset or database or data-generating function within the Entity. This person needs to understand the value and risks associated with their data so that they can effectively prioritise, classify and catalogue it. They will be responsible for determining whether the data should be Open or Shared and setting out the access rules. Generally, this is a role within a business unit (data generator) who has a business responsibility (not a technical or a legal one) for ensuring that the data is used effectively to meet both the business needs of the department and the wider goals of the Digital Data Interoperability program. The Data Custodian does not necessarily need to be the creator or the primary user of the dataset but should understand its value to the Entity.
Key Responsibilities	**Leadership** • The management of the assigned data including inventorying, prioritizing and describing datasets • Recommending changes to data management policy and procedures, digital data quality and the implementation of UAE Digital Data Interoperability Standards. • Understanding and promotion of the value of digital data for Entity-wide purposes and facilitation of digital data sharing and integration. **Governance** • Ensuring the quality, completeness and up-to-date of their digital data • Working in collaboration with the Data Management Officer to determine priorities and associated risks of making data accessible by third parties • Engaging with the external developer community to determine how enhancements to the data set could facilitate greater levels of re-use.

	Conformance • The collection and updating of the assigned digital data • Management of any third party use of the digital data in accordance with UAE policies and processes • Advising and reporting on data management issues • Suggesting the terms and conditions upon which Shared Digital Data should be made available.
Skills and competencies	• Collaboration skills within the business and Data Management Officer to help provide effective solutions to data issues and problems • Displays mastery of the portions of business processes executed by their business area • Proficiency with MS Office, basic and some more advanced data analysis and process control methods/techniques • Understands the fundamentals of digital data bases and data structures (tables, hierarchical structures, flat files, etc.) • Demonstrates understanding of all the functions performed by their business area • Displays familiarity with systems used within/by their business area

Title	**Data Specialist - Reports to Data Management Officer or Senior Data Custodian**
Role Profile	This is a role with technical responsibility over digital data, and in particular with responsibility for preparing digital data for publication as open data or for exchange as shared digital data. Probably based within IT or database administrator teams, Data Specialists will need to facilitate between the Information Technology, Information Security and business teams and ensure the digital data they are responsible meets the format, schema and quality requirements in the Digital Data Interoperability Framework standards. They should also be able to provide support to the Entity for cross-business definition of data standards, rules, and hierarchy and refinement of data processes in accordance with defined standards.
Key Responsibilities	**Leadership** • Assists with the resolution of data integration issues as requested by the Data Custodian • Assists the Data Custodian in the definition of data requirements and data rules • Supports projects and initiatives in development and refinement of data processes and metrics in accordance as requested by Data Custodian **Governance** • Supports definition, approval and execution of the Digital Data Quality program • Understands the Information Technology Landscape and has the ability to identify what digital data is stored in what systems • Supports efforts to provide digital data awareness education for senior and upper management • Works with the Data Custodian in the identification of root causes of major data problems and supports the implementation of sustainable solutions

	• Resolves routine data problems **Conformance** • Assists the Business Data Custodian with data problem resolution when requested • Reviews digital data deletion and archiving requests for digital data in their span of responsibility and forwards to approver with appropriate recommendations
Skills and competencies	• Displays mastery of fundamentals of problem solving and basic digital data quality analysis • Displays understanding of the portions of business processes executed by their business area • Strong understanding of the Systems Development Life Cycle and methodologies, and familiarity with process improvement frameworks. • Proficiency with MS Office, basic data analysis and process control methods/techniques • Proven ability to work well and contribute to cross-functional teams • Proven ability to present to peers and supervisors • Displays familiarity with systems used within/by their Entity • Displays familiarity with functions performed by their business area

GUIDANCE NOTE 2: BUILDING A DIGITAL DATA INTEROPERABILITY ROADMAP

Purpose	This Guidance Note provides a recommended process, templates and supporting guidance for each Government Entity to build its own Roadmap for implementing the UAE Digital Data Interoperability Framework. An Entity-level Roadmap that follows this guidance will, if effectively managed, ensure that the Entity: • Achieves significant business benefits from shared and open data • Converges its data management practices over time to conform with the UAE Digital Data Interoperability Principles and Digital Data Interoperability Standards
When to use	At the start of each Entity's Digital Data Interoperability program.
Responsibility	Director of Data, with close involvement and Roadmap sign-off by the Entity's Management Board.

Overview

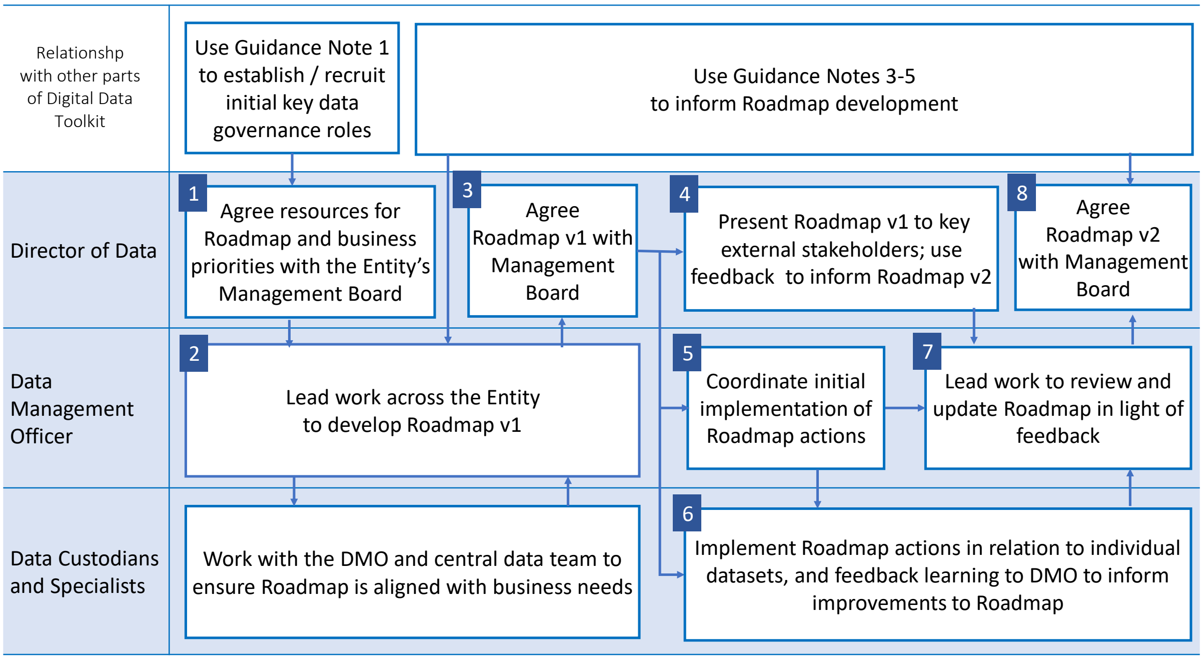
A single, undifferentiated approach for implementing Digital Data Interoperability across all Government Entities will not work. There are a number of factors that will influence the content of an Entity's Roadmap, such as the type and size of Entity, the complexity of its delivery ecosystem, and the Entity's level of maturity in current data sharing practices. The advice in this Guidance Note is therefore not mandatory and Government Entities should tailor it to their own business requirements.

The Guidance Note recommends good practices on:

- The **process** that Entities should follow in developing a Digital Data Interoperability Roadmap
- The **scope and content** of an effective Digital Data Interoperability Roadmap, with a recommended template setting out:
  - The **purpose** of that section
  - **Issues** to address
  - **Actions** that the Entity should take to inform development and documentation of that Section
  - **Resources** that are available to support the Entity in developing this part of the Roadmap.

Recommended process for developing a Digital Data Interoperability Roadmap

The diagram below summarizes the process that the Entity should follow for developing its Roadmap, illustrating which of the key data governance roles will normally have lead responsibility for each step of the process.


Relationshp with other parts of Digital Data Toolkit
Use Guidance Note 1 to establish / recruit initial key data governance roles
Use Guidance Notes 3-5 to inform Roadmap development
Director of Data
1
Agree resources for Roadmap and business priorities with the Entity's Management Board
3
Agree Roadmap v1 with Management Board
4
Present Roadmap v1 to key external stakeholders; use feedback to inform Roadmap v2
8
Agree Roadmap v2 with Management Board
Data Management Officer
2
Lead work across the Entity to develop Roadmap v1
5
Coordinate initial implementation of Roadmap actions
7
Lead work to review and update Roadmap in light of feedback
Data Custodians and Specialists
Work with the DMO and central data team to ensure Roadmap is aligned with business needs
6
Implement Roadmap actions in relation to individual datasets, and feedback learning to DMO to inform improvements to Roadmap

Throughout this process, you should seek to take an approach which is:

- **Iterative and collaborative:** You should not develop the Roadmap in isolation or see this as a one-off exercise. Once an initial Roadmap has been developed, you will want to:
  - Share it with other key stakeholders: the Entity Overseeing Digital Data, other Entities addressing similar customer groups, other Entities using your data or supplying you with data and so on
  - Improve it in the light of implementation experience within the Entity.
- The process diagram above summarizes this collaborative process in terms of producing a first version of the Roadmap and then a second. In practice, the Entity will want to keep the Roadmap updated on an ongoing basis.
- **User-focused:** the Roadmap should be user-centric, i.e. it should identify and address the needs of key internal and external digital data users and should allow for regular engagement with them
- **Practical:** the Roadmap should be achievable within the timeframe, supported with adequate resources to deliver it and appropriate project management disciplines to ensure high quality and timely delivery
- **Phased:** the Roadmap should be developed to be delivered in a phased manner, ensuring that work is closely informed by **Guidance Note 4: Prioritization Criteria and process**.

Scope and content of an effective Digital Data Interoperability Roadmap

Overview

The table below sets out the key elements that should be covered in each Entity's Roadmap. It is not obligatory to follow this precise structure, and you may wish to add other elements to the Roadmap. However, it is important to ensure that – whatever structure is used - all the elements shown below are covered.

Recommended structure	**Overview of each section**
1. **Objectives**	• Sets out the scope and purpose of the Roadmap, and describes the Entity's vision for how it will manage its data in future to align with UAE Digital Data Interoperability Principles.
2. **Gap analysis**	• Highlights the key areas in which current ways of working within the **Entity** are not currently aligned with the future vision
3. **Governance**	• Describes key governance roles and processes for managing implementation of the Roadmap.
4. **Delivery plan**	• Describes work streams in the Roadmap, mapping out key milestones, deliverables, and dependencies.
5. **Risks**	• Sets out the key risks associated with the Roadmap, their likely impact and the proposed mitigation strategies.
6. **Impact measurement**	• Sets out the key benefits that the Entity seeks to deliver through implementation of its Digital Data Interoperability Roadmap: – How will success be measured and when? – What will success look like? – How will learnings be incorporated?

The more detailed tables below provide guidance on the scope and purpose of each of these six recommended sections of the Digital Data Interoperability Roadmap, issues to address within it, the actions you should take to inform its development, and the resources that are available to help you.

Section 1: Objectives

Scope and purpose of the section

In order for an Entity to comply with the UAE Digital Data Interoperability Framework, it needs to establish a new operating model for its digital data management, ensuring that all data is managed as a cross-government asset using open standards. This means each Entity should consider and fully plan the changes that will need to be made internally to drive the transformations that are needed.

This introductory section should set the scene for each Entity, so that they can describe:

- The changes that will be delivered, in terms that resonate with the Entity's own internal and external stakeholders.
- The benefits that the Entity will achieve through delivery of the plan, and ultimately how these will contribute towards the wider goals of the UAE Digital Government strategy.

Issues to address

When developing this section of the Roadmap, you will need to consider:

- Why does the plan exist and what is it trying to achieve?
- What are the required behaviours/actions that need to be taken?
- Which areas of the Entity's activities does this Roadmap cover?[1]
- What is the Entity's future vision for how it will manage its digital data?

Actions you should take to inform development of this section

All team members involved in developing the Digital Data Interoperability Roadmap should have a clear understanding of the Digital Data Interoperability Principles, and of the standards and guidance that exists to support them. So before they embark upon developing the Entity-level Roadmap they should read all the UAE Digital Data Interoperability Framework documents.

And it is vital that the Management Board of the Entity understands the key principles, the degree of change that is involved, and understands and commits to the Roadmap process. While it will not be necessary for all top managers in the Entity to read the full Digital Data Interoperability Framework, it is important that they review the Digital Data Interoperability Principles and give a steer on how they see these best being implemented within the organisation.

It may be worthwhile for your immediate digital data management team to draft this section, review it with the Management Board, and then to circulate it to the wider stakeholder community to validate whether this resonates with them.

Resources to help you

There are a number of key documents that will help the Entity to prepare this section of the Roadmap:

- **UAE Digital Data Interoperability Framework: overview and principles**: this sets out the key purpose and principles of the UAE Digital Data Interoperability initiative, and provides an easy to assimilate guide to the business changes that are required of Government Entities
- **The UAE Digital Data Interoperability Standards:** this sets out the minimum mandatory standards that Government Entities should deliver with their data
- **Guidance Notes 3 – 5 of this Digital Data Interoperability Implementation Guide** give detailed 'how to' guides on implementation of the standards. Use of this guidance is not mandatory, but it provides a good starting point for thinking through the Entity's own approach.

Section 2: Gap analysis

Scope and purpose of the section

To identify the key areas in which current ways of working within the Entity are not currently aligned with the future vision.

Issues to address

You should make clear the scale of the changes that will be required – informed by real evidence of gaps and challenges at two levels:

- **The organisational level:** to what extent does the Entity have the governance, culture, processes and infrastructure needed to manage and reap benefits from digital data?
- **The dataset level:** to what extent are datasets in the Entity currently already aligned with the best practices set out in the UAE Digital Data Interoperability Standards?

[1] In general, the Roadmap should cover the whole organization. But there may be cases where it is sensible either to exclude some elements of the organization, or to cover activities that fall outside the organizational boundary of the Entity, but which nevertheless are most sensibly covered within the Entity's Roadmap.

Actions you should take to inform development of this section

You should:

- Undertake a self-assessment of organisational digital data capabilities
- Identify a sample of key datasets used by the Entity, and audit them against the seven Digital Data Quality principles described in **[DQ1] Data Quality Interoperability Principles.**

Resources to help you

Appendix B provides a Digital Data Quality Maturity Matrix for auditing quality of a dataset against 5 levels of maturity – of which Level 3 represents full conformance with all mandatory requirements of the Digital Data Interoperability Standards. An organisational self-assessment tool is also available.

Section 3: Governance

Scope and purpose of the section

This section should cover:

- **Governance model:** describing functions, roles and accountabilities – both for ensuring successful delivery of activities and milestones in the Roadmap and also for realisation of the targeted benefits – and the processes within which these will operate.
- **Resourcing:** staff, financial and other resources that will be deployed in delivering the Roadmap.

Issues to address

When developing this section of the Roadmap, you will need to consider a number of points:

- Do we have all roles filled across our digital data team?
- If not, how are we proposing to plug the gaps in the short term?
- Do these resources have all the skills and knowledge that they need to carry out their roles successfully?
- What is our own RACI (Responsible, Accountable, Consulted, and Informed) for all digital data management processes?
- How will we manage our Entity's involvement in the wider governance for UAE Digital Data?

Actions you should take to inform development of this section

You will need to have a clear data team structure in place, who are appropriately linked to the Entity-level governance mechanisms.

You will also need to involve the 'data practitioners' within your Entity to map out workflow and processes for data conformance, based on the guidance given in the Digital Data Interoperability Implementation Guide for Government Entities.

Resources to help you

Guidance Note 1: Governance roles and processes gives advice on how to establish governance to develop and deliver a Digital Data Interoperability Roadmap for a Government Entity.

Section 4: Delivery plan

Scope and purpose of the section

This section should set out:

- The work streams and activities that will be taken forward by the Entity to deliver digital data
- The deliverables that will result from this work, with milestones
- A Gantt plan, providing a graphical illustration of the work streams, activities, and tasks that the Entity will manage, with key milestones mapped to a delivery timeline, showing how each contribute to one or more deliverables.
- Key dependencies that the Entity needs to manage, including with other Government Entities and with private-sector delivery partners.

Issues to address

When developing this section of the Roadmap, you will need to consider a number of points:

- Are we clear on what deliverables we need to deliver?
- Will our resources change over time, and how will this impact our approach?
- Are we starting from scratch, or do we need to take into consideration some existing progress on Digital Data Interoperability in some areas of the Entity?
- Do the suggested work streams (see below) cover all the areas that are relevant to our Entity?
- As we map out specific tasks under each work stream,
  - Does each of the tasks have a clear owner?
  - Do we have enough resources to be able to deliver on this plan?
  - Are the timescales for each task/activity realistic and achievable?
  - Are we clear on which milestones are cascaded to us by the overall program, and therefore the dates must remain static?
  - Is the sequencing logical and aligned with the data preparation and management processes?
  - Are there any other Entity-specific limitations that may prevent us from being successful in delivering this plan?

A high-level view of the work streams in a typical Delivery Plan is shown on the following page. This is divided into four main phases:

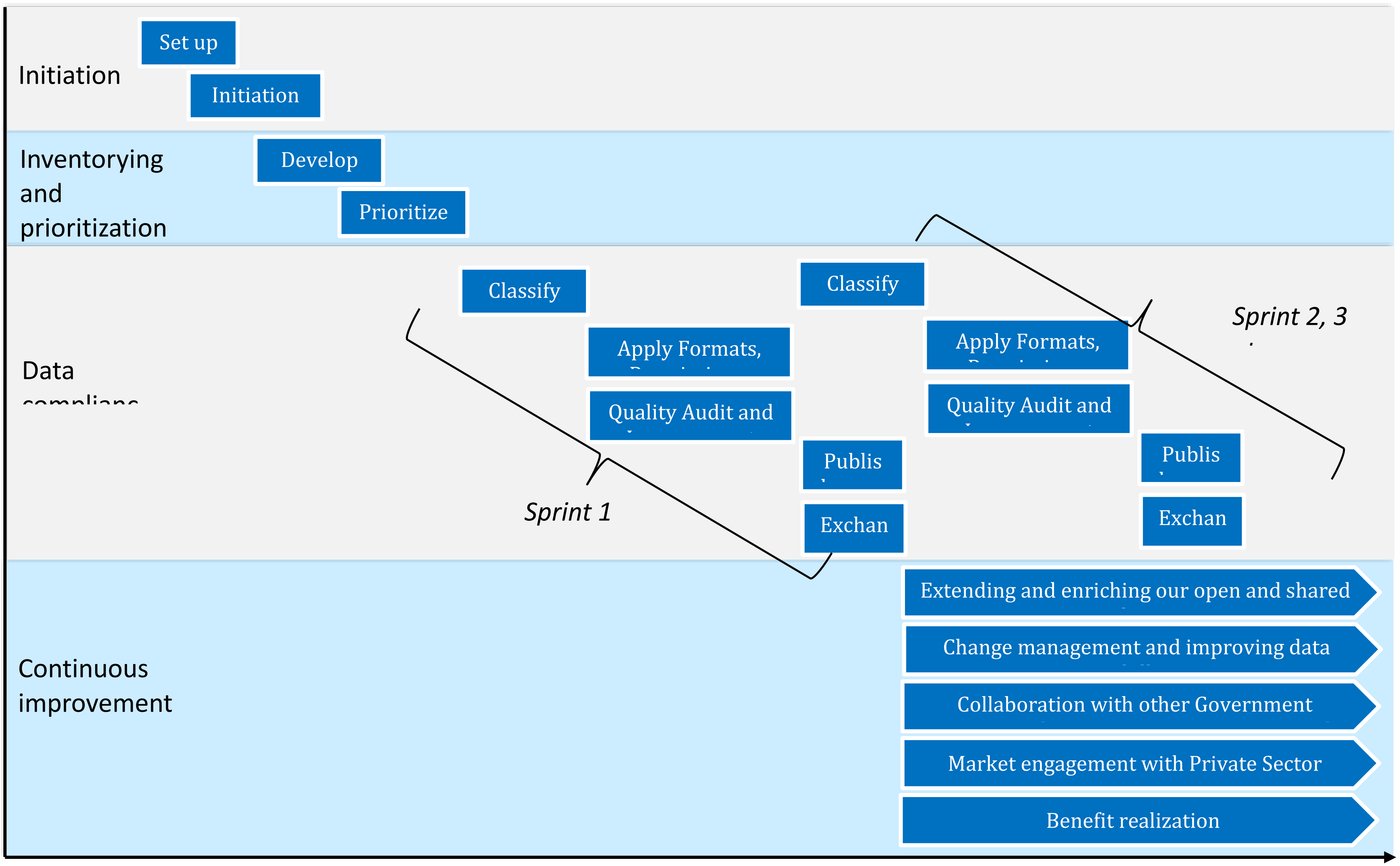
1. **Initiation:** when the initial planning is undertaken and governance systems established

2. **Inventorying and prioritization:** when the Entity pulls together an inventory of the key digital data assets it manages, and prioritises which should be addressed first

3. **Digital Data conformance:** when prioritized datasets are taken through a systematic process to ensure they meet the mandatory requirements of the UAE Digital Data Interporpibility Standards for Classification, Quality and Exchange, in a series of ‘digital data sprints’.

4. **Continuous improvement:** when the Entity drives forward longer-term improvements (beyond the mandatory minimum in the UAE Digital Data Standards) – improvements both to digital data quality and to the level of re-use of the Entity’s digital data by digital data users across the public and private sectors.

Top-level Digital Data Delivery Plan



Initiation

Set up

Initiation

Inventorying and prioritization

Develop

Prioritize

Data complianc

Classify

Apply Formats,

Quality Audit and

Publis

Exchan

Sprint 1

Classify

Apply Formats,

Quality Audit and

Publis

Exchan

Sprint 2, 3

Continuous improvement

Extending and enriching our open and shared

Change management and improving data

Collaboration with other Government

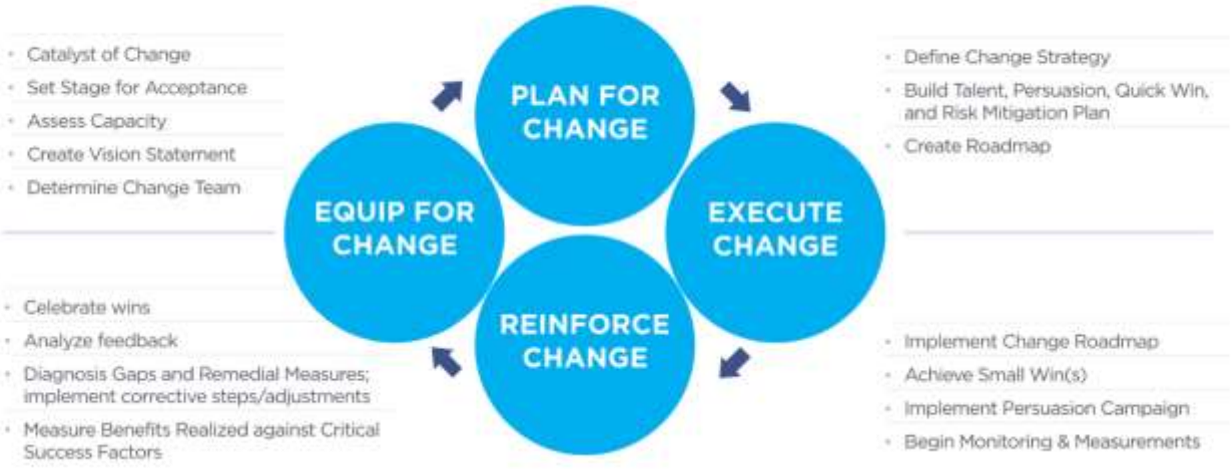
Market engagement with Private Sector

Benefit realization

Detailed guidance on the tasks you will need to manage in Phase 1, 2 and 3 are set out in the other Guidance Notes of this Digital Data Interoperability Implementation Guide. Implementation of these should get the Entity to the point where all its data is compliant with the mandatory requirements of the UAE Digital Data Interoperability Standards. But to get the fullest benefits from digital data, Entities should also look at setting themselves stretch targets for continuous improvement beyond that level. Key longer term work streams that you should take forward during this continuous improvement phase include:

- ***Extending and enriching your open and shared data.*** Actions in this phase should include:
  - Driving forward the quality of key datasets – and in particular Primary Registries – beyond the minimum requirements of the UAE Digital Data Interoperability Standards
  - Engaging with existing and potential users of the Entity's data to identify any barriers to greater levels of re-use
  - Progressively modernising legacy IT and information architectures to facilitate interoperability and data sharing
  - Tackling any contractual barriers that prevent key data being shared or published as open data (for example where ownership of key data currently lies with a private-sector contractor)
  - Establishing systems and processes to embed the use of UAE Digital Data Interoperability Standards into all future procurements and technology implementations

- ***Change management and improving data skills.*** Many aspects of Digital Data Interoperability implementation will require new skills within the Entity, at both a business management level and a technical level. Initial training will be facilitated by the Entity Overseeing Digital Data, but you will also want to lead a program of cultural change within the Entity, aimed at embedding understanding of open and shared digital data, promoting the new ways of working that it opens up, and championing the benefits. The diagram below highlights key phases and issues to consider when doing so.



- ***Collaboration with other Government Entities to drive service improvement with Shared Digital Data.*** Actions in this phase should include:
  - Re-configuring your services and business processes to take advantage of relevant Primary Registries being developed by other Government Entities
  - Using 'joined-up' data with other Entities that provide services to the same customer groups to facilitate more user-centric, integrated service delivery to those customers
  - Providing more responsive and personalised services, informed by richer and more real-time digital data
  - Using Open Data to drive a culture of service co-creation, in which the Entity's customers play a significant role in driving service enhancements.

- ***Market engagement with the Private Sector.*** Actions in this phase should include:
  - Pro-active engagement with the developer community to seek views on the quality of the digital data published already, seek priorities for future releases, and facilitate exploration and experimentation with your Open Data
  - Encouraging small businesses, community organisations and individual citizens to use your Open Data to create new sorts of services
  - Engaging with Private-sector Entities to foster the development of innovative services and new business models based on sharing and integration of private-sector and public-sector data
  - Promoting demonstrator projects to showcase and champion the benefits being achieved by early adopters of open data in the private sector
  - Building public trust in the privacy and security of personal data held by the Entity.

- ***Benefit realization.*** Actions in this phase should include the key tasks and milestones flowing from the approach to Benefit Realization set out in Section 6 of the Entity's Digital Data Interoperability Roadmap. (See the Table on Section 6 below for further advice).

Actions you should take to inform development of this section

In order to develop a realistic and achievable plan, you will need to:

- Involve the 'data practitioners' within your Entity. They will need to have a clear understanding of the processes that will need to be followed, using the training materials and policy products for preparation and cataloguing.
- Undertake sample Digital Data Quality Audits using the Data Quality Maturity Matrix provided in Appendix B, to get an early sense of the current state of data quality across the Entity and key areas where improvement will commonly be needed.
- Engage with current users of any digital data that you currently publish as open data or share with other Entities, to help understand user priorities.

Resources to help you

Guidance Note 1: Digital Data governance roles and processes gives advice on steps to take during the Project Initiation phase of the Roadmap.

Guidance Note 3: Developing a Digital Data Inventory and **Guidance Note 4: Prioritization criteria and process** give guidance on the steps to take during the Inventorization and Prioritization phase of the Roadmap.

Guidance Note 5: Digital Data Conformance Process gives guidance on the steps to take during the Digital Data Conformance phase of the Roadmap.

Section 5: Risk

Scope and purpose of the section

This section should do two things:

1. set out the current list of key risks associated with delivering the Roadmap, including their likely impact and proposed mitigation strategies
2. Set out the process that the Entity will follow to raise and manage delivery risks.

Issues to address

When developing this section of the Roadmap, you will need to consider:

- **Ownership:** who is the most suitable owner for each risk, responsible for driving forward the agreed mitigating action?

- **Governance:** what is the most appropriate escalation route for any high impact risks that are not being managed satisfactorily?
- **Coordination:** who within our Entity will coordinate the whole risk management process, to regularly review the register is being managed effectively?
- **Tools:** how will we manage reviewing and updating our risk register, so that the latest version is visible to all relevant team members and they can contribute updates quickly and easily?

Actions you should take to inform development of this section

We recommend that you review the resources and guidance on risk described below in 'Resources to help you' with all members of the data management team. Once you have a draft risk register in place, this should be reviewed with the Entity Management Board – and also with the Entity Overseeing Digital Data in order to facilitate joint working across government to mitigate risks that are common to multiple Entities.

Resources to help you

The global standard on ICT-enabled, digital data-driven service transformation ('The Transformation Government Framework', or TGF[2]) identifies nine Critical Success Factors. These nine CSFs provide an evidence-based framework, developed following international best practice research and consultation, into the key reasons why ICT-enabled change programs such as UAE Digital Data Interoperability are mostly likely to fail. We recommend that all Entities review what risks they face in relation to each of these nine categories, using the checklist tool in the TGF standard.

Strategic clarity	**Leadership**	**User focus**
Collaborative engagement	Skills	Supplier partnership
Achievable delivery	Future-proofing	Benefit realization

Section 6: Impact measurement

Scope and purpose of the section

This section should set out:

- The key benefits that the Entity seeks to deliver through implementation of its Digital Data Interoperability Roadmap
- How and when these will be measured by the Entity.

Issues to address

When developing this section of the Roadmap, you will need to consider several issues:

- What will success look like for our Entity in its use of digital data?
- How will success be measured and when?
- Who are the likely owners of each of the Entity-level benefits?
- How will learnings be incorporated?
- What systems and tools can be put in place to monitor the ongoing delivery of benefits?
- What quick wins can we deliver early in the program that will start the ball rolling?

[2] Refer to V2 of the standard published by international open standards consortium OASIS in 2014.

- What are the longer term benefits that we are seeking to achieve, and how do we sustain and embed the business changes required to achieve the desired impacts?

Actions you should take to inform development of this section

Your starting point should be the strategic objectives for UAE Digital Governance as a whole, as described in UAE Digital Data Interoperability Framework: overview and principles. These encapsulate the impact that the government as a whole is seeking to achieve.

Determine how you can best measure your own Entity's impact on the delivery of these objectives. Seek to develop success criteria and targets that are SMART:

1) Specific – clear and unambiguous;
2) Measurable – quantifiable;
3) Achievable – realistic and attainable;
4) Relevant – applicable and worthwhile;
5) Time-bound – delivered within a specific timeframe.

Resources to help you

Further advice on developing an effective approach to Benefit Realization is set out in the global standard on ICT-enabled, data-driven service transformation ('The Transformation Government Framework', or TGF[3]).

[3] Refer to V2 of the standard published by international open standards consortium OASIS in 2014.

GUIDANCE NOTE 3: DEVELOPING A DIGITAL DATA INVENTORY

Purpose	This document describes how to create a list of datasets which are collected, managed or maintained by the Entity. While it may not be possible to create a complete list in one step, this Guidance Note helps Entities ensure that the most valuable data assets are listed as an initial priority, and then to expand the Inventory over time.
When to use	When the Entity has a management structure and a team responsible for data in place (for example by using **Guidance Note 1: Digital Data governance roles and processes**)
Responsibility	Data Management Officer, reporting to the Entity's Director of Data.

Overview

To realize the strategic vision of efficient and effective digital data management in government that enables better decisions and better services, each Entity requires a good understanding of its current data assets and data processes. The first step of this is to produce an inventory of all datasets in the entity. This allows the entity to identify gaps where data is currently not fit for purpose, to spot and address duplication and become more standardized.

This Guidance Note provides help in listing and inventorying the data an entity holds and covers:

- What types of digital data should be considered;
- The process for producing the initial version of the inventory;
- The process for expanding and enriching the inventory over time (annual review).

Types of digital data

Structured digital data

Structured digital data that is machine-readable digital data such as a table in a spreadsheet, digital database, and digital data on a geospatial map is the main set of data which needs to be inventoried. Entities should list **existing** digital data which the Entity uses, maintains or collects. This could be data frequently used by the departments within this Entity which might not have a clear owner. It should also include all data where the Entity is responsible for collecting and updating the data, even if this work is done by others on its behalf.

Digital Data should be listed in the form of digital datasets**.** A **digital dataset** consists of digital data with its metadata. The metadata provides context and information about the data. Therefore, a digital dataset should be an individual object that makes sense as a whole by itself.

A dataset may be a database or spreadsheet along with its name, location, description. It could also be a map or a table from a report moved to a spreadsheet.

It may be more practical to count a collection of digital data, such as a digital database, as one dataset or as several. You should count it as one digital dataset if the digital data within a digital database is:

- thematically related
- easiest to describe as a whole
- Interrelated.

Otherwise, it likely consists of several digital datasets. The split depends on the existing and potential use of the digital data. It's up to the data owners who understand the digital data best to make the judgement decisions on how digital data should be listed as datasets.

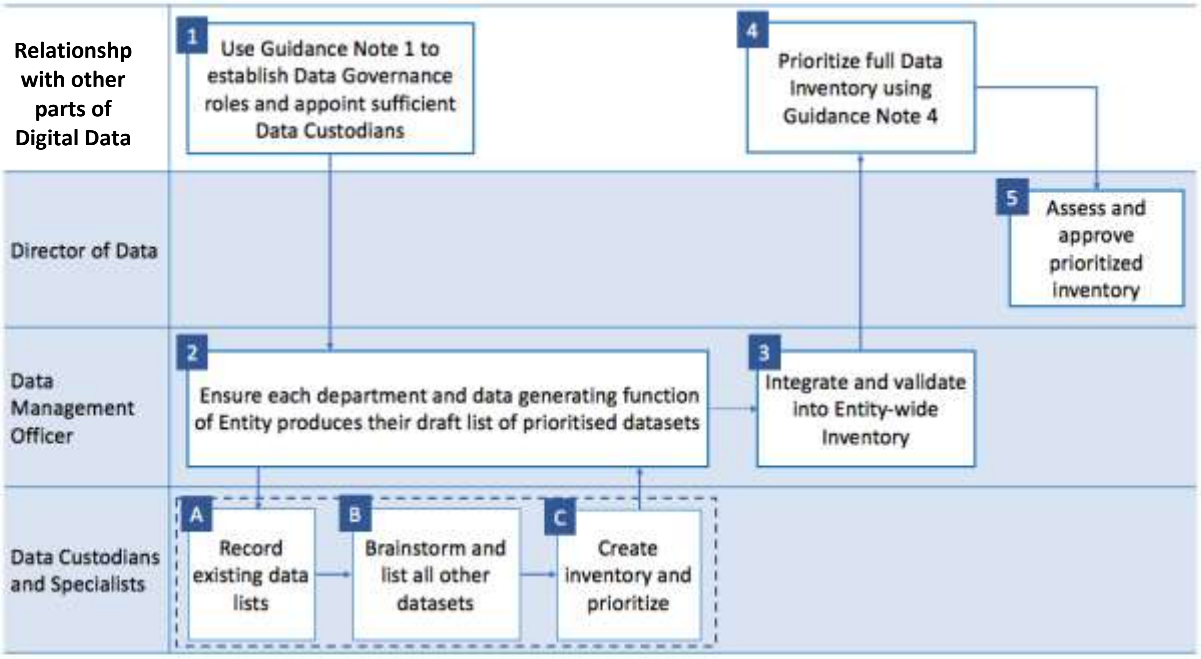
Unstructured digital data

While we recommend primarily focusing on structured data, unstructured data or information such as text documents, diagrams, pictures or media can also be important to publish or share effectively. Entities will generally have a lot more unstructured content and it will be difficult to inventory all of it, so two key steps are recommended:

1. Identify opportunities to turn unstructured data into structured digital data. For example, by putting tables in Word documents into a spreadsheet or seeing if there's a geodata version of a map picture available. Then deal with this structured data as explained in the process below.
2. Identify information that's particularly relevant for re-use – within the Entity itself or externally by the public or other Entities. This might be a report, presentation or videos which impart important information or can be utilized in new ways, and list these in the Inventory.

Recommended process for developing the initial Inventory

The diagram below summarises the process Entities should follow for developing the initial version of the Entity's Digital Data Inventory, illustrating which of the key data governance roles will normally have lead responsibility for each step of the process. Each step is then described in more detail below.



Relationshp with other parts of Digital Data

1 Use Guidance Note 1 to establish Data Governance roles and appoint sufficient Data Custodians

4 Prioritize full Data Inventory using Guidance Note 4

Director of Data

5 Assess and approve prioritized inventory

Data Management Officer

2 Ensure each department and data generating function of Entity produces their draft list of prioritised datasets

3 Integrate and validate into Entity-wide Inventory

Data Custodians and Specialists

A Record existing data lists

B Brainstorm and list all other datasets

C Create inventory and prioritize

1. Identity a digital data representative per department within the Entity

Each department or business unit within the Entity should have a named responsible Data Custodian who has a good understanding of the data their department produces, uses and manages. This person should be someone who is in a senior role (or appointed directly by a senior role) and regularly deals with data and is aware of the variety of data which exists within their department. They may be supported by a Data Specialist who has technical ownership or understanding of the data. For larger departments that handle a lot of data, this role may be covered by more than one person. Further guidance on the role of Data Custodian and Data Specialists is set out in **Guidance Note 1: Digital Data governance roles and processes**.

2. Ensure each department produces a draft list of digital datasets

Under the co-ordination of the Data Management Officer, the Data Custodian / Specialist within each department should go through the following process to develop an initial list of what data they know or expect to be managed within their department.

There is no need to change or rearrange data before adding information about the dataset to the list, or to collect any data which is not already held.

A: Record existing digital data lists

Draw together existing lists of datasets that are collected, maintained or managed by the Entity. These may include:

- The list of Primary Registries identified by the Entity Overseeing Digital Data.
- Data which has been previously requested by other Entities, external bodies, Entity Overseeing Digital Data or other departments within the current Entity.
- Data listed in the Entity's Information Asset Register (as required by the Information Security Regulation)
- Existing digital data catalogs or lists: e.g. digitaldata available in a catalog or portal, documentation of previous information audits, datacenter inventory, management databases or software asset lists.

B: Brainstorm and list all other datasets

Next the Data Custodian and/or Data Specialist should think about and list any datasets the Entity:

- Collects
- Stores
- Maintains and updates
- Commissions externally

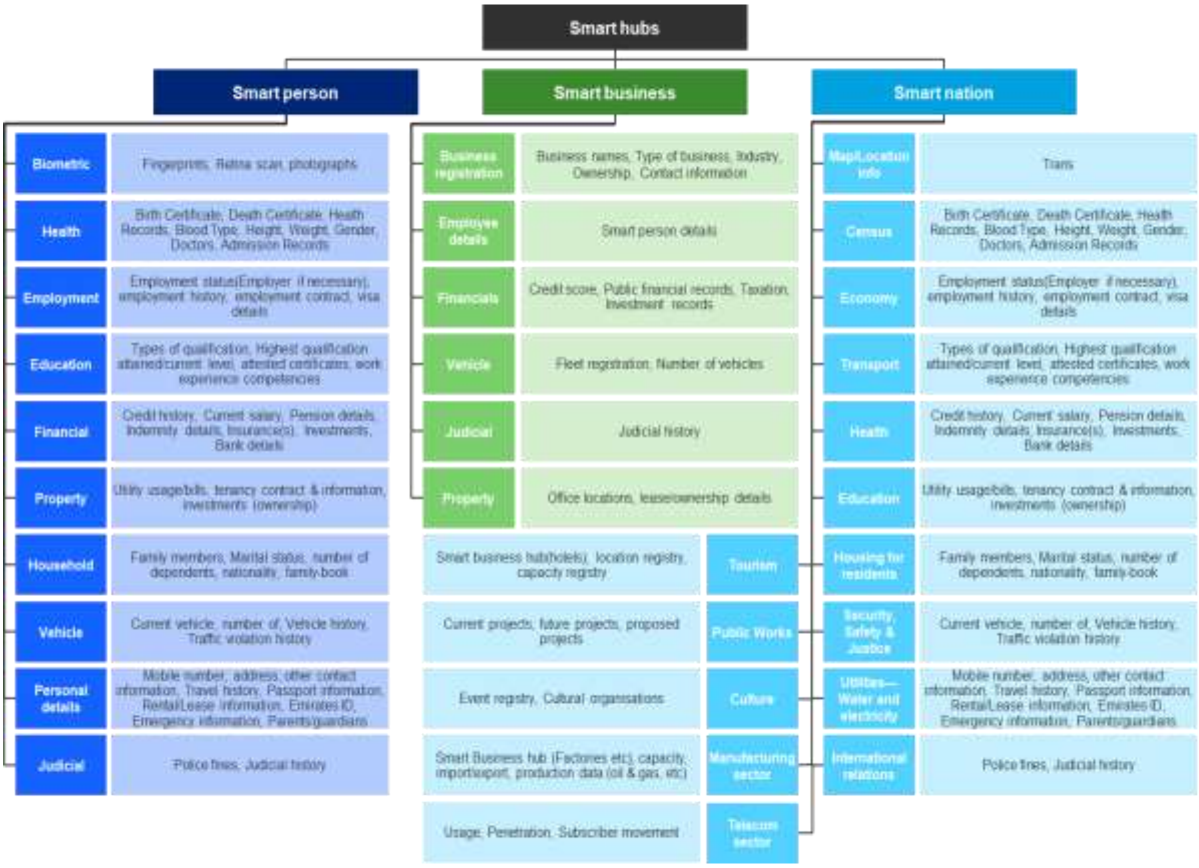
Aim to be as comprehensive as possible, but it is not expected that all datasets will be captured in the initial version of the inventory. It may be useful to consider:

- Any **digital datasets that have already been openly published or are currently shared** with other Government Entities.
- **Obvious or high-value** datasets: for example, data which can be used to provide a service to individuals or businesses (such as setting up a business, hiring a car, managing insurance, choosing where to live), make government more transparent, or any data which is expected to be managed by this Entity

- What datasets exist for each **type of digital data**: e.g. real-time, operational, reference data, aggregated data (see illustrative table below).

TYPE OF DATASET	DEFINITION	EXAMPLES	METRO SPECIFIC EXAMPLE
REAL-TIME DATA	Constantly updating data - often high volume and high velocity.	weather data; footfall through airport; cars passing toll booths; pollution levels; real-time location data; electricity usage	• Current location of metro trains • Current passenger numbers
OPERATIONAL DATA	These are the records that are made as part of an organisation carrying out its **day-to-day business.**	Entity organisation chart; forecast or modelling data; buildings owned/ maintained; budget; spending; staff levels; performance against metrics	• Staff numbers • Duty rota for staff at metro stations
REFERENCE DATA	**Authoritative or definitive data that rarely changes** about things, and that is necessary to help understand other data – often because it includes identifiers for those other things. Often produced by the public sector as a service in itself due to its high importance and value.	Timetables; names and locations of schools, hospitals, bus stops, metro stations; tax codes; land holdings; mapping data; indicators; address data; citizen id	• Metro timetables • Capacity per metro line • Metro station locations
AGGREGATED DATA	**Analysed and summarized data,** which provides overview information of reference or administrative data	hospital operation success rates; school exam pass rates; population statistics; housing; tourist numbers by month / year; nationalities of visitors	• Metro usage per month • Metro trends over time

- Any strategic **reference digital data** the Entity may hold, see table of examples below:



Smart hubs

Smart person

Biometric	Fingerprints, Retina scan, photographs
Health	Birth Certificate, Death Certificate, Health Records, Blood Type, Height, Weight, Gender, Doctors, Admission Records
Employment	Employment status(Employer if necessary), employment history, employment contract, visa details
Education	Types of qualification, Highest qualification attained/current level, attested certificates, work experience competencies
Financial	Credit history, Current salary, Pension details, Indemnity details, Insurance(s), Investments, Bank details
Property	Utility usage/bills, tenancy contract & information, investments (ownership)
Household	Family members, Marital status, number of dependents, nationality, family-book
Vehicle	Current vehicle, number of, Vehicle history, Traffic violation history
Personal details	Mobile number, address, other contact information, Travel history, Passport information, Rental/Lease information, Emirates ID, Emergency information, Parents/guardians
Judicial	Police fines, Judicial history

Smart business

Business registration	Business names, Type of business, Industry, Ownership, Contact information
Employee details	Smart person details
Financials	Credit score, Public financial records, Taxation, Investment records
Vehicle	Fleet registration, Number of vehicles
Judicial	Judicial history
Property	Office locations, lease/ownership details
Smart business hub(hotels), location registry, capacity registry	Tourism
Current projects, future projects, proposed projects	Public Works
Event registry, Cultural organisations	Culture
Smart Business hub (Factories etc), capacity, import/export, production data (oil & gas, etc)	Manufacturing sector
Usage, Penetration, Subscriber movement	Telecom sector

Smart nation

Map/Location info	Trams
Census	Birth Certificate, Death Certificate, Health Records, Blood Type, Height, Weight, Gender, Doctors, Admission Records
Economy	Employment status(Employer if necessary), employment history, employment contract, visa details
Transport	Types of qualification, Highest qualification attained/current level, attested certificates, work experience competencies
Health	Credit history, Current salary, Pension details, Indemnity details, Insurance(s), Investments, Bank details
Education	Utility usage/bills, tenancy contract & information, investments (ownership)
Housing for residents	Family members, Marital status, number of dependents, nationality, family-book
Security, Safety & Justice	Current vehicle, number of, Vehicle history, Traffic violation history
Utilities—Water and electricity	Mobile number, address, other contact information, Travel history, Passport information, Rental/Lease information, Emirates ID, Emergency information, Parents/guardians
International relations	Police fines, Judicial history

C: Create inventory and prioritize

Within a spreadsheet or table, add the following for each identified dataset from the previous two steps:

- A name or title for the dataset. If it doesn't have an existing name that you are aware of, choose a short descriptive name.
- A brief description to clarify what data is being referred to and its scope.
- The department or business unit responsible for managing the data
- The list of data attributes (normally column headings for tabular data) that are used in this dataset. Entities do not need to list data attributes for unstructured data.
- The Data Custodian, if known, i.e. the role or person within the department responsible for the data
- The Data Custodian's initial assessment of the extent to which this dataset should be a priority for initial publication as open or shared digital data (using the process and prioritization scoring set out in the **Guidance Note 4: Prioritization criteria and process**).

Example

Dataset name	Description and any notes	Responsible department	Data Attributes	Priority score	Data Custodian
Bus Transport timetables	All current bus timetables	Operations department	Bus number, bus stop location, arrival time, departure time, frequency	34/38	Mark Jones

3. Integrate and validate

Each draft list from each department should be reviewed by the head of that department and Data Management Officer, and then combined into a single Entity-wide list.

The Data Management Officer should then:

1. Co-ordinate with the responsible Data Custodian to repeat steps A, B and C to identity any missing datasets that may not sit within any individual department or which have been missed.
2. Ensure all datasets have an assigned Data Custodian.
3. Check and ensure dataset names are unique and their descriptions are clear.
4. If any information is missing, contact the relevant person to get this resolved.
5. Ask department data leads to confirm that in their view, the inventory includes all **existing** open datasets, all **obvious** datasets and all **high-value** datasets held by their unit.
6. Identify any core reference data used by the Entity where this Entity is not the owner. Take these out of the Inventory and forward to the Primary Registries team within the Entity Overseeing Digital Data for resolution. They will decide which Entity has ultimately responsible for maintaining these datasets and will be the authoritative source of the data. If this Entity is identified, those datasets will go back in the Inventory.
7. Ensure there are no duplicates in the inventory.

4. Prioritize

This Inventory should be taken through the prioritization process described in **Guidance Note 4: Prioritization criteria and process**. The Data Management Officer should ensure:

- all datasets have a prioritization score
- the scores have been reviewed and adjusted so that the prioritization of the whole list makes sense
- All digital datasets listed on the Inventory have been prioritized on a consistent basis.

5. Assessment and approval

The full initial Prioritized Inventory should be reviewed by the Director of Data who should verify that:

- the inventory contains a reasonably comprehensive list of data held by the Entity
- no key datasets are missing
- it was carried out by the appropriate staff members
- It contains the prioritization information specified in **Guidance Note 4: Prioritization criteria and process.**

Annual review – expanding the inventory

The inventory process should be repeated at regular intervals (for example, annually) in order to:

- Identify new datasets managed by the Entity (these could be completely new or extensions and reformulations of existing digital data)
- Respond to user demand for digital data
- Review the existing inventory in light of publication and sharing: both lessons learned and feedback received from other Entities, the public, external stakeholders and internal staff.

The process to follow should be similar, but instead of listing all possible datasets it should involve using the existing inventory as a basis and using each step of the process to see how the inventory can be expanded or amended. Expansion should cover both:

- **Extending** the inventory, by adding new datasets
- **Enriching** the inventory, by increasing the proportion of datasets that have been catalogued (using the Digital Data Exchange Standard and Digital Data Quality Standard).

GUIDANCE NOTE 4: PRIORITIZATION CRITERIA AND PROCESS

Purpose	It will not be possible to ensure all the Entity's datasets meet the Digital Data Interoperability Framework requirements at once. This Guidance Note provides criteria and a clear process for prioritising the order in which digital datasets should be made conformant to the standards prior to their publication or exchange.
When to use	After producing an inventory of the Entity's data assets, and before proceeding to take the initial highest priority group of datasets through the Digital Data Conformance Process described in Guidance Note 5.
Responsibility	Data Management Officer

Overview

This Guidance Note helps Entities focus resources on making the most important datasets conform to the Digital Data Interoperability Framework standards first and ensuring there is a clear prioritized plan for which datasets will be ready for publication and exchange next.

Once a Government Entity has prepared an initial draft of its Digital Data Inventory (using **Guidance Note 3: Developing an Entity-wide Digital Data Inventory**), it should not seek to ensure full conformance with the UAE Digital Data Interoperability Standards for all its data at once. We recommend prioritising which of the inventoried datasets it should prepare first for publication as open data or for exchange with other Entities.

By starting with a subset of its data inventory, Entities can:

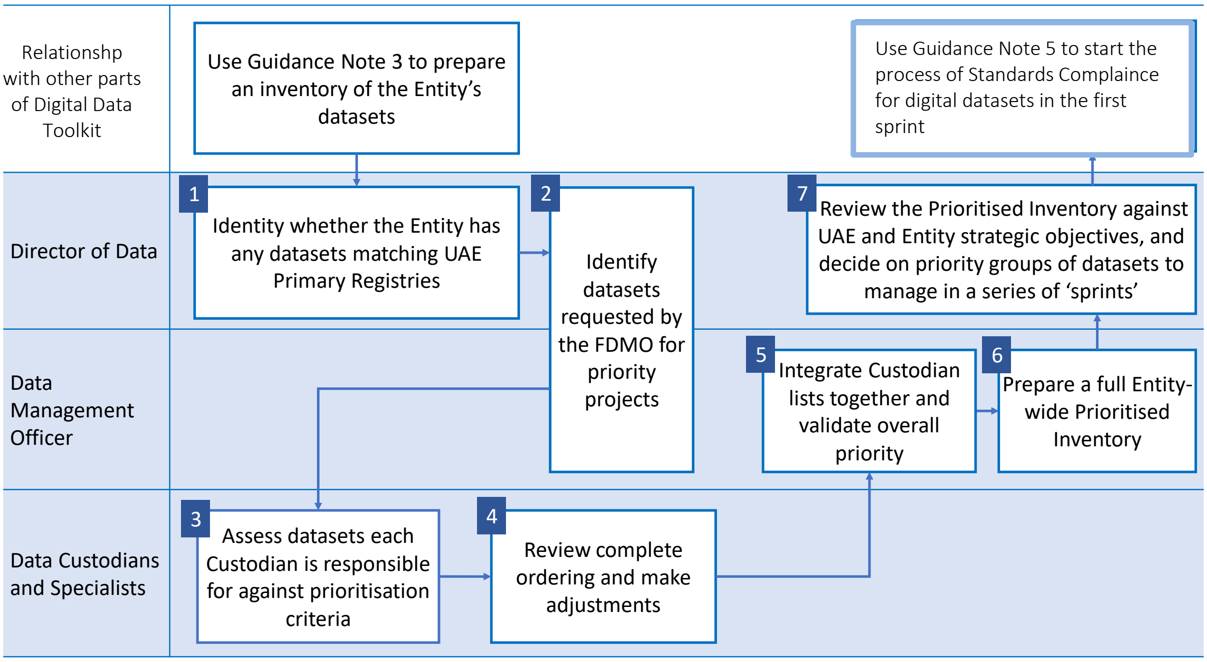
- Quickly publish and exchange high value and low effort digital data
- Go through the process faster, learn from it and adopt desired changes to the process in future.

The guidance below looks in turn at:

- The recommended process for prioritising digital datasets
- The criteria which are recommended for use within the prioritisation process.

Recommended process for prioritizing the inventory

The diagram below summarises the process that Government Entities are recommended to follow when prioritising their data, illustrating which of the key data governance roles will normally have lead responsibility for each step of the process. Each step is then described in more detail.


Relationshp with other parts of Digital Data Toolkit
Director of Data
Data Management Officer
Data Custodians and Specialists
Use Guidance Note 3 to prepare an inventory of the Entity's datasets
Use Guidance Note 5 to start the process of Standards Complaince for digital datasets in the first sprint
1
Identity whether the Entity has any datasets matching UAE Primary Registries
2
Identify datasets requested by the FDMO for priority projects
7
Review the Prioritised Inventory against UAE and Entity strategic objectives, and decide on priority groups of datasets to manage in a series of 'sprints'
5
Integrate Custodian lists together and validate overall priority
6
Prepare a full Entity-wide Prioritised Inventory
3
Assess datasets each Custodian is responsible for against prioritisation criteria
4
Review complete ordering and make adjustments

1. Identify primary registries digital datasets

Identify whether the Entity has any datasets relevant to the UAE Primary Registries as identified by the Entity Overseeing Digital Data. If so, these and any dependent datasets should be given top priority.

2. Identify digital datasets required for priority projects

Identify digital datasets required or requested by the Entity Overseeing Digital Data for priority national projects. This could be for:

- Cross-Entity projects,
- Digital Datasets relevant to achieving national indicators and realizing Digital Governance goals.

Discuss requirements with the Entity Overseeing Digital Data and the Digital Data Interoperability Electronic Platform team. Place recommendations next on the priority order.

3. Assess Inventory against prioritization criteria

Each Data Custodian should now assess the datasets they are responsible for (using their Digital Data Inventories list), against the prioritization criteria described this Guidance Note. These criteria assess both data readiness and the benefit or value of sharing that data, resulting in a balanced score (with the overall score for 'priority' being an equally-weighted average of the scores for benefit and readiness).

The priority order should be noted in each Custodian's departmental inventory.

4. Review Ordering

Data Custodians should review their complete prioritized list (which includes primary registries, datasets needed for projects and results of applying the prioritization criteria to the digital data in their inventory) and re-arrange as needed. For example, if there are many datasets with the same score, use judgement to prioritize between them or if a dataset looks out of place and feels like it should be above or below others, rearrange as needed.

5. Integrate and validate overall priority

The Data Management Officer should review the prioritized inventories of each department and combine them together. Then, follow steps 1 and 2 to additionally prioritize other datasets which do not belong within a specific department.

The combined whole prioritized list now needs to be judged for whether the ordering makes sense.

As part of this, the Entity may also want to consider having a spread of types of datasets in its initial batch of priority datasets – some very easy to publish, some very valuable to test the comparative effort and impact of taking these through to publication.

6. Add priority to Entity's full Digital Data Inventory

The Data Management Officer should ensure that the validated prioritization scores are included within the Data Inventory then reviewed and signed off by the Director of Data.

7. Review and prepare sprints

The Director of Data should satisfy themselves that the prioritized list makes sense and aligns with the strategic aims of the Digital Data Interoperability Framework as well as within the aims of the Entity itself. Then the final approved list should be split into batches of prioritized datasets that make sense to tackle together through a series of 'sprints' through the Digital Data Conformance process described in **Guidance Note 5** of this Implementation Guide.

Recommended criteria for assessing priority

There are two broad sets of criteria to consider:

1. **Benefit** criteria for evaluating the potential value of opening a particular dataset to individuals and Private Sector Entities, or sharing it with other Government Entities.
2. **Readiness** criteria for evaluating the effort involved in getting the dataset ready for publication or exchange.

In combination, these two sets of criteria help identify the datasets which will have the most impact with the least effort. They have been chosen because they balance each other. If a dataset is very high value, but a lot of work must go into making it publishable or reusable, then it is still fairly high on the priority list, but below digital data which is both. Similarly, just because the data will be very easy and quick to publish, does not mean it should be focused on first unless there's some potential benefit in its publication and sharing.

The tools below provide simple-to-use recommended approaches for quantifying both of these dimensions.

1. Assessing potential benefit

First, the benefit criteria measure the *potential value* or *degree of benefit* that could be created for individuals, Government Entities and Private Sector Entities in opening up or exchanging each digital dataset. This provides a simple way to evaluate the comparative impact that publishing different data would have on the strategic goals of UAE's Digital Data Interoperability program.

Each dataset should be given a score out of 5 for each of the following four questions, recording the score in the right hand column:

Benefit criteria	**Question**	**Choose and record answer in score**	**Score**
User demand for digial data	How likely is it that individuals, Government Entities, Private Sector Entities would want to use or have access to this digital data?	1 – highly unlikely: no evidence and no plausible reasons this would be relevant 2 – unlikely 3 – possible 4 – highly likely: we can think of good reasons others may want this data 5 – Definite: we've already seen requests for this data	
Economic impact	If we open up this data, how likely is it that private-sector Entities could use it - perhaps combined with other data - to create commercially valuable products and services?	1 – highly unlikely 2 – unlikely 3 – possible 4 – highly likely 5 – Definite: we have clear evidence that private-sector Entities want to exploit this data commercially	
Better services	How likely is it that exchanging this data will lead to innovations and services that improve the quality of life for people in UAE?	1 – highly unlikely 2 – unlikely 3 – possible 4 – highly likely 5 – Definite: we have clear evidence of the quality of life gains that could be made	
Better governance	How likely is it that will exchanging this data will improve the efficiency, transparency and accountability of Government Entities?	1 – highly unlikely 2 – unlikely 3 – possible 4 – highly likely 5 – Definite: we have clear evidence of efficiency; transparency or accountability gains that could be made	
		Total score out of 20:	

2. *Assessing digital data readiness*

Secondly, the readiness criteria assess the state and quality of the digital data. This is to establish how much work is needed to prepare it for publication or exchange. Digital Data which is of high quality, already documented, up to date and with a clear owner can be more easily published or exchanged. These criteria help to identify ‘quick wins’ for the Entity.

Please assess each dataset for the following, recording the score in the right-hand column:

Readiness criteria	Question	Scoring criteria (pick most suitable score and 1 – medium if unsure)	Score
Accuracy	How accurate is the digital data?	2 – High accuracy (we review and check accuracy) 1 – Medium accuracy 0 – Low accuracy (there are known errors in the data)	
Completeness	How complete is the digital data?	2 – High completeness (we have all the data at current granularity) 1 – Medium completeness 0 – Low completeness (there is known missing data, or this data will not make sense by itself)	
Timeliness	How up to date is the digital data?	2 - Latest month / week / year is available 1 – Digital Data is not time sensitive OR we have all the data apart from latest month / week / year 0 – Digital Data is out of date	
Validation	Does the digital data use a schema or is standardised?	2 – Yes, digital data is published with same headings / fields (schema) each time 1 – The digital data does not use a schema AND is not published regularly (i.e. it is one off data) 0 – Digital Data is regularly updated, but does not use a set schema	
Ownership	Is there a clear specific digital data owner?	2 – Yes 0 – No	
Description	Does the data have existing metadata - that is, information on what the digital data is about, how it was generated etc.?	2 – Yes 0 – No	
Accessibility	Is the digital data already published somewhere or available on the web / through an API?	2 – Yes 0 – No	
Interoperability	Is the digital data in an open machine-readable format?	2 – Yes	

		0 – No	
License	Does the digital data have a license?	2 – Yes 0 – No	
		Total Score out of 18	

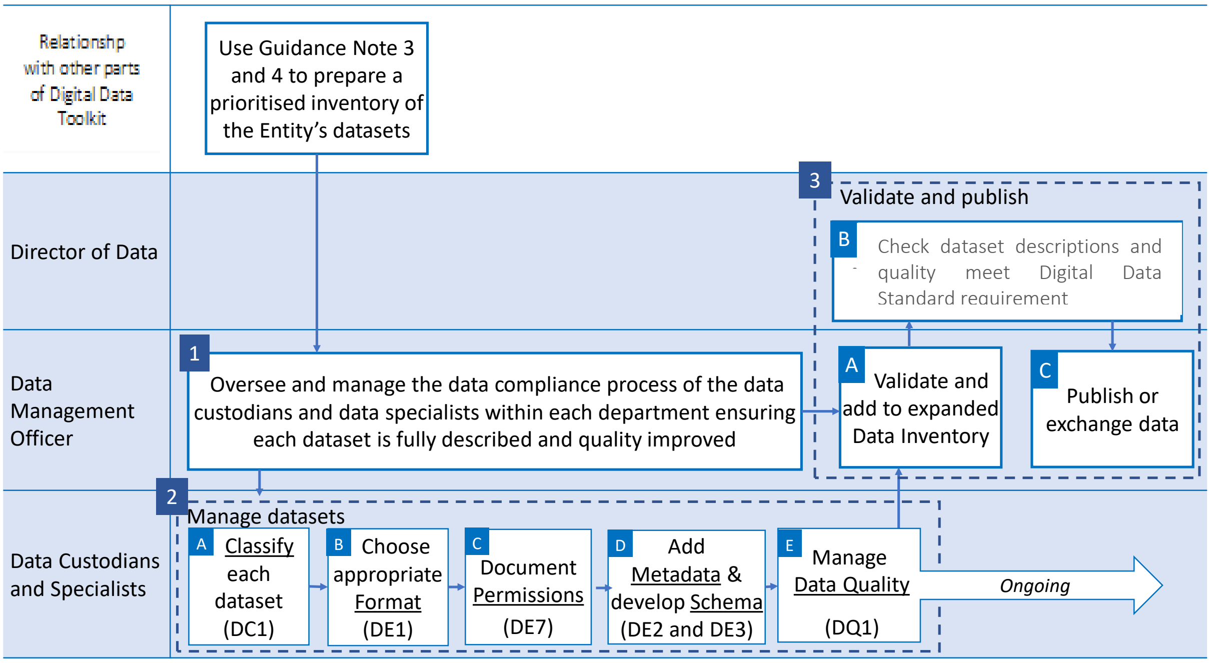
Combine the two scores to get an overall priority score out of 38.

GUIDANCE NOTE 5: DIGITAL DATA CONFORMANCE PROCESS

Purpose	This Guidance Note outlines the process that Entities are recommended to follow to ensure their data conforms with the Digital Data Interoperability Standards, and is ready for publication and exchange.
When to use	Before publishing digital data as open data or exchanging shared data with other Entities. Entities should focus in turn on successive batches of digial data that have been prioritised for data conformance in line with the advice in **Guidance Note 4: Prioritization process and criteria**.
Responsibility	The Director of Data has overall accountability for ensuring effective systems are in place to manage data conformance, but the lead responsibility for operating these systems will lie with the Data Management Officer

Overview of recommended process to ensure data conforms to standards

The diagram below summarises the process that Government Entities are recommended to follow as they seek to ensure that their datasets comply with the mandatory requirements of the UAE Digital Data Interoperability Standards. The diagram illustrates which of the key data governance roles will normally have lead responsibility for each step of the process – and is followed by a brief description of each step in the process.


Relationshp with other parts of Digital Data Toolkit
Use Guidance Note 3 and 4 to prepare a prioritised inventory of the Entity's datasets
Director of Data
3
Validate and publish
B
Check dataset descriptions and quality meet Digital Data Standard requirement
Data Management Officer
1
Oversee and manage the data compliance process of the data custodians and data specialists within each department ensuring each dataset is fully described and quality improved
A
Validate and add to expanded Data Inventory
C
Publish or exchange data
Data Custodians and Specialists
2
Manage datasets
A
Classify each dataset (DC1)
B
Choose appropriate Format (DE1)
C
Document Permissions (DE7)
D
Add Metadata & develop Schema (DE2 and DE3)
E
Manage Data Quality (DQ1)
Ongoing

1: Manage and coordinate

It is recommended that the Data Management Officer is responsible for overseeing and helping the Entity's Data Custodians and Data Specialists complete steps 2[A] – 2[E] above, by:

- Establishing a clear internal timetable for completing the data conformance process, aligned with Entity and Federal milestones for publishing or exchanging data
- Ensuring that Data Custodians and Data Specialists are fully briefed on their roles and on the requirements of relevant Digital Data Interoperability Framework standards
- Facilitating opportunities for Data Custodians and Data Specialists to come together and exchange experiences and lessons learned through the process.

A summary is set out below of the steps that need to be taken as part of this coordinated approach:

- Steps 2[A] to 2[E] look at the actions which Data Custodians and/or Data Specialists should take to ensure that an individual dataset is conformant with the UAE Digital Data Interoperability Standards
- Steps 3[A] to 3[C] then look at the actions which the Data Management Officer and Director of Data should then take to validate and approve for publication the datasets that have come through this process.

2: Management of datasets by Data Custodians and Data Specialists

2A: Classify

The Data Custodian should classify the digital dataset as *Open*, *Confidential, Sesnitive or Secret*, in accordance with the **[DC1] Digital Data Classification** specification. Detailed guidance on the process to follow is given below in **Guidance Note 5.1: Classifying digital data**.

Once this is done, you might be left with the original dataset and one or more derived (or 'child') datasets which have been modified to allow an Open classification. Both the original and derived digital datasets should be catalogued separately in the following steps.

2B: Choose format

Decide on an appropriate format in which to make the digital data available that complies with the **[DE1] Digital Data Formats** specification and produce a sample dataset in that format. Detailed guidance on the process to follow is given below in **Guidance Note 5.2: Formatting digital data**.

2C: Decide on and document Shared Digital Data Access Permissions

For data which will be exchanged with other Enitities rather than published as Open Data, you will need to comply with **[DE7] Shared Digital Data Access Permissions**. This will involve determining and then documenting the appropriate permissions model. Detailed guidance on the process to follow is given below in **Guidance Note 5.3: Documenting a permissions model for shared digital data**.

2D: Add Metadata

Describe each dataset with metadata ensuring that all Core Metadata fields required in **[DE2] Metadata** are complete and as many Optional Metadata fields as can be easily filled in. Detailed guidance on the process to follow is given below in **Guidance Note 5.4: Adding metadata and schema**.

2E: Manage data quality

Assess the data against the **[DQ1] Digital Data Quality Principles**. The Data Custodian should then:

- Identify and implement any 'quick wins'
- Then develop a longer term plan for improving the quality of the dataset to better meet user requirements.

The Data Custodian's work on this will need to feed into broader work on improving data quality in the Entity, as detailed in **Guidance Note: 2 Building a Digital Data Interoperability Roadmap**.

Detailed guidance on the process that Data Custodians should follow is given below in **Guidance Note 5.5: Managing data quality**.

3: Validation and publication

Once a digital dataset has gone through the process described in Steps [2A] to [2E] above, the Data Management Officer and Director of Data will need to validate and approve the dataset either for publication or for sharing and exchange with other government entities over the Digital Data Interoperability electronic platform. Further guidance on this – along with guidance on when these decisions need also to be approved by the Entity Overseeing Digital Data) – is given below in **Guidance Note 5.6: Validation and publication of digital data**.

5.1 Classifying Digital data

Purpose	This Guidance Note outlines the recommended process for ensuring that a dataset has been correctly classified in accordance with UAE Digital Data Interoperability Standards.
When to use	Before a dataset is published as open data or exchanged with other Entities, the data should be correctly classified.
Responsibility	The Data Custodian that the Entity has identified as accountable for a particular dataset should be the lead person responsible for applying the **Digital Data Classification Standard** to the dataset. The Entity's Data Management Officer is responsible for supporting all Data Custodians across the Entity as they undertake this task, and for ensuring a consistent approach at the Entity-wide level. The Entity's Director of Data is responsible for reviewing and approving classifications of all datasets.

Overview

At the start of the Digital Data Conformance Process, it may be that a dataset has already been classified as *Open*, *Confidential, Sensitive* or *Secret* – because FGEs have been required to use such a classification for several years (see for example the 'Regulation of Information Security at the Federal Entities of UAE Cabinet Resolution' No. (21), 2013). Previously, however, Entities were free to establish their own criteria for determining what sort of digital digitdata they assigned to each class. Now, following agreement of the UAE Digital Data Interoperability Standards, there is a common government-wide set of criteria which all Government Enities should apply. These criteria are intended to enable much greater levels of open data publication and data exchange between organisations than has historically been the practice in the UAE.

The table below gives criteria for assessing what data falls into each class, with examples. Deciding the classification level depends on a risk assessment to assess the level of damage that may result from unrestricted disclosure of the data (to privacy, security, commercial confidentiality etc).

Data classification

Open	**Criteria:** Data that can be openly disclosed to individuals, governmental, semi-government entities and private sector for use, re-use and sharing with third parties. This should be the default classification for all non-personal data, and exceptions to this should have a documented rationale that clearly explains why open publication of the data would contravene specific criteria listed below that require classification as Confidential, Sensitive or Secret
	Examples: Open data can include: • **Real time data:** constantly updating data, often high volume and high velocity

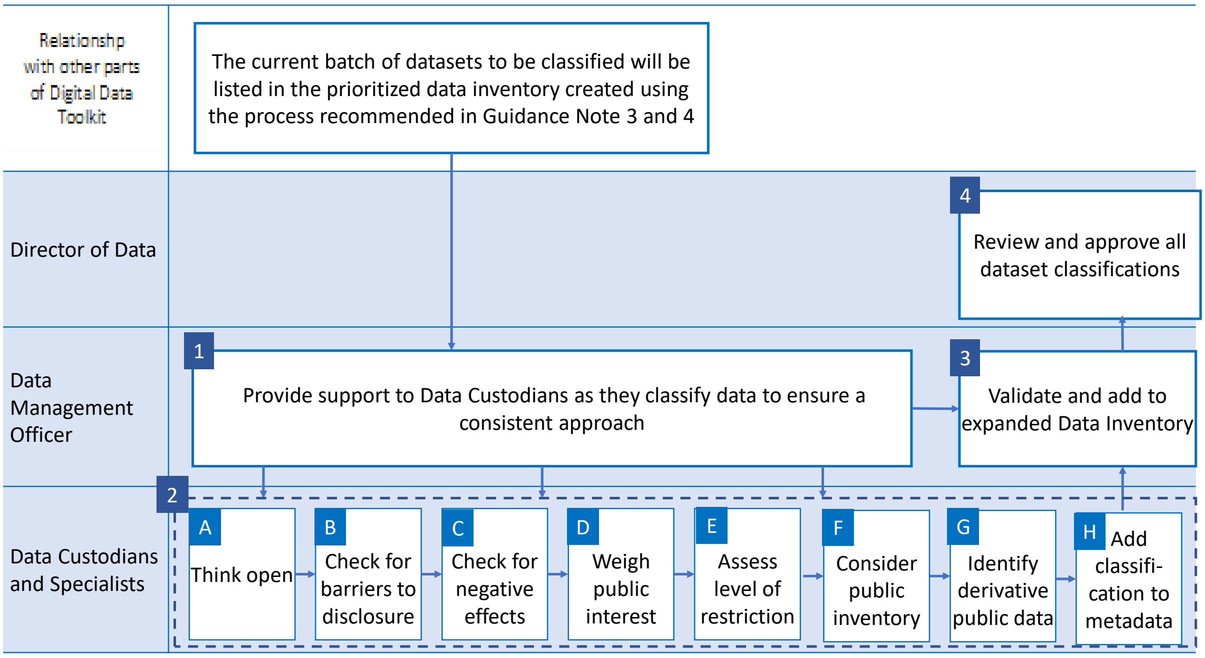
	Examples include: weather data; footfall through airport; cars passing toll booths; pollution levels; real-time location data; electricity usage • **Operational data:** the records that are made as part of an Entity carrying out its day to day basis *Examples include: Entity organisation chart; forecast or modelling data; buildings owned/maintained; budget; staff levels; performance against metrics* • **Reference data:** authoritative or definitive data that rarely changes about things *Examples include: timetables; names and locations of schools, hospitals, bus stops; tax codes; land holdings; mapping data; indicators; address data* • **Aggregated data:** analysed and summarised data, which provides overview information in relation to other types of data *Examples include: hospital operation success rates; school exam pass rates; population statistics; housing; tourist numbers by month/year; nationalities of visitors*
Confidential	**Criteria:** This is the default classification for datasets containing personal data which is non-sensitive. "Personal data" means any information relating to an identified or identifiable natural person; an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that person. Non-sensitive personal data refers to all types of personal information which are not 'confidential' (as defined in the criteria for Sensitive Data below). In addition, data should be classified as *Confidential* if unrestricted disclosure or exchange of the data cause No damage to government bodies, companies or individuals such as: • Adversely affecting or preventing the ability of a Government Entity to carry out its day to day duties • No damage to assets, or limited financial loss of an Entity, company or individual • Limiting the competitiveness of companies and negatively affecting the principle of equal opportunities • Adversely affecting public safely, criminal justice and enforcement activities.
	Examples: Typically, non-sensitive personal data will include information which is personal but does not impact on the reputation of the person. Examples include name, date of birth and address. Examples of other types of *Confidential* information include: • Minutes of meetings, internal regulations and policies, and government-body performance reports • Correspondence within a government body or with other government bodies or third parties • Financial transactions and financial reports • Company data such as tenders or contracts which provide for non-disclosure clauses • Individual's dealings with the government, which include personal data (details of ownership of properties of various kinds, commercial or professional licenses, personal documents, residence permits, visas, and leases).

Sensitive	**Criteria:** This is the default classification for datasets containing sensitive personal data. Sensitive personal data are personal data that directly or indirectly reveal an Individual's family, racial or ethnic origin, sectarian origin, political opinions, religious or philosophical beliefs, their union membership, criminal record, health, sexual orientation, genetic data or biometric data In addition, data should be classified as *Sensitive* if unrestricted disclosure or exchange of the data may cause **limited damage** to government bodies, companies or individuals such as: • Infringing Intellectual Property Rights • A significant decline in the ability of one of the bodies to carry out its functions, limited damage to its assets, or significant financial loss • Causing limited damage to companies that could lead to loss of competitiveness, or loss of some of its core cognitive and intellectual advantages or incurring heavy financial loss • Limited damage to the operational effectiveness of the police, security forces, military forces, intelligence services or the administration of justice • Limited damage to relations with friendly governments or damages to international relations resulting in formal protest or sanctions.
	Examples: For example, this might be the details and content of: • Draft government laws and policies and legislation • Audit reports of a government body • Employees' complaints and investigation minutes • Staff salaries and performance reports • Confidential financial expenses • Data, plans or technical documentation for technological information systems and networks of a governmental body • Credit card or bank accounts data • Judgments, irregularities or violations under investigation relevant to individuals • Attachment orders over assets and property of individuals and companies.
Secret	**Criteria:** Data the unrestricted disclosure or exchange of which may cause **significant damage** to the supreme interests of the United Arab Emirates and **very high damage** to government bodies, companies or individuals, such as: • Disclosing any personal information of a VIP (very important person) or infringing any Intellectual Property Rights of a VIP • A significant or noticeable negative impact to the supreme interests of the United Arab Emirates • A sharp decrease in the ability of one of the vital bodies to carry out its functions, or very high damage to its assets, heavy financial loss, clear negative impact on the image of the body and a loss of public confidence in such body and in the government in general • Causing significant damage to private sector entity that have vital and strategic roles in the national economy, which may lead to heavy financial losses, bankruptcy or loss of its leading role • Seriously endangering the safety and lives of certain individuals associated with a security role (e.g., security forces and police) or as parties to serious judicial cases (e.g. witnesses)

	• Information the disclosure of which would negatively affect the maintenance of security and the administration of justice, or cause major, long-term impairment to the ability to investigate or prosecute serious crimes.
	Examples: Examples include details and content of: • Security reports, minutes or orders • Sensitive minutes and reports of the Council of Ministers or its committees • Agreements or contracts of a secret nature between the United Arab Emirates with other countries or individual Emirates • Government Entity's data, plans, operating systems which would significantly damage the production of energy or water, infrastructure networks or traffic control or communications systems • Security forces data, including the facilities, equipment, personnel and operation systems • Data and regulations of individuals and entities under control or blacklisted • Data of control and surveillance systems and entry and movement control systems at vital institutions • Data relevant to security detectives, spies or witnesses in serious lawsuits • Data relevant to Government strategic financial investments of nature (national companies, investment funds, off-shore companies) • Attachment or travel ban orders.

It is therefore vital that every dataset prioritised for open publishing and for inter-Entity exchange has its classification status reviewed against the requirements of the Digital Data Classification Standard.

The diagram below summarises the process that the Government Entity is recommended to follow when classifying a dataset against the Digital Data Classification Standard mandated in the UAE Digital Data Interoperability Framework.


Relationshp with other parts of Digital Data Toolkit
The current batch of datasets to be classified will be listed in the prioritized data inventory created using the process recommended in Guidance Note 3 and 4
Director of Data
4
Review and approve all dataset classifications
Data Management Officer
1
Provide support to Data Custodians as they classify data to ensure a consistent approach
3
Validate and add to expanded Data Inventory
Data Custodians and Specialists
2
A
Think open
B
Check for barriers to disclosure
C
Check for negative effects
D
Weigh public interest
E
Assess level of restriction
F
Consider public inventory
G
Identify derivative public data
H
Add classifi-cation to metadata

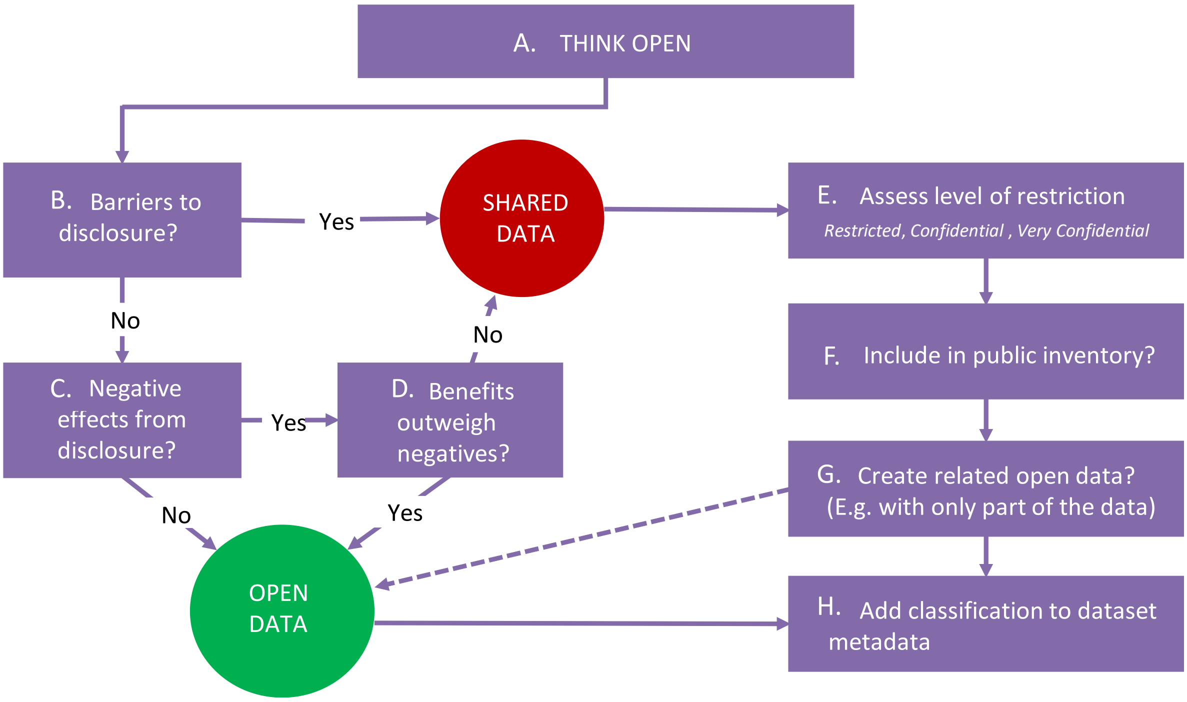
Step 1: Manage and coordinate

It is recommended that the Data Management Officer is responsible for overseeing and helping the Entity's Data Custodians and Data Specialists complete steps [2A] – [2B] above, by:

- Establishing a clear internal timetable for completing the digital data classification process, aligned with Entity and Federal milestones for publishing exchanging digital data
- Ensuring that Data Custodians and Data Specialists are fully briefed on their roles and on the requirements of Digital Data Classification Standard
- Facilitating opportunities for Data Custodians and Data Specialists to come together and exchange experiences and lessons learned through the process.

Step 2: Classify each digital dataset

For each dataset, the responsible Data Custodian should classify the dataset using the eight step process recommended above. These steps that an individual Data Custodian should go through can also be visualized as a logical decision model, as illustrated below.



A. THINK OPEN
B. Barriers to disclosure?
Yes
SHARED DATA
E. Assess level of restriction
Restricted, Confidential , Very Confidential
No
No
C. Negative effects from disclosure?
Yes
D. Benefits outweigh negatives?
F. Include in public inventory?
G. Create related open data? (E.g. with only part of the data)
No
Yes
OPEN DATA
H. Add classification to dataset metadata

Each of the steps on this process are described below.

2A. Think Open

It is vital to recognize the UAE Government's strategic commitment to high levels of openness. When following the steps of this procedure, the default assumption about a dgital dataset should be that it will be classified as open. Exceptions require a compelling case linked to clear criteria, which should be documented and then personally signed off by the Entity's Director of Data.

‘Thinking Open’ is often the most difficult part of the classification procedure, especially if the Entity is inexperienced with open data. Staff may be concerned that publication will reflect badly on them where, for example, some of the data may be interpreted as unfavorable, or the data may have gaps or inaccuracies. It is vital that staff understand that they will be have the backing and support of the management for the decision to publish digital data in which problems are later found. Such problems plague all Entities and all digital data, and publication should be seen as an opportunity to help find and improve errors and problems.

For these reasons, it is very helpful if at the start of the data conformance process, the senior management communicate to the staff their and the Entity’s commitment to openness. The Director of Data should be available to respond to any concerns raised by staff.

The following steps should be carried out by the person(s) most familiar with the data, such as the Data Custodian, for each of the datasets in the current batch going through the data conformance process.

2B. Check for barriers to disclosure

There are certain criteria that may preclude a dataset being classified as ‘*Open’* and then disclosed as Open Data. These include two absolute barriers to disclosure. A digital dataset cannot be Open if its publication would:

- violate existing legislation or laws; or
- Represent a significant threat to the supreme national interest and/or national security.

Check that your digital dataset’s publication would not violate one of these conditions. In most cases it should be obvious if one of these barriers applies, but in cases of doubt you may need to consult your Entity's legal department. If a digital dataset is barred from publication by one of these barriers, then it cannot be classified as Open. Proceed to Step [E] to determine whether it should be classified as *Confidential*, *Sensitive* or *Secret*.

2C. check for harmful effects of disclosure

If the barriers to disclosure in Step [B] do not apply, then there are other possible harmful effects to consider before the digital data can be confirmed as open. Consider whether release of the dataset would entail a significant risk of one or more of the following by checking whether the answer ‘yes’ to any of the questions listed in the checklist below.

Risk	Questions and notes
A breach of the privacy of any individual	**1. Consider whether any individuals can be identified from this digital data?** This would apply if the dataset includes digital data about identifiable individuals - for example, their address, medical history, date of birth, or tax information. **Note**: A person does not need to be named to be identifiable. If the digital data contains information about individuals, even if the individuals cannot be easily identified, they may become identifiable when the data is combined with other publicly available information or digital datasets. Any release of data at the level of individuals, or small groups of individuals such as households, is likely to run this risk. **If the answer is yes**, then the Entity should: • Classify the digital dataset as Confidential if the digital data relates to Personal Information or Confidential if it relates to Sensitive Personal Information • Try to create a derivative dataset which can be classified as 'Open'. This could be achieved by anonymizing the digital data, taking out digital data referring to small sample sizes, aggregating or summarizing the data, or taking out attributes which hold the personal data. Once one or more derivative datasets have been created: – Check whether any of the other risk assessment criteria apply. – If not, then classify as 'Open' and add to the Digital Data Inventory. **Note**: It should almost always be possible to create a version of the data which does not breach privacy. In many cases, these summary, anonymized digital data sets will already exist within the Entity: providing statistical, analytical and management information for use in running the service which has generated the more detailed personal digital data.
A breach of legal rights or agreements (such as Non-Disclosure Agreements, Intellectual Property rights or release of commercially sensitive information)	If publication of the digital data would breach an existing legal agreement, the dataset should initially be classified as Shared and steps taken to see if parts could be published openly or such agreements renegotiated to allow publication in future. Specifically, consider the following questions: **2. Does the legal statute under which the Entity is empowered to collect the digital data from or relating to a Private-Sector Entity place a duty on the Entity to keep that digital data confidential?** If the answer is yes, then the Entity should classify the Digital Data set as *Confidential or Sensitive*, depending on the degree of damage that would be caused by breach of confidentiality (see Step E).

	3. Does the Entity have a Non-Disclosure Agreement or other contract in place with one or more Private-Sector Entities that places contractual obligations to keep the data confidential? If the answer is yes, then the Entity should: • Consider whether it would be helpful to approach the relevant Private-sector Entities to seek their agreement to voluntarily agreeing to waive their non-disclosure rights in relation to some or all of their data • If not, classify the digital dataset as *Confidential or Sensitive*, depending on the degree of damage that would be caused by breach of confidentiality (see Step E) • At any future review points or renewal points in the contract, consider the scope for re-negotiating the contract to enable greater disclosure of Open Data in future. **4. Does a Private-Sector Entity hold Intellectual Property Rights in some or all of the digital data?** If the answer is yes, then the Entity should: • Engage with the IPR holder to establish whether it will give consent to opening up the data, potentially with some license restrictions • If not, classify the digital dataset as *Confidential or Sensitive*, depending on the degree of damage that would be caused by breach of confidentiality – unless there is an overriding public interest in publishing. (See Steps D and E) • In cases where the Private-Sector Entity's IPR arises from the performance of a commercial contract on behalf of the Government Entity, seek to re-negotiate these contract terms, particularly at any contract review or renewal points. **Note:** In answering this question, Government Entities should note that it is not acceptable to classify a digital dataset as *Confidential* on the grounds that a Government Entity has Intellectual Property Rights in the data, even in cases where it is currently exploiting that IPR on a commercial basis. Rather, the dataset should be classified as Open Data, albeit with consideration given to the nature of the licensing and pricing basis on which it is made Open.
Risk to the safety of individuals and society	Consider: **5. Would disclosure of this data pose risks to the health and safety of individuals or to public health and safety?** **6. Would disclosure of this digital data pose other risks to society?** If any risks identified under these two questions are: ✓ Specific and clear, not general and vague ✓ Evidence-based

	…. then the digital dataset should be classified as *Confidential or Sensitive*, with reasoning documented with sufficient detail that external stakeholders will be able to understand the rationale and subject it to challenge. **Note**: greater transparency is in general a force for social good rather than a social risk.
Risk of negatively affecting the administration of justice and maintenance of security	**7. Consider whether disclosure of this data pose risks to the administration of justice and maintenance of security?** If any risks identified under these two questions are: ✓ Specific and clear, not general and vague ✓ Evidence-based Then the digital dataset should be classified as *Confidential or Sensitive*, and reasoning documented with sufficient detail that external stakeholders will be able to understand the rationale and subject it to challenge.
A significant negative impact on the work and effectiveness of government	**8. Consider whether the disclosure of this data cause significant negative impact on the effectiveness with which Entity or other Government Entities can deliver its work and objectives?** **Note:** • It is not acceptable to treat "potential for Open Data to embarrass the government because it may reveal poor performance" as a risk under this heading • Any risks identified should be: ✓ Specific and clear, not general and vague ✓ Evidence-based ✓ Documented within the Data inventory with sufficient detail that external stakeholders will be able to understand the rationale and subject it to challenge.

If one or more of these harmful effects applies:

- Add the risk you have identified to the inventory
- Proceed to Step [D].

If none of the above negative effects apply,

- Classify the digital data as *Open* and add this classification to the inventory
- Move on to another digital dataset, or proceed to step [H].

2D. Weigh risk of harm against public interest

If harmful effects of publishing are identified in Step [C], then there is a presumption not to publish, but they are not absolute barriers to disclosure. In some instances, the public interest in publishing a digital dataset may outweigh the negative consequences.

Consider whether there is a high economic value or public interest in publishing the digital data. For example, would making the digital data open:

- Have significant economic benefits, e.g. could the digial data be used in the provision of new high-value services?
- Increase transparency of government spending or decision making?

If so, it should provisionally decide whether it would be reasonable and proportionate to publish the data, in spite of the negative effects identified in Step [C]. The final decision will lie with the Entity Overseeing Digital Data.

If you consider that the public interest outweighs the risk of harm:

- Record this in the inventory
- Classify the digital data as *Open* and add this classification to the inventory
- Move on to another digital dataset, or proceed to Step [H].

If the public interest does not outweigh the risk of harm:

- Proceed to Step [E] to classify the digital data as *Confidential, Sensitive or Secret*.

2E. Assess level of restriction

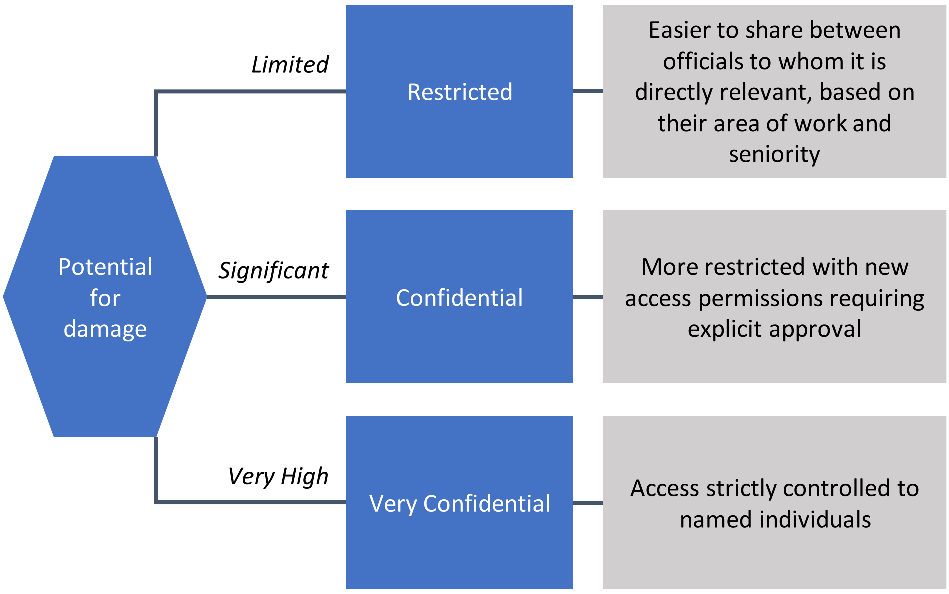
Where a dataset cannot be classified as *Open* after following Steps [A]-[D], it should be classified as *Confidential, Sensitive or Secret*, depending on the damage that would be risked by disclosure.

- Where there are No potential for damage, classify the digital data as ***Confidential***
- Where the potential for damage is **limited**, classify the digital data as ***Sensitive***
- Where the potential for damage is **very high**, classify the digital data as ***Secret***.

The **[DC1] Digital Data Classification Criteria** within the UAE Digital Data Interoperability Standards sets out the criteria to be applied when making this classification.

As illustrated below, this classification will affect who can see the digital data. *Confidential* data will be easier to share between officials to whom it is directly relevant, based on their area of work and seniority. *Sensitive* data is more restricted with new access permissions requiring explicit approval. *Secret* data will have access strictly controlled to named individuals.

Where a dataset cannot be classified as Public after following Steps A-D



Potential for damage

Limited

Restricted

Easier to share between officials to whom it is directly relevant, based on their area of work and seniority

Significant

Confidential

More restricted with new access permissions requiring explicit approval

Very High

Very Confidential

Access strictly controlled to named individuals

2F. Justify exclusions from the Open data inventory

By default, all *Confidential and Sensitive* datasets should be included in the published version of the Entity's Data Inventory. That is, it will be a matter of public record that the Entity holds the digital data, even though the digital data itself will not accessible except from authenticated and authorized users.

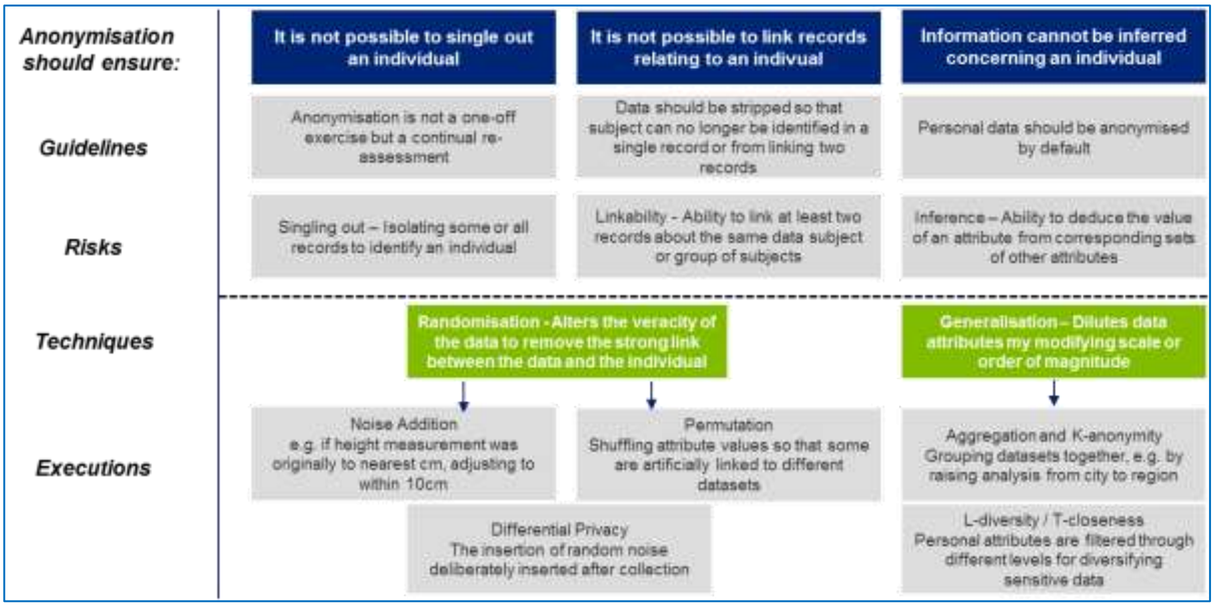
If an Entity wishes to make an exception to this, it should demonstrate that simply putting into the public domain the fact that the dataset exists (as opposed to the data itself) will cause negative impacts of the type considered in Step [C]. This decision should be agreed personally by the Entity's Director of Data.

2G. Consider whether a derivative dataset could be published

Where digital data has been categorized as *Confidential or sensitive* a balance needs to be stuck between the need for confidentiality and the benefits of openness. It may be possible to publish a summary, redacted version, extract, or other derivative of the data, which would have value as open data but avoid the negative effects identified at Step [C].

For example, personal data can be removed from a dataset through a range of anonymization techniques as illustrated below. In order to anonymize (or conceal identity in the information), and in accordance with the guidelines of the European Union, "Digital Data should be stripped of sufficient elements so that the author of such data cannot be identified." Specifically, data should be processed in such a way as to make it impossible to identify a natural person by using "all possible means and reasonable use." It should be borne in mind that the process of processing information to strip it of identifiable information is not reversible.

Anonymization Process (Hiding Personal Identity)



Anonymisation should ensure:	It is not possible to single out an individual	It is not possible to link records relating to an indivual	Information cannot be inferred concerning an individual
Guidelines	Anonymisation is not a one-off exercise but a continual re-assessment	Data should be stripped so that subject can no longer be identified in a single record or from linking two records	Personal data should be anonymised by default
Risks	Singling out – Isolating some or all records to identify an individual	Linkability - Ability to link at least two records about the same data subject or group of subjects	Inference – Ability to deduce the value of an attribute from corresponding sets of other attributes
Techniques	Randomisation - Alters the veracity of the data to remove the strong link between the data and the individual		Generalisation – Dilutes data attributes my modifying scale or order of magnitude
Executions	Noise Addition e.g. if height measurement was originally to nearest cm, adjusting to within 10cm	Permutation Shuffling attribute values so that some are artificially linked to different datasets	Aggregation and K-anonymity Grouping datasets together, e.g. by raising analysis from city to region
	Differential Privacy The insertion of random noise deliberately inserted after collection		L-diversity / T-closeness Personal attributes are filtered through different levels for diversifying sensitive data

Example: creating a derivative dataset

Consider a dataset of school students' educational results. The data would be of value in various ways: for example, to researchers looking at variation in educational achievement between different genders or different areas, or economic and social value through an app provided by a startup to help parents compare different schools. However, the dataset has been labelled as *Confidential* because the records include personal information about students and releasing the dataset would breach their privacy.

In this example, there are a number of ways that a derivative dataset could be prepared and published, depending on the details. It may be that simply anonymizing the records would be sufficient, as individual students could no longer be identified. If the data is very granular and specific, it may need to be aggregated or small number suppressed to ensure that individual results or performance can't be traced to particular people. In this case, results could be shown by year group, gender and school or with particular attributes / fields removed.

Data Custodians should therefore consider whether it would be possible to publish a modified version of the data. If there is the possibility of a derivative dataset that avoids the barriers and negative effects in Steps [B]-[C], or where the negatives are outweighed by public interest as in Step [D], then the Data Custodian should list a new derivative (or 'child') dataset in the Digital Data Inventory, noted as such and linked from the original dataset, but classified as *Open* Data.

This digital dataset should then also be catalogued following the rest of the Digital Data Conformance process.

There may also be cases where *Confidential* or *Very Confidential* datasets could be summarized or otherwise adapted in ways which, while still not allowing open publication, might enable less restrictive sharing across Government Entities. Again, if this is the case, then a new dataset should be created on the Data inventory, at the appropriate lower classification level.

2H. Add classification and documentation to digital dataset metadata, as part of the Digital Data Inventory

The classifications and supporting reasons process should be documented, in order to inform the Digital Data Inventory. The classification will form a mandatory part of the metadata for the digital dataset, along with the other elements specified in the **[DE2] Metadata** specification within the UAE Digital Data Interoperability Standards. For a smaller Entity this could be done in a standalone document or spreadsheet, but a large Entity with sufficient technical resources may wish to install their own data catalogue, allowing data custodians from each department to enter and edit metadata on the digital datasets for which they are responsible.

Steps 3 and 4: validation and review

Once all digital datasets have been fully catalogued, the classification and its supporting documentation (along with the rest of the Metadata and Format sample) will be collected by the Data Management Officer and reviewed. The Data Management Officer should then include all relevant results and metadata for those digital datasets in the Data Inventory for the Entity.

The resulting catalogued Inventory should then be reviewed internally by the Director of Data. The Director of Data should confirm that:

- Every digital dataset being catalogued in the current batch (as identified in the Prioritization process) has been classified correctly
- Where a digital dataset has been classified other than *Open*, a proper consideration has been given to whether a derived dataset could be recorded as *Open* or with a less *Confidential* classification (evidenced by the documented reasoning)
- The reasons for classifying any digital data as non-Open are documented in the Digital Data Inventory.

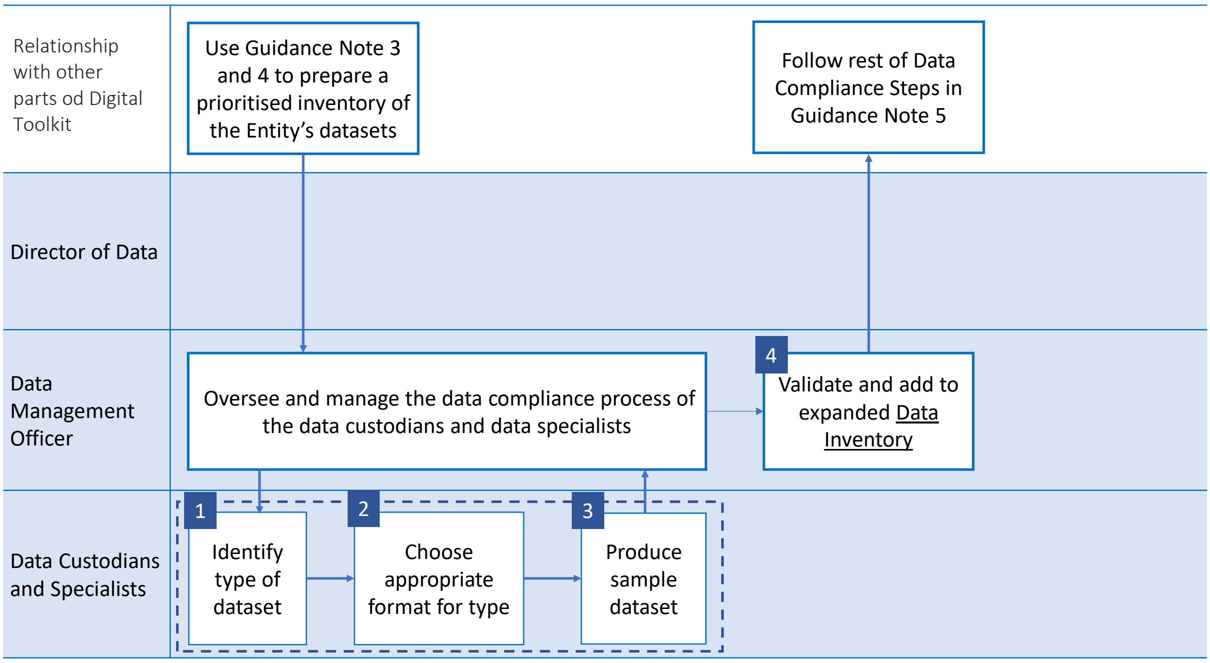
5.2 Formatting digital data

Purpose	This Guidance Note outlines the recommended process for ensuring that a digital dataset is correctly formatted in accordance with UAE Digital Data Interoperability Standard: **[DE1] Digital Data Formats**.
When to use	Before a digital dataset is published as open data or exchanged with other Entities, the data should be correctly formatted.
Responsibility	The Data Custodian is accountable for ensuring the correct formatting of dataset for which he or she is responsible, but may delegate responsibility for the work to a Data Specialist.

Process

Each digital dataset that is published openly or exchanged with other Entities by a Government Entity should comply with the **[DE2] Digital Data Formats** specification. To achieve this, we recommend the responsible Data Specialist should:

1. Identify the type of digital dataset that is being prepared for conformance
2. Choose an appropriate format to match that type
3. Produce a sample digital dataset which can be easily shared, shown and approved
4. Lastly: add the format to the digital dataset metadata and continue with conformance process


Relationship with other parts od Digital Toolkit
Use Guidance Note 3 and 4 to prepare a prioritised inventory of the Entity's datasets
Follow rest of Data Compliance Steps in Guidance Note 5
Director of Data
Data Management Officer
Oversee and manage the data compliance process of the data custodians and data specialists
4
Validate and add to expanded Data Inventory
Data Custodians and Specialists
1
Identify type of dataset
2
Choose appropriate format for type
3
Produce sample dataset

1. Identify type of digital dataset

The first step in choosing a digital data format is to determine what kind of data you are dealing with. Different types of data have different properties and need to be formatted in different ways.

Tabular digital data

Most government data are tabular data. If the digital data you are dealing with is a list, or would make sense to record in a spreadsheet then it is almost certainly tabular digital data.

Tabular digital data consists of rows, each of which is an individual **record** in the digital dataset, and columns, each of which represents one **field** of the record. For instance, a dataset about schools might be:

Unique Id	Name	Highest age	Lowest age
AB292	Blue Water High	11	5
HG383	Green Tree Academy	11	5

The second row containing "Blue Water High" is the **record** about Blue Water High, and the column titled "Highest age" contains the digital data from the highest age **field** about each school.

Geospatial digital data

Geospatial data relates to information about how you would draw things on a map.

We know that digital data is geospatial when:

- It contains the coordinates used to point to something on a map - for instance a latitude and longitude pair - for example the location of parking spaces, or public libraries.
- It contains the shape that we would draw onto a map to represent a particular area. For instance: data about the catchment area for a school; the boundaries of an electoral district; administrative regions for school districts; or zoning areas for planning permission.

Real time or service digital data

Real time digital data is generally provided immediately via an API (Application Programming Interface) that can be consumed by other software applications. Digital Data is real time if it changes so frequently that most questions you would ask about it would be quickly out of date.

One example would be the status of trains on a rail network, or information about current flights - departures, arrivals and delays at an airport.

Digital Data being provided to power real-time services which frequently access or need to update records automatically should also be provided via an API.

Structured non-tabular digital data

Some digital data is structured, but does not fit into a tabular form in a natural manner. If your digital data is hierarchical or contains many levels, then it is likely structured non-tabular digital data. Examples would include the organization chart for your department, or a project plan.

2. *Choose an appropriate format*

Once you have identified the type of digital data you are dealing with in any specific dataset, use the following table sets out format requirements for common types of data to use:

Digital Data type	Mandatory	Recommended
Tabular data	CSV	CSV and Excel file with definitions and commentary
Geospatial data (the coordinates and information about the point)	CSV	
Geospatial data (shapes of areas)	GeoJSON or KML	GeoJSON and KML
Real time data or data used for responsive services	Via an API	
Structured non-tabular	An appropriate open, machine readable format, conforming to an open standard where available. E.g. JSON, XML, RDF, GTFS.	
Unstructured data		Using an open format where such exists

In addition to the above generic criteria, there are format-specific criteria detailed below.

Digital Data Format	Conformance requirements
CSV data	The format of a CSV dataset will be conformant if: • It contains a header row which includes the name of the column • The formatting of dates or numbers is consistent throughout the whole file • It does not include empty rows • It does not include rows with missing or extra cells • It does not use header names more than once in the same file • It does not include any commentary or explanatory text
Structured non tabular data	The format of structured non tabular data is conformant if: • It conforms to a pre-existing open standard for representing such data, such as GTFS, Popolo, the Schema.org job posting standard or b) It is in a valid open machine-readable standard such as JSON, XML and: • The structure of this data is clearly documented and published alongside it • The structure of the data is appropriate for re-use given the nature of the domain to which the data relates.
Geospatial data	The format of a geospatial dataset will be conformant if: • It is published in valid GeoJSON or KML (see Digital Data Interoperability Implementation Guide for further advice and recommendations on validating conformance).

Real-time and service data	An API is conformant if the API endpoint and API documentation is available. API documentation should include: • Clear reference information providing the functions, remote call and methods for the API • Guidance to help developers experiment with the API • Information about security, versioning and rate limiting so users can plan their commitment to using the API Entities may provide an API to their data in addition to publishing or exchanging the data in one of the other formats.

Alternative formats

If you want to use an alternative format not listed above, you should have a clear reason for why that is the most appropriate format to publish in. You should also check you've selected the most open, standardized and machine-readable format available to meet the requirements.

3. Produce a sample digital dataset and check it's well formatted

Once you have determined the format, the next stage is to produce a sample dataset.

This is likely to be a file, or collection of files, that represent what you would publish each time that the digital dataset is updated. It could be the full digital data file or some of the data. If the digital data is constantly updating it will be a particular slice of the digital data.

Ensure your sample is well formatted. The table below recommends tools and guides for avoiding common formatting mistakes:

Digital Data Format	Tools
CSV data	• Linting tools (programs that analyze code, data, etc. for potential errors) such as http://csvlint.io/ are available for CSV data and can be used to identify errors with files early. • If a field contains a comma, a line ending or a double quote then the field is escaped by wrapping it in double quotes. Within a field that is escaped like that, any double quotes are doubled up.
Geospatial data	• Linting tools are available for GeoJson such as http://geojsonlint.com/ which will help catch errors in your data files. For KML it is possible to validate your data against the KML Schema (https://developers.google.com/kml/schema/kml21.xsd?csw=1) • High quality open tools exist to convert geospatial data between formats, and can be included in automated dataset generation pipelines to easily publish in multiple formats. • One good example is Ogr2Ogr http://www.gdal.org/ogr2ogr.html
Structured non-tabular data	• For many types of common dataset there exist open standards for representing that information as structured data which should be re-used as much as is possible. • Examples of such standards include: – Schemas found on http://schema.org/ – The Popolo data standard for people, organizations and voting http://www.popoloproject.com/

	• Non tabular structured data should in general use JSON, unless there is a clear reason to use an alternative format, such as a common standard in an alternative format (e.g. GTFS for transport data)
Real-time data and APIs	• APIs should be designed to meet the requirements for your use-case and with privacy and security built in. Where possible, ensure data minimization – giving access to the smallest amount of information required for the service outcome or to enable a decision. For example, sending 'yes', 'no' or 'not found' in response to a query of whether a citizen or user is over 18 or has a valid driving license instead of sending personal information. • Guidance on good practice when designing and documenting APIs can be found in here: – UK Government Service Manual – US White House API standards • An example of data API documentation for the UK Government Registers is here.

The sample is representative of what will be available for users. Its purpose it to help the Digital Data Custodian, Data Management Officer, Director of Data and the Entity Overseeing Digital Data see that the data:

- Conforms to the Digital Data Interoperability Framework standards
- Makes sense as a dataset

4. *Continue with digital data conformance process*

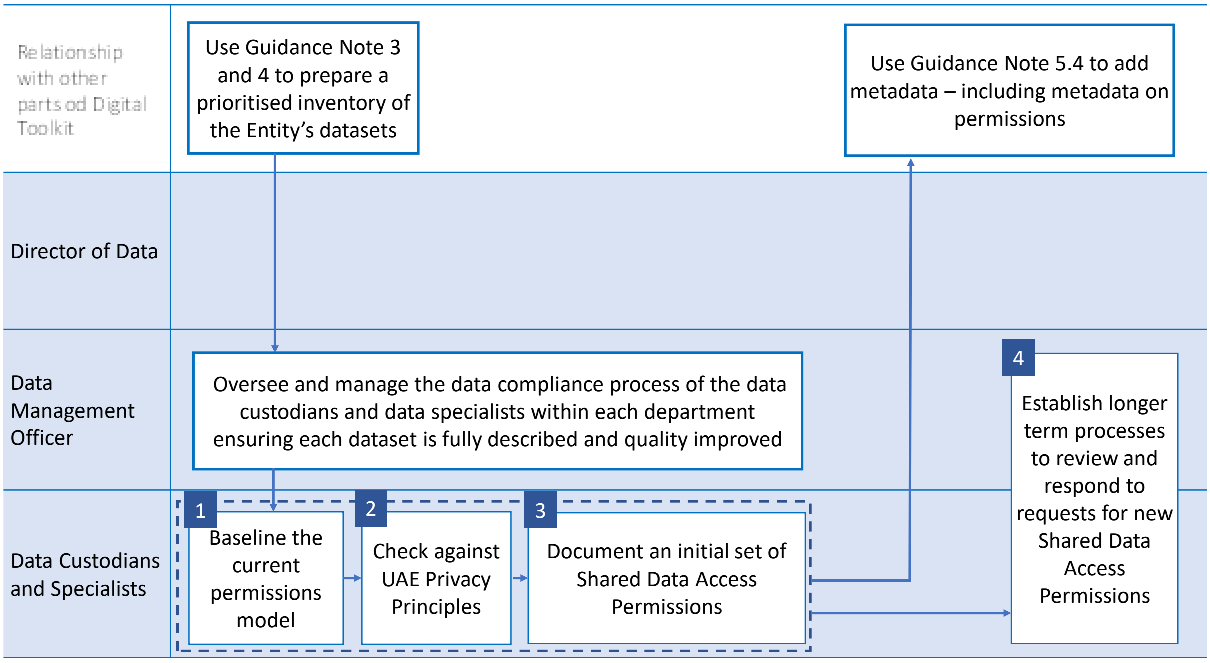
Record the format and add the sample to the expanded inventory and add the chosen format to the appropriate metadata field. This ought to be reviewed by the Data Custodian responsible for the digital data and Data Management Officer.

5.3 Documenting a permissions model for shared digital data

Purpose	This Guidance Note outlines the recommended process for determining who may access a digital dataset and with what level of access in conformance with the **[DE7] Shared digital data access permissions** specification.
When to use	When preparing Confidential or Sensitive digital data for exchange with another Entity for the first time. Also when responding to future requests for additional access permissions.
Responsibility	Data Custodian.

Process

For each digital dataset in the current prioritized batch being catalogued, the responsible Data Custodians need to ensure that the requirements of the **[DE7] Shared digital data access permissions** specification are met in respect of Confidential and Sensitive Data. The following process is recommended.


Relationship with other parts od Digital Toolkit
Director of Data
Data Management Officer
Data Custodians and Specialists
Use Guidance Note 3 and 4 to prepare a prioritised inventory of the Entity's datasets
Use Guidance Note 5.4 to add metadata – including metadata on permissions
Oversee and manage the data compliance process of the data custodians and data specialists within each department ensuring each dataset is fully described and quality improved
1
Baseline the current permissions model
2
Check against UAE Privacy Principles
3
Document an initial set of Shared Data Access Permissions
4
Establish longer term processes to review and respond to requests for new Shared Data Access Permissions

1. Baseline the current permissions model

Your dataset will already have a set of permissions associated with it, even if this is simply current practice rather than a documented policy. So the starting point is to document the baseline position around who is permitted to access the data:

- Who currently has access to the dataset within the Entity?
- Are these specific named individuals only, or groups of individuals (e.g. people in a specific business unit of the Entity, or people in a specific grade or function)
- Are there any restrictions on the purposes for which these people may access the digital data?
- What kind of access do they have? Is it full access to the digital dataset, access to specific entries or records, or ability to edit/update records?

- Similarly, are there individuals or groups of people in other organisations who are permitted access to the digital data, and on what basis?

2. Check against UAE Privacy Principles

The **[DE6] Digital Data protection and privacy** specification sets out a set of UAE Privacy Principles which all Government Entities should seek to apply when managing digital data that contains personal data or commercial data. Having documented current practices around granting access permissions to data, you should check that those practices are compliant with these Privacy Principles. Particular issues to consider are:

- Does the Entity have the consent of the digital data subject to share their digital data with all those people who currently access it?
- If not, is the dataset covered by sector specific regulations which mean that such consent is not required?
- Are there controls in place to ensure that people permitted access to the data may only use it for specified business purposes?
- Is the level of access proportionate to the stated purpose? (For example, if an official has a business need to check whether an individual is over 18, they should be permitted yes/no query access to the digital data rather than being able to see the digital date of birth of the individual.)

Entities should embed the following UAE Digital Data Privacy Principles in their data management practices, and in those of third parties contracted to manage digital data and services on their behalf.

Digital Data Privacy Principles		**Description**
1.	**Consent**	• Personal Digital Data in relation to individuals and Commercial Data in relation to Private Entities should not be disclosed or shared without the data subject's consent. • When providing a service to an individual or a Private Entity, Government Entities should seek the consent of that digital data subject for the data to be exchanged with other Government Entities for the purpose of enabling any Government Entity to provide services to the data subject without the need for the digital data subject to provide the same information again.
2.	**Transparency**	• Digital Data subjects should be informed - at the point of digital data collection - when and by whom their digital digital data is being collected, why it is needed, and how it will be used.
3.	**Purpose**	• Data should only be used for limited and explicitly stated purposes and not for any other purposes without first gaining informed consent from the digital data subject.
4	**Proportionality**	• When data is requested and stored, the type of data collected should be the minimum required to carry out the stated purpose, individual users of the digital data should only be given the minimum access to that data that they need, and the digital data should not be kept for longer than is necessary for that purpose.
5	**Personal access and control**	• Digital Data subjects should be enabled to: – Access and take copies of data that is held about them – Correct inaccuracies in data that is held about them

		− Request removal of data that is held about them, but is no longer relevant or applicable to the business of the Entity
6.	**Security**	• Collected digital data should be protected by robust and tested security safeguards (technical and organizational) against such risks as loss and unauthorized access, destruction, use, modification or disclosure. • To help achieve this, Government Entities should − Apply the latest version of the UAE's information security standard − Ensure compliance with the Payment Card Industry (PCI) Security Standards for information systems that store or process credit card data − Ensure that cloud suppliers to the Government Entity meet ISO/IEC 27017 Cloud Security Standards and ISO/IEC 27018 Handling of Personally Identifiable Information Standards. − Notify the Entity Overseeing Digital Data in the event of any perceived conflict between other provisions of this Standard and the standards listed in the three bullets above..
7.	**Sectoral compliance**	• Each sector has its own laws and regulations, some of which relevant to the basis on which digital data can be shared with other entities or with public. Examples of these laws include the United Arab Emirates Penal Code, the Copyrights' Act, and the Telecommunications' Act. • Entities should ensure they comply with both relevant sectoral regulations and this Standard, and should notify the Entity Overseeing Digital Data in the event of any perceived conflict.
8.	**Documentation**	• Entities should document who is permitted to access each digital data set, either in the form of the **[DE5] Open Data License** (for all Open Data) or through a documented set of **[DE7] Shared Digital Data Access Permissions**. • Entities should produce and maintain privacy metadata in relation to these access permissions, as part of their broader work on **[DE2] Metadata**, and store this in their Digital Data Inventory.
9.	**Awareness**	• Entities should develop an awareness programme for their digital data privacy policy, which shall be disseminated to all staff within the Entity who manage digital data (both from business and technical areas) in order to remind them of the Entity's obligations and their personal responsibilities concerning digital data privacy.
10	**Accountability**	• Entities should establish and publicise effective complaints and redress mechanisms for data subjects who believe they are failing to manage their digital data in accordance with the above principles.

3. Document an initial set of Shared Digital Data Access Permissions

Develop a documented set of initial Shared Digital Data Access Permissions. Normally, this will simply codify the existing data sharing practices that are in place for the dataset - perhaps modified following the privacy conformance assessment at Step 2. These Shared Digital Data Access Permissions should cover:

- **Who may have access to the shared digital data**. These permissions may be given to either:
  - Named individuals
  - One or more classes of individuals, such as government employees:
    - In a specific professional function (such as finance, HR, operations, IT)

        - In a specific grade
        - In specific positions (such as Head of Finance)
        - In specific Entities, or departments within Entities
        - With specific levels of security clearance
    - A combination of the two.

- **What purpose this access is for**. This documentation is particularly important to ensure conformance with the 'purpose' and 'proportionality' principles of **[DE6] Digital Data protection and privacy** and to enable effective auditing.

- **The level of access that they may have**. These permissions (which may be different for different digital data users) should specify whether access to the dataset is permitted as:
    - Query-only access
    - Read-only access
    - Read-write full access.

Key elements of this documentation should then be codified in the metadata for the digital dataset – see **Guidance Note 5.4: Adding metadata and schema** – as the digital dataset moves to the next stage of the Digital Data Conformance process.

Government Entities should embed the following Access Permission Principles in their digital data management practices,

Access Permission Principles	**Description**
1. Entities should facilitate cross-government data sharing of their digital data	• Access to shared data shall be approved by the Government Entity responsible for that digital data. • However, data ownership does not mean the monopoly of digital data by any Government Entity, or entitle it to obstruct the reasonable needs of other parties to access that data in pursuit of their legitimate functions. This means that: – Whenever a Government Entity wishes to use data that is owned and managed by another Government Entity, the digital data-owning Government Entity has a duty to respond rapidly and positively to that request – Data-owning Entities have a duty to invest in systems and process which facilitate rapid, effective and secure data-sharing – in particular in respect of digital datasets that have been identified as Primary Registries – Government Entities may not charge other Government Entities to access their Shared Digital Data.
2. Use of the Federal Digital Data Platform	• Wherever possible, Government Entities should exchange digital data via the Federal Digital Data Platform, or Emirate-level electronic platforms that are securely inter-connected with the Federal Digital Data Platform. For Federal Government Entities use of the Federal Digital Data Platform is mandatory, and exceptions to this require prior written approval from the Entity Overseeing Digital Data.
3. Digital Data Sharing Access Permissions should be documented	• Government Entities that share and exchange non-open data with other Entities should document (and record within their **[DE2] Metadata**): – **Who may have access** to the shared digital data. These permissions may be given to either: ▪ Named individuals ▪ One or more classes of individuals, such as government employees: ▪ In a specific professional function (such as finance, HR, operations, IT)

		▪ In a specific grade ▪ In specific positions (such as Head of Finance) ▪ In specific Entities, or departments within Entities ▪ With specific levels of security clearance ▪ A combination of the two. – **What purpose this access is for.** This documentation is important to ensure compliance with the 'purpose' and 'proportionality' principles of **[DE6] Digital Data protection and privacy** and to enable effective auditing. – **The level of access that they may have.** These permissions (which may be different for different data users) should specify whether access to the dataset is permitted as: ▪ Query-only access ▪ Read-only access ▪ Read-write full access. • For most digital data users, query-only access that returns the minimum necessary information will be sufficient for their business purposes. Access permissions should therefore be designed to give access to the smallest amount of information required for the service outcome or to enable a decision. (For example, sending 'yes', 'no' or 'not found' in response to a query of whether a citizen or user is over 18 or has a valid driving license instead of sending personal information.)
4	**Access to shared digital data should be secured and audited**	• Entities should establish systems to ensure that: – A shared digital dataset can only be accessed by identified individuals, who have been appropriately authenticated as being permitted such access under the terms of the Digital Data Sharing Access Permissions – All Shared Digital Data access via electronic platforms should store an audit log of what data was accessed, when and by whom.

Mandatory actions

In applying these principles, each Government Enterprise should:

- **Develop a detailed Digital Data Sharing Plan, setting out how they will implement the Digital Data Exchange Principles** described in this Standard**,** including any investments in systems and processes that they will need. They should share this Plan with the Entity Overseeing Digital Data.
- **Respond in writing within a reasonable time** to requests for digital data sharing from other Entities, giving either:
  - Agreement to the request, and a clear timetable for implementation
  - A refusal of the request, accompanied by a clear rationale for the refusal that is rooted in the principles of this Standard.
- **Notify the Entity Overseeing Digital Data** of all requests from other Entities for sharing and exchange of *Confidential* or *Sensitive* Data. This includes requests both for access to a digital dataset which is currently not shared, and requests to add new individuals or classes of individual to the Shared Data Access Permissions for a digital dataset that is already being shared across Entities. Notification should be made as follows:

  - **Approved Confidential:** For *Confidential* data sharing requests from other Entities which the data-owning Entity approves, it may notify the Entity Overseeing Digital Data after the event, for example by giving a quarterly update on all data-sharing initiatives it has approved.

- **Refused Confidential and Sensitive for** digital data sharing requests which the data-owning Entity proposes to refuse, it should inform the Entity Overseeing Digital Data at the same time as declining the requesting Entity, documenting its rationale for declining the request. The office has the power to issue binding decisions to change data access decisions in the cases of dispute between Government Entities or with third parties.
- **Approved Confidential:** Given the extra sensitivity of such data, where a data-owning Entity believes that sharing *Confidential* Data with another Entity is in the public interest and follows the principles of both this Standard and **DE6] Digital Data protection and privacy**, it should consult the Entity Overseeing Digital Data before giving approval.

Recommended actions

When implementing Access Permission Principle 5 (“Access to shared data should be secured and audited”), Government Entities are recommended to make this audit functionality openly available for use by individual digital data subjects. This means:

- Configuring digital platforms and supporting business processes so that individual digital data subjects (citizens, residents and businesses) can see an audit trail of who accessed their digital data, and for which documented purpose (excluding security service or law enforcement access)

Providing mechanisms by which digital data subjects can raise concerns / escalate if they believe access has been misused.

4. Establish longer-term processes to modify Shared Digital Data Access Permissions

In addition, the Data Custodian should work with the Data Management Officer to establish longer term processes to review and respond to new requests in future for digital data access – for example, from other business units within the Entity or from other Government Entities who wish to access the digital data for new business purposes that were not envisaged within the initial Shared Digital Data Access Permissions.

The service standard set out in the **[DE7] Shared Digital Data Access Permissions** specification is that Government Entities should respond to such requests within rreasonable period, giving either:

- Agreement to the data sharing request, and a clear timetable for implementation
- A refusal of the request, accompanied by a clear rationale for the refusal that is rooted in the principles of the **[DE7] Shared Digital Data Access Permissions** specification.

In certain cases, the Specification requires this response to be agreed with the Entity Overseeing Digital Data.

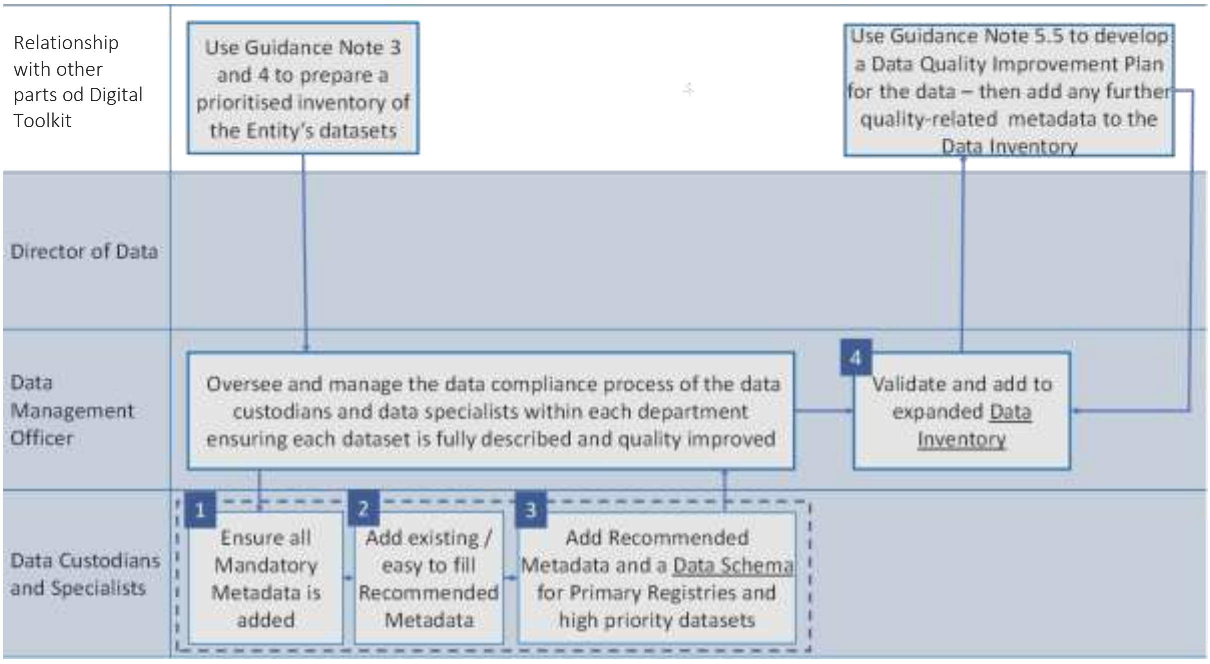
In the early stages of sharing a digital dataset with other Entities, such responses will be managed on a case-by-case basis. As the Entity develops more experience of assessing the privacy implications of data sharing, it will increasingly want to codify that experience into a set of rules that can be simplified and automated -to speed up access permission management in a risk-based way

5.4 Adding metadata and schema

Purpose	This Guidance Note outlines the recommended process for ensuring that a dataset has all required metadata in accordance with UAE Digital Data Interoperability Standard: **[DE2] Metadata** and **[DE3] Schema**.
When to use	Before a dataset is published as open data or exchanged with other Entities, the data should be appropriately described so that it is discoverable and users understand its reliability.
Responsibility	The Data Custodian is accountable for ensuring the dataset contains all required metadata. The Data Specialist is responsible for providing a data Schema if applicable.

Process

For each digital dataset in the current prioritized batch being catalogued, the responsible Data Custodians and Data Specialists need to ensure the Metadata and Schema and Digital Data Quality requirements are met. The following process is recommended.


Relationship with other parts od Digital Toolkit
Use Guidance Note 3 and 4 to prepare a prioritised inventory of the Entity's datasets
Use Guidance Note 5.5 to develop a Data Quality Improvement Plan for the data – then add any further quality-related metadata to the Data Inventory
Director of Data
Data Management Officer
Oversee and manage the data compliance process of the data custodians and data specialists within each department ensuring each dataset is fully described and quality improved
4
Validate and add to expanded Data Inventory
Data Custodians and Specialists
1
Ensure all Mandatory Metadata is added
2
Add existing / easy to fill Recommended Metadata
3
Add Recommended Metadata and a Data Schema for Primary Registries and high priority datasets

1. Ensure all Mandatory Metadata is complete for the dataset – this should include the title, description, subject, format, size, publisher, custodian, classification, access permissions, license, coverage (temporal and geospatial) as well as the data files and last updated timestamp.

	Metadata field	Definition / description	Requirement level
Discoverability	**Title**	Brief descriptive name for the digital dataset. Should communicate subject and scope.	Mandatory
	Description	A description of the digital dataset. This could provide more detail about what the data contains and what it's about, how and why is was collected, any known errors or limitations. Ideally the description covers all the relevant context that would be useful for users to help them decide if this digital data is fit for their purpose.	Mandatory
	Subject	The top level theme or category for the digital data. For example: health, transport, business, education. This should be a pre-defined taxonomy or vocabulary that is common across UAE. It could have one level or include sub categorizations.	Mandatory
	Tags	Keywords related to this dataset, such as: 'schools', 'location', 'class sizes', 'parking'.	Recommended
Technical information	**Data files**	Links to or uploads of the data relevant to this dataset. Might be in multiple formats (for example as CSV and Excel). If providing an API, ensure the API endpoint and API documentation is linked to or uploaded.	Mandatory
	Format	Describes the technical format in which the data is currently held (e.g. CSV, GeoJSON) – See **[DE1] Formats** for more guidance. May be auto-filled if publishing in a catalogue depending on ingestion method.	Mandatory
	Size of the dataset	Size of the dataset files (in MB, kB, etc.) If using a platform to publish the data, this can be configured to display automatically.	Mandatory
	Schema	The schema defines the parameters of each attribute in the data. This should describe the attributes, clarify whether each is required, the type (string, number, date), the vocabularies used (if any) and so on. More detail can be found on creating a schema in **[DE3] Data Schema**.	Recommended
	Last Updated	Timestamp of when this dataset was last updated. If using a platform to publish the data, this can be configured to display automatically.	Mandatory
	Unique identifier (URI)	Each digital dataset published or exchanged through an electronic platform or on the web should have a unique identifier. Ideally this would be a public identifier (such as a URI).	Recommended

Source	**Publisher**	The name of the Entity that owns the dataset. This should be in the format “Entity, Business Unit”, for example “TRA, Wireless Networks & Service Section”.	Mandatory
	Custodian	Name of the Data Custodian responsible for this dataset.	Mandatory
	Contact information	The email address or web form that should be used to contact the Entity for queries, feedback or requests concerning this dataset	Recommended
	Source system	Name of the source system (upstream database) this data comes from, if applicable.	Recommended
Applicability	**Classification**	The Data Classification of the dataset, as defined in the Classification Standard. Should be one of: Open, Confidential, Sensitive or Secret	Mandatory
	Temporal Coverage start date	Indicates the earliest digital date that the data in this dataset relates to. Should use ISO 8601 date format.	Mandatory
	Temporal Coverage end date	Indicates the latest date that the digital data in this dataset relates to. Should use ISO 8601 date format.	Mandatory
	Geographic coverage	Region covered by this digital data – for instance, the name of a city, district, council or country.	Mandatory
	Language	Language used in the dataset. Should use ISO 639 codes.	Mandatory
Access	**License**	Link or copy of the license terms under which the data may be used. By default, this should use **[DE4] Open data licensing** for Open Data.	Mandatory
	Access Permissions	Documentation of who has access to this dataset and the level of access they have, as described in the **[DE7] Shared Digital Data Access Permissions** standard.	Mandatory
	Personal data?	Does this digital data contain any personal digital data? Yes/No. Personal digital data means any information relating to an identified or identifiable natural person; an identifiable person is one who can be identified, directly or indirectly, in particular by reference to **an identifier** such as a name, an identification number, **location data**, online identifier or to one or more factors specific to the physical, physiological, **genetic**, mental, economic, cultural or social identity of that person.	Recommended
	Sensitive personal digital data?	Does this digital data contain Sensitive personal data? Yes/No.	Recommended

		Sensitive personal data are personal data that directly or indirectly reveal an individual's family, racial or ethnic origin, communal origin, political opinions, affiliations, religious or philosophical beliefs, their union membership, criminal record, health, sexual life, genetic data or biometric data.	
	Intellectual Property	Description of any IP contained in the digital data and conditions/ rights of use and distribution.	Recommended
Reliability	**Provenance**	Details of how thedigital data was collected, processed, redacted or amended. Digital Data provenance documents the sources, inputs, organizations, systems, and processes that have formed and influenced the digital data, in effect providing a historical record of the digital data and its origins. This allows data-dependency analysis, awareness of limitations and coverage, error detection and auditing. It helps other users (including the Entity) understand the limitations and level of trust they can place in the digital data. The Entity should aim to standardize its provenance descriptions over time and provide it in a machine readable method – for example by using the World Wide Web Consortium standard for data provenance. Example of what to include in methodology: • Where the digital data came from (survey, third party etc.) • Sample size (if survey) • Data collection method (face-to-face interviews, online, requests from authorities) • Exclusions (what data what not included and why) • Statistical aggregation methods used (small number suppression, averaging, etc.) The office of National Statistics in the UK regularly publishes provenance and methodology data which can be reviewed for a real data example.	Recommended
	Publishing Frequency	The rate at which the digital data in the dataset will be updated. Responses should correspond to a value contained in the Dublin Core Collection Description Frequency Vocabulary (that describes frequency periods from “triennial” through to “continuous”). Updates are expected to be additional digital data files following the same schema, but new temporal coverage (e.g. latest month). To ensure good practice, the Data Custodian should: 1. decide on a publication schedule appropriate to the digital dataset	Recommended

		2. create a publication calendar for that digital dataset 3. establish a specific role or individual with responsibility for publishing the dataset on that date This allows users build reliable processes, tools and services using the digitlized data.	
	Known issues	Description of any known errors or limitations of the digital data. For example if there was unreliable data collection or if particular fields are unvalidated and rely on the data subject self-reporting.	Recommended
	Digital Data completeness	Description of any known gaps in coverage of the data. Are there missing geographic areas or time periods for which there is no digital data?	Recommended

2. Add any existing or easy to fill Recommended Metadata fields out of tags, schema, unique ID, contact information, source system, provenance, publishing frequency, known issues and data completeness as well as details on whether the data contains personal or sensitive personal digital data or intellectual property and associated terms of use.
3. Add all Recommended Metadata and a digital data Schema, using the relevant standards, for Primary Registries datasets or high priority (as defined in the **Guidance Note 4: Prioritization Criteria and Process**), structured and regularly updated datasets. These might be datasets needed for cross-entity service delivery projects or which have been frequently requested by users and deliver on strategic objectives.

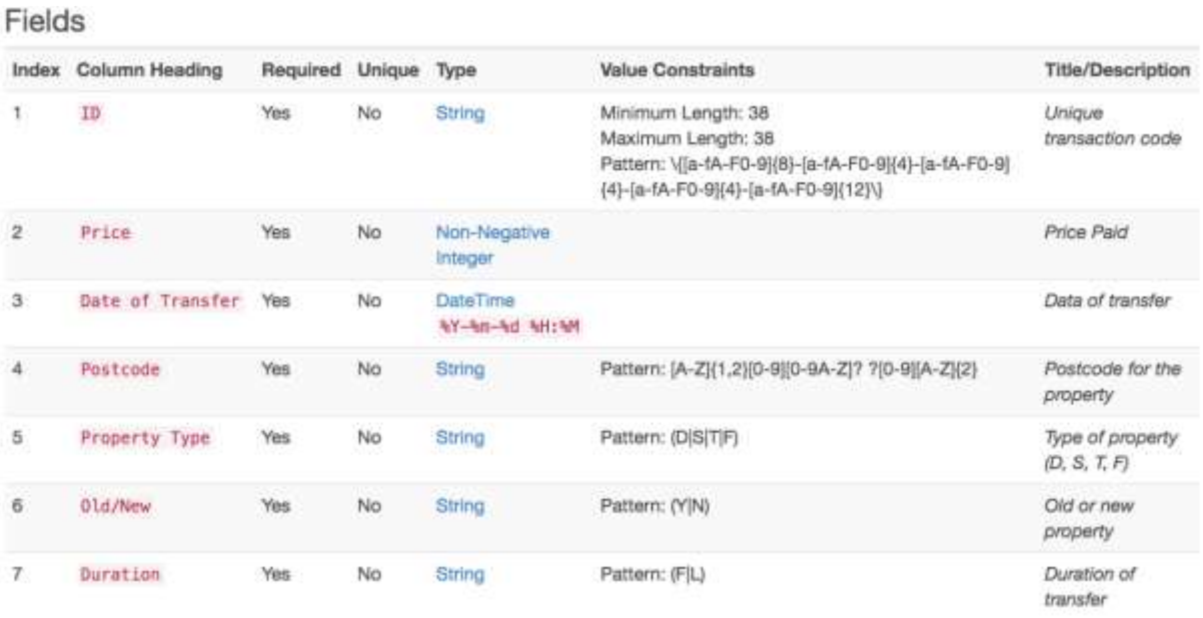
Establishing a schema allows for automatic validation of digital datasets, as well as making it significantly easier for third parties to build tooling around digital data and re-use it.

SQL databases will already have a schema in place, although Entities may want to ensure these are fit for purpose by modelling the data to be stored and deciding on the relationships, vocabularies, validation and range(s) to be applied.

Structured non-tabular (e.g. JSON) data should provide a schema in JSON Schema format according to the specification here: http://json-schema.org/.

Tabular (e.g. CSV) data should be expressed as a JSON Table Schema according to the open specification here: https://frictionlessdata.io/specs/table-schema/.

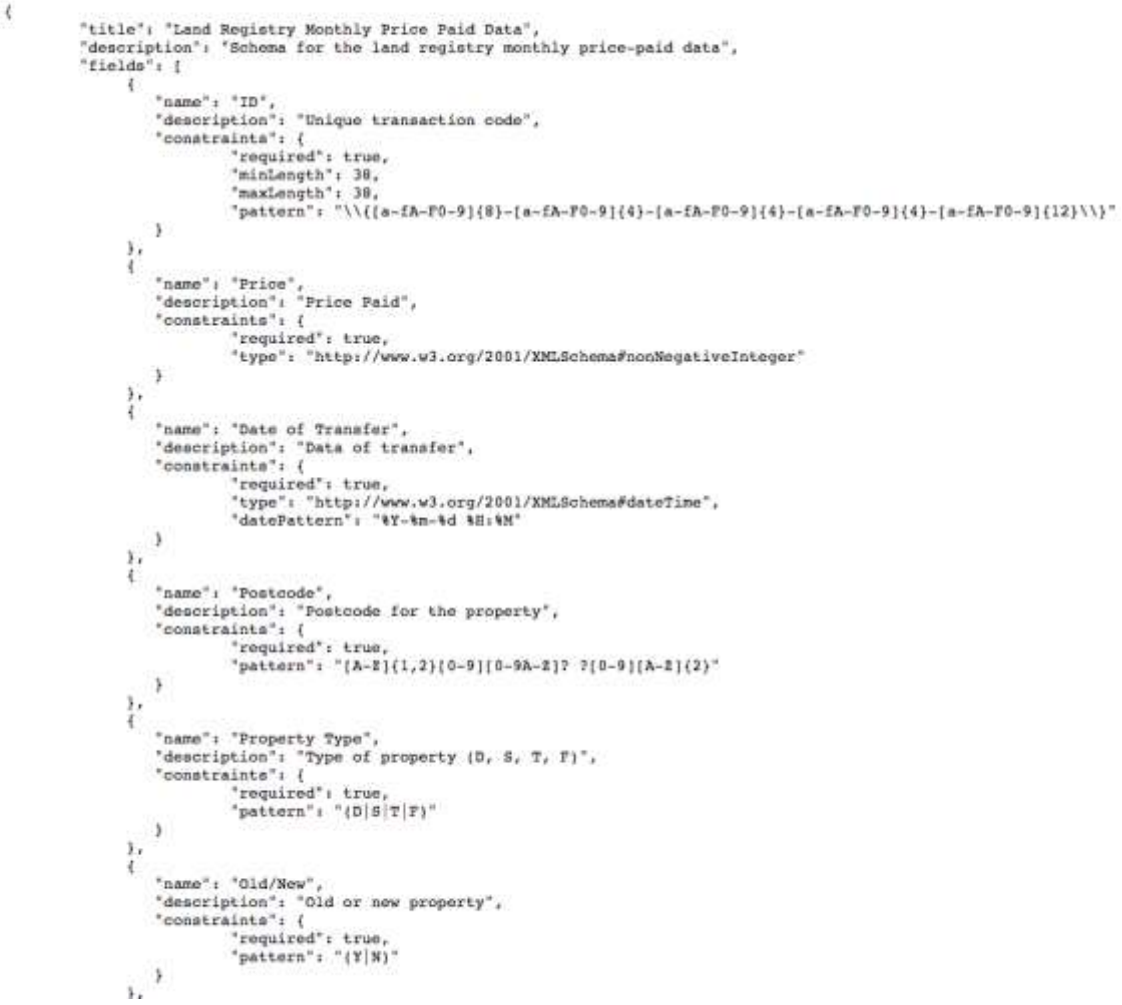
A human-readable version of a schema looks like:



Fields

Index	Column Heading	Required	Unique	Type	Value Constraints	Title/Description
1	ID	Yes	No	String	Minimum Length: 38 Maximum Length: 38 Pattern: \{[a-fA-F0-9]{8}-[a-fA-F0-9]{4}-[a-fA-F0-9]{4}-[a-fA-F0-9]{4}-[a-fA-F0-9]{12}\}	*Unique transaction code*
2	Price	Yes	No	Non-Negative Integer		*Price Paid*
3	Date of Transfer	Yes	No	DateTime %Y-%m-%d %H:%M		*Data of transfer*
4	Postcode	Yes	No	String	Pattern: [A-Z]{1,2}[0-9][0-9A-Z]? ?[0-9][A-Z]{2}	*Postcode for the property*
5	Property Type	Yes	No	String	Pattern: (D|S|T|F)	*Type of property (D, S, T, F)*
6	Old/New	Yes	No	String	Pattern: (Y|N)	*Old or new property*
7	Duration	Yes	No	String	Pattern: (F|L)	*Duration of transfer*

A machine-readable version in JSON of this schema looks like:



```
{
        "title": "Land Registry Monthly Price Paid Data",
        "description": "Schema for the land registry monthly price-paid data",
        "fields": [
             {
               "name": "ID",
               "description": "Unique transaction code",
               "constraints": {
                        "required": true,
                        "minLength": 38,
                        "maxLength": 38,
                        "pattern": "\\{[a-fA-F0-9]{8}-[a-fA-F0-9]{4}-[a-fA-F0-9]{4}-[a-fA-F0-9]{4}-[a-fA-F0-9]{12}\\}"
               }
             },
             {
               "name": "Price",
               "description": "Price Paid",
               "constraints": {
                        "required": true,
                        "type": "http://www.w3.org/2001/XMLSchema#nonNegativeInteger"
               }
             },
             {
               "name": "Date of Transfer",
               "description": "Data of transfer",
               "constraints": {
                        "required": true,
                        "type": "http://www.w3.org/2001/XMLSchema#dateTime",
                        "datePattern": "%Y-%m-%d %H:%M"
               }
             },
             {
               "name": "Postcode",
               "description": "Postcode for the property",
               "constraints": {
                        "required": true,
                        "pattern": "[A-Z]{1,2}[0-9][0-9A-Z]? ?[0-9][A-Z]{2}"
               }
             },
             {
               "name": "Property Type",
               "description": "Type of property (D, S, T, F)",
               "constraints": {
                        "required": true,
                        "pattern": "(D|S|T|F)"
               }
             },
             {
               "name": "Old/New",
               "description": "Old or new property",
               "constraints": {
                        "required": true,
                        "pattern": "(Y|N)"
               }
             },
```

Source: http://csvlint.io/schemas/530b16c163737676e9260000

Having completed this process, proceed to assess whether the dataset meets the requirements of the Digital **Data Quality Standard**, following the process described below in **Guidance Note 5.3 Managing digital data quality**. If Guidance Notes 5.1, 5.2, 5.3 and 5.4 have been followed properly, then all the mandatory quality requirements will already now be met. But there may be additional steps you take in improving data quality which will generate additional metadata requirements.

5.5 Managing digital data quality

Purpose	This Guidance Note provides Entities with guidance on the process and steps for improving and managing data quality over time in a structured, prioritised and appropriate to the intended and potential use of the Entity's digital data assets. It covers both: • initial steps to meet minimum quality requirements ahead of initial publication of open data or initial exchange of shared digital data • longer term actions to drive forward digital data quality
When to use	At the start of each Entity's Digital Data Interoperability program
Responsibility	• Data Custodians at the level of individual datasets • Data Management Officer for Entity as a whole

The Entity-level context for data quality

Managing digital data quality should be a strategic priority for the Entity as a whole. Having discoverable, reliable, trusted and well managed digital data enables the Entity to be more efficient, effective and accountable. It allows for the automation of processes as well as enabling better service delivery and decision making. The **[DQ3] Digital Data Quality Improvement Plan** specification requires all Government Entities to develop a Digital Data Quality Plan. Guidance on how to do this at an Entity-wide level is given in:

- **Guidance Note 1: Establishing digital data governance roles and processes**, which gives advice on specific roles which will have key responsibilities for data quality and suggest there needs to be at minimum:
  - A full time dedicated role of a Data Management Officer or equivalent who day-to-day manages and tracks the quality of the Entity's digital data
  - Data Custodians and Data Specialists who own and are responsible for the digital data quality of the digital data assets they manage, undertaking digital data quality assessments and data cleansing
  - A Director of Data to set the strategic direction and requirements for digital data quality management across the Entity
- **Guidance Note 2: Building a Digital Data Interoperability roadmap**, which gives advice on how to embed data quality within a broader roadmap for managing and improving data

Managing the quality of an individual dataset

The **[DQ1] Digital Data Quality Principles** makes clear that data quality should be 'appropriate for purpose'. This means that the requirements for quality in respect of a specific dataset depend on the current and potential use of that digital dataset – not all digital data needs to be at the highest quality levels. Therefore, it is recommended that before assessing and planning data quality improvements, the Entity has a good understanding of its data assets and their function and importance for the Entity as well as other potential users. Use **Guidance Note 3: Developing a Digital Data Inventory** and **Guidance Note 4: Prioritization criteria and process** to do this.

The Digital Data Quality Principles are listed in full below.

UAE Data Digital Quality Principles	Description
Ownership and authority	• The digital data is managed by an accountable Data Custodian, who is responsible for the quality of the data and ensuring it meets user needs. • The digital data being published or exchanged is not a copy, but is made accessible at source. Users are accessing and re-using the original data, thus reducing duplication and errors. • Entities should identify duplicate versions of digital data and designate a specific owner and authoritative source of that data. Further, they should ensure other functions re-use this authoritative and maintained source directly, for example, via an API or downloading a copy from where the digital data lives as needed and not redistributing their copy.
Accessibility	• The data is easy to find and use, because it: − Has comprehensive metadata for discoverability − Uses appropriate open machine readable formats (reducing the requirement to buy specific proprietary software and ensuring it is interoperable with other digital data) and, − Is made available for bulk download or via an API either on the web or through a platform with reliable lasting permanent access that is supported over time.
Accuracy	• The data is sufficiently accurate for its intended use and any gaps, known limitations, approximations or errors are clearly described so that re-users understand the limitations of the digital data. • Users both inside and outside the Entity should have a way to communicate their requirements for greater accuracy and have those acted on by the Data Custodian / Data Specialist responsible.
Descriptiveness	• Digital Data has context to that potential re-users know what is in the digital data and how reliable it is so that they can effectively judge whether it is fit for their purpose. • This means all datasets should have associated metadata and ideally a schema specifying the ranges and values of each field. • Re-users should be able to understand how the data was created and processed, it's temporal and geographic coverage, granularity and limitations.
Timeliness	• Digital Data is published or made accessible in real-time or soon after the digital data has been generated. The data being published for re-use as open data or exchanged with other Entities should be the same data as that being used for its intended purpose within the digital data generating Entity. • If it is regularly updating data (such as a monthly report) the update schedule should be clear in the metadata and should be reasonably followed closely to ensure re-users can rely and trust this data for operational needs and decision making purposes.
Completeness	• The digital data should make sense as a complete dataset. It should be usable without requiring other digital data (other than Primary Registries data) to make sense or use of it. This means digital data should be published or exchanged as datasets which are comprehensive and relevant missing records should be flagged.
Validation	• The digital data should be valid and effort made to ensure it is accurate and reliable over time. • This means for core and frequently updated digital data:

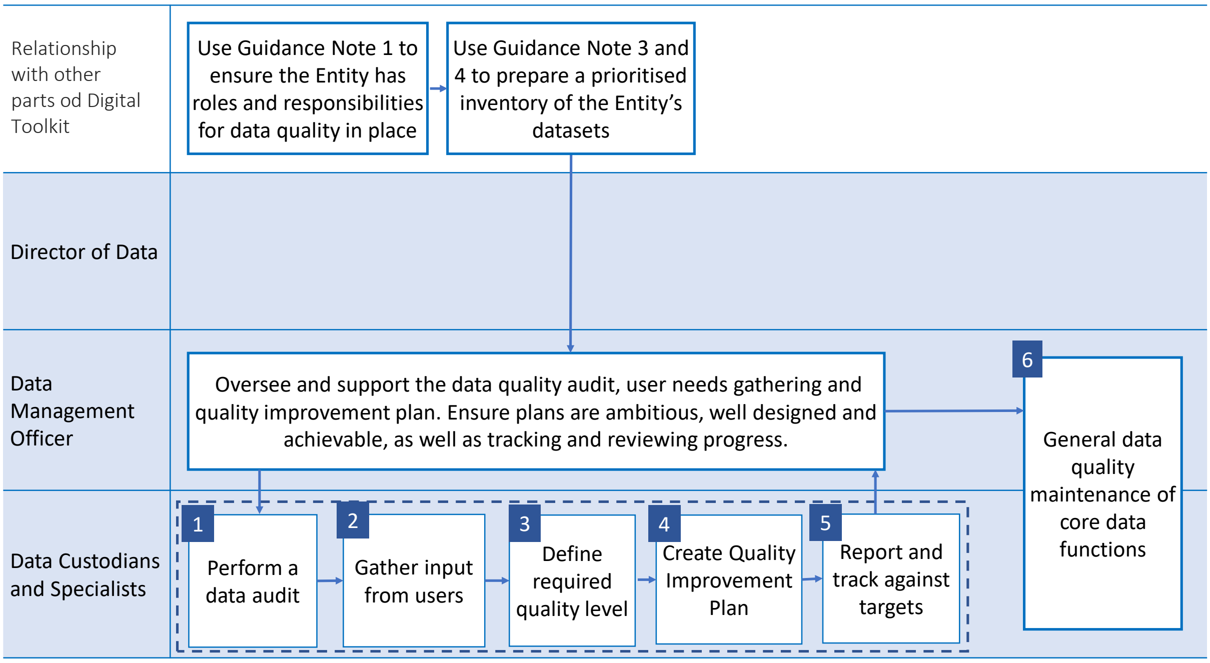
	– Using a schema – Having a clear data model with unique identifiers for the main objects in the data (for example National ID for citizens) – Regularly cleaning and testing the digital data to remove errors or duplicates.

Not all of these Principles will need to be applied in full to every digital dataset in order for digital data to be appropriate for purpose. Digital Data quality may be appropriate for current purposes even if one or more of the principles is overlooked. (For example, digital data may be collected with a lower accuracy level to provide digital data in a timely manner if time is a priority.) This means that these Principles should be balanced against the importance and intended use of the relevant digital data.

However, some quality characteristics are essential in order to enable effective data publication and exchange. These characteristics have been built in as mandatory elements of the Digital Data Exchange Standards, as summarised in the table below.

Core quality requirements	**Detailed standard which sets out mandatory requirements in this area**
Ensuring digital data is published or exchanged in appropriate formats	**[DE1] Digital Data Formats**
Recording and publishing mandatory metadata: the **title**, **description**, **subject**, **format**, **size**, **publisher**, **custodian**, **classification**, **access permissions**, **license**, **coverage (temporal and geospatial)** and **last updated timestamp**	**[DE2] Metadata**
Publishing and validating against a data schema for high value, structured and regularly updated datasets as well as any primary registries the Entity holds	**[DE3] Schema**

Once roles are in place and there is a prioritized inventory of the Entity's digital data assets, the assigned Data Custodians should first assess the state of data quality, then assess the quality level required by data users and then make a plan to close the gap between the two over time, as illustrated below.


Relationship with other parts od Digital Toolkit
Use Guidance Note 1 to ensure the Entity has roles and responsibilities for data quality in place
Use Guidance Note 3 and 4 to prepare a prioritised inventory of the Entity's datasets
Director of Data
Data Management Officer
Oversee and support the data quality audit, user needs gathering and quality improvement plan. Ensure plans are ambitious, well designed and achievable, as well as tracking and reviewing progress.
6
General data quality maintenance of core data functions
Data Custodians and Specialists
1
Perform a data audit
2
Gather input from users
3
Define required quality level
4
Create Quality Improvement Plan
5
Report and track against targets

1. Assess current data quality by performing a digital data audit

Use the Digital Data Quality Maturity Matrix provided in Appendix B to assessing the digital dataset against the UAE Data Digital Quality Principles, looking in turn at:

- Does the digital data have a clear owner? Is this the authoritative source of digital data?
- How accessible is the digital data?
- How accurate is the digital data?
- How well described is the digital data?
- Is the digital data up to date and has a publishing schedule?
- Is the digital data complete? Can it be used and make sense by itself?
- Has the data been validated against a schema or checked for duplication, errors, and inaccuracies?

maturity:

- **Level 1: Initial** – unmanaged digital data, no owner, no open format, no metadata etc
- **Level 2: Partially conformant** – the digital dataset has an identified owner and is making progress towards conformance with the Digital Data Quality Standard
- **Level 3: Conformant** – the dataset meets all core requirements of digital data quality and UAE Data Standards
- **Level 4: Improving** – thedigital dataset meets all core requirements and also is implementing additional good practices
- **Level 5: Optimizing** – digital data quality fully meets the needs of current and potential future users, with clear systems for driving continuous improvement.

	1 = Initial	2 = Partially conformant	'1 – Conformant	4 = Improving	5 = Optimizing
Ownership and authority					
Accessibility					
Accuracy					
Descriptiveness					
Timeliness					
Completeness					
Validation					

The detailed tool for use when completing this matrix is at Appendix B of the UAE Digital Data Interoperability Implementation Guide.

2. Collect input from existing and potential users

Run consultations and workshops with existing internal and external users of digital data. Ask whether the digital data is fit for their purpose, whether they have any problems or concerns with the digital data and whether they have any requests around data quality. Record and group the results by most voices/concerns/requests.

It is also important to assess the needs of potential users who may not yet have access to the digital data or would use the data if it was of a higher quality or reliability. Therefore, it is recommended to invite input and feedback from potential future users. The Entity should publish its Digitall Data Inventory on its digital portal, including digital datasets that have not yet been prepared for publication and exchange, in order to give potential users visibility of its digital data assets. It should also provide online channels for data users to give feedback on their priorities for expanding the number of digital datasets that are available on the portal and improving the quality of existing open data.

3. Define and determine required digital data quality per dataset or data source

Using the feedback from users, define what good digital data quality looks like for the digital dataset. Develop a documented statement of Digital Data Quality Requirements – including what the appropriate target level should be for each element of the Digital Data Quality Maturity Matrix provided in Appendix B. Record any specific measures or indicators which are key to ensuring the digital data is reliable and fit for purpose for the majority of users.

4. Create a plan to close the gap between existing and required quality level

Create a plan to reach defined quality level. This may require new processes, change management, upskilling and training, better tools or other steps. Ensure your plan is ambitious, but realistic. Milestones should be specific, measurable and time bound with clarity on who is responsible for the milestone being achieved and how this will be measured and tracked.

Data Specialist and the Entity's IT and security teams should create automated tracking for quantitative measures such as % of metadata complete, use of open machine-readable formats, % of publishing frequency dates realized, whether datasets have a schema, results of data validation against schema, results of scripts to check duplicate records or unconformity data entries etc.

Data Custodians should also track qualitative measures such as feedback from users and impact of use of digital data.

- Embedding the **[DQ1] Digital Data Quality Principles** across all digital data within an Entity will require a phased, prioritised and Entity-wide plan of action.
- Development and delivery of such a plan is a mandatory requirement for Government Entities.
- An effective Digital Data Quality Improvement Plan will be **prioritized**, **baselined**, **user-focused**, **SMART**, **managed** and **reported** as described in the table below.

An effective Digital Data Quality Improvement Plan is	**Description**
Prioritised	• The Plan should focus first on driving up the quality of digital data needed for **Primary Registries**, the Entity's own **core business functions**, and other **high priority digital datasets**. • Within these priority areas, it should focus first on fixing known quality issues – and in particular focused on ensuring that the **Core Quality Standards** are met.
Baselined	• The Entity should ensure that its plans are informed by Digital **Data Quality Audits** that give a clear assessment of current performance against the [**DQ1] Digital Data Quality Principles**. • These should include **quantitative and qualitative measures of digital data quality**, including both use of the **[DQ2] Digital Data Quality Maturity Matrix** and additional measures that are relevant to the specific dataset.
User-focused	• For all priority datasets, the Entity should develop clear statements of Digital **Data Quality Requirements**. These should be evidence-based and reflect the documented quality needs of users. • In developing these user requirements, the Entity should engage with existing internal and external data users – but also consider the wider potential re-use of their data (either as open data or shared data exchanged with other Entities). • These Digital Data Quality Requirements should specify and define required **digital data quality measures** for their different types of data sources and business processes, aligned with the [**DQ1] Digital Data Quality Principles**.
SMART	• For each priority dataset, the Entity should: – **Identify the gaps** between the current baseline perfromance level as revealed in the Data Quality Audit and the data quality requirement as expressed by users – Set quantitative and/or qualitative **targets for improvement**. Targets should be SMART (Specific, Measurable, Achievable, Relevant and Time-bound) • A on-size-fits-all approach to data quality targets across the Entity is not recommended: rather these should be related to the current and potential use of the specific digital dataset, to ensure the quality is appropriate for that use.
Managed	• The Entity should set out an overall Entity-wide plan for how it will deliver its targets for quality improvement. This should include: – Establishing clear accountabilities for data quality, at the Entity-wide level and for each dataset – Establishing systems and processes that guarantee data quality as part of the normal business activity of the Entity – Building data quality requirements into any contracts and outsourcing of digital data management or data generation – Assessing digital data quality of third party suppliers (which could include another government Entity, business partner, customer, service provider or other stakeholder), and performing spot checks (ideally against Service Level Agreements with the data supplier).
Reported	• Entities should establish **systems to track and report** on data quality status, with the Entity's Management Board receiving regular progress reports (for example, on a

	quarterly basis) showing progress across the Entity as a whole and by individual business units. • Ideally, elements of this reporting will be **automated and managed in real-time**, for example through: – automated reports on dogotal data quality indicators (such as completeness of metadata fields, use of schemas, success rate of validation, use of open machine readable formats, update frequency % met, etc) – Regular analysis of structured data against its data schema.

5. Report and track digital data quality against targets

Data quality should be tracked and assessed against required level as well as being reviewed regularly (at least annually). User requirements may change over time and the Entity should respond to this.

6. General digital data quality maintenance

Outside of meeting specific user-needs-based digital data quality requirements, Entities should ensure the overall health of their digital data. This means regularly reviewing and improving how digital data is modelled, use of schemas and digital data entry validation processes, use of unique identifiers, links and use of reference data (vs. introducing un-authoritative copies), instances of record duplication, errors or uncompliant data entries (for example mis-spellings or different date/telephone formats being used for the same data) and so on.

Entities should decide on their own processes for managing this. An example end-to-end data cleansing process is detailed below. Typically this should be seen as an iterative process that should be repeated to improve and maintain digital data quality as business and technical requirements change.

Digital Data-cleansing step	**Description**
1. Extract data from operational data sources for profiling	Digital Data profiling tools perform complex analysis on digital data, and to perform this analysis directly against live data sources is not recommended. Digital Data extraction may be performed using separate ETL tools, or may be a capability of the digital data profiling tools themselves.
2. Perform digital data profiling analysis	This shall occur as part of a regular data audit process, enabling digital data quality issues to be identified. The output of data profiling shall be used to build the technical knowledge base for digital data cleansing.
3. Build cleansing knowledge base for each digital data profile	The cleansing knowledge base includes mappings and correction rules that may be automatically applied. For example, the range of mobile phone formats identified by data profiling may include (nnn) nnn nnnn, +nnn nnnnnnn, nnn nnn-nnnn. The knowledge base should include the rules for converting these formats into a single format. A knowledge base may include the ability query external digital data services, such as telephone number validation, reference data management systems, and digital data enriching systems, such as an Emirates ID service to provide more Citizen profile digital data. Physically, the knowledge base may be one or more systems, and may include master digital data management tools, reference data management tools, and vendor specific digital data cleansing solutions.

4. Automated cleansing using knowledge base	Automated cleansing may be performed in batch against live systems, typically out of hours, and subject to sufficient testing. The size of the batch chosen should be determined by the smallest batch of digital data that can reasonably be completed within the time window allowed. The choice of records that form part of the each cleansed batch shall be defined, for example, through order of insertion, age based (newest/oldest) first, or most active records first. Automated cleansing can also be applied to data extracts; however, the plan to refresh the live data with the cleansed digital data should be considered carefully to avoid conflicts where the live data has since changed.
5. Interactive digital data cleansing	Automatic matching will reject data that cannot be cleansed. The Data Custodian shall use this rejected data to perform manual cleansing. The recording of cleansing decisions should be fed back into the knowledge base to improve the quality of automated matching. This iterative cycle will initially occur often during the development of the knowledge base.
6. Automated cleansing services	Automated cleansing services can then be delivered as interactive services, allowing information systems to have data validated and cleansed at the point of digital data entry. For example, a CRM system for capturing a citizen's name and address may make a service request to the automated cleansing service to enrich the address, validate the telephone number, and match the individual citizen with their other records stored in digital datasets elsewhere within the Entity.

5.6 Validation and publication of digital data

Purpose	This Guidance Note provides Entities with guidance on the process and steps for validating and then publishing digital datasets that have been prepared for digital data conformance using Guidance Notes 5.1 to 5.5.
When to use	Before publication of Open Data or exchange of Shared digital Data
Responsibility	Data Management Officer, reporting to the Director of Data who has overall accountability for conformane with UAE Data Exchange Standard.

Once the relevant Digital Data Custodian has taken a digital dataset through the process described in Guidance Notes 5.1 to 5.5 (that is: classify; format; document the permissions model; add metadata and develop schema; manage data quality), the Data Management Officer and Director of Data will need to validate and approve the digital dataset either for publication or for sharing and exchange with other government entities over appropriate electronic networks.

First, the Data Custodian / Specialist should provide all of the relevant information (on classification, format, metadata and quality) to the Data Management Officer in a single ‘conformant dataset’ with the conformant data sample file, the metadata, and the details of the business processes needed to support quality publication (what needs doing, who is responsible, timelines).

This complete dataset should be reviewed by the Data Management Officer and added to an Entity-wide Data Inventory. The Data Management Officer should check that each digital dataset:

- ✓ Has a classification. In cases where the classification is not *Open*, then a clear rationale for this isdocumented and an Open derivative dataset provided
- ✓ Has a digital data quality assessment report
- ✓ Has a sample digital dataset in the appropriate format
- ✓ Contains all Mandatory Metadata as defined in the **[DE2] Metadata** specification
- ✓ Has easy-to-add and appropriate Recommended Metadata
- ✓ That the above have been provided by a qualified person familiar with the digital data

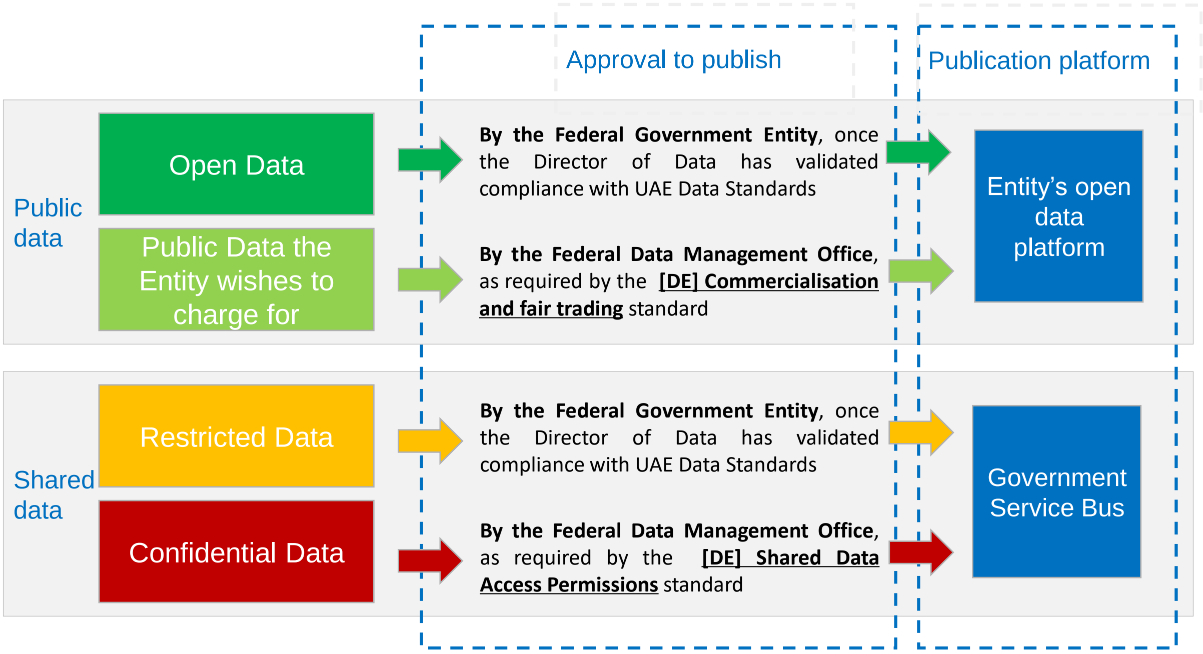
The Director of Data should then satisfy themselves that each digital dataset is conformant to the Digital Data Interoperability Framework standards. In most cases, this decision will be taken within the Entity itself by the Director of Data.

In some cases, however, the UAE Digital Data Interoperability Standards require that the approval of the Entity Overseeing Digital Data is given ahead of publication or exchange. In particular, the consent of the Office will need to be given in advance of publication for:

a) Any Open data which the Entity believes it should charge users a fee to access, despite the general principle that open data will be published as Open Data

b) Any digital data that Entities wish to exchange which has been classified as Confidential.

A summary of these approval requirements is given below.



By the Federal Government Entity, once the Director of Data has validated compliance with UAE Data Standards

By the Federal Data Management Office, as required by the **[DE] Commercialisation and fair trading** standard

By the Federal Government Entity, once the Director of Data has validated compliance with UAE Data Standards

By the Federal Data Management Office, as required by the **[DE] Shared Data Access Permissions** standard

Once satisfied that the relevant approvals are in place, the Director of Data should communicate to the relevant teams that the compliant datasets can be published as open data or exchanged with external Entities. As illustrated above, publication to the Digital Data Interoperability electronic platform will be through one of two routes depending on the type of data being published.

For *Open* data, compliant digital datsets should be published through the data-owning Entity's online portal. By default, this should be as Open Data - except in exceptional circumstances where an access fee is charged following the approvals process illustrated above and in conformance with the **[DE5] Data commercialization and fair trading** specification. The portal should:

- Include a full list of digital datasets on the Entity's Digital Data Inventory (including as yet unpublished digital data, in order to facilitate feedback on future publication priorities from data users)
- For Open Data, the portal should:
  - Provide a clear and user-friendly Open Data Licence, which complies with the **[DE4] Open Data Licensing** specification giving a clear and unambiguous license to use and distribute including for commercial purposes. (The UAE Federal Open Data License at Appendix A is recommended as the ideal way of meeting this requirement.)
  - Enable anonyomous access to the digital data, without requiring users to register any personal details or fill out forms
  - Not charge any access fees
  - Not discriminate between types of user
  - Provide data in open, machine-readable formats that comply with the **[DE1] Digital Data Formats** specification, or enable direct API access to the digital dataset
  - Provide bulk download functionality for data and gurantee a level of permanence by not breaking URIs and ensuring original URIs redirect if the dataset location (URL) changes.
  - Provide metadata that complies with the **[DE2] Metadata** specification
  - Provide measures of digital data quality, including use of the **[DQ2] Digital Data Quality Maturity Matrix**
  - Provide online mechanisms for users to give feedback about digital data quality and to express their views on future priorities for expanding the number of digital datasets that are available on the portal and improving the quality of existing open data
- Or, in exceptional cases where the Entity Overseeing Digital Data have approved the Entity to charge an access fee for Open Data, the Entity should:
  - Ensure that the rationale for charging, and the principles that the Entity applies to ensure fair and competitive provision in line with the requirements of the **[DE5] Commercialization and fair trading** specification, are published clearly on the website
  - Provide access to effective complaints and redress mechanisms, again in conformance with the **[DE5] Commercialization and fair trading** specification.

For **Shared Dgital Data**, the Entity should use the Government Service Bus as the key platform for exchanging its data with other Entities. No charge is required for Entities to integrate with the Government Service Bus. For Shared Datasets that have been identified as Primary Registries, the data should be made available:

- Via API
- Under the terms of a Service Level Agreement setting out the expectations that digital data users can have in terms of the Entity's commitment to the quality of the digital data
- With privacy 'designed-in'. This means that the API should give access to the smallest amount of information required for the service outcome or to enable a decision. (For example, sending 'yes', 'no' or 'not found' in response to a query of whether a citizen or user is over 18 or has a valid driving license instead of sending personal information.)

Charging digital data users to access raw Open Data

- Open Data which a Government Entity collects and manages in the course of its normal duties should be published as Open Data with no access fee
- Where there is demand from data users (whether from other Government Entities, the private sector or individual citizens) for access to digital data that the Government Entity does not currently collect and that would require significant additional action and investment by the Government Entity, then there may be a case for charging fees to digital data users in order to help finance this investment.
- For example, there may be cases when a Government Entity wishes to develop a **public-private-partnership (PPP) funding model,** aimed at ensuring up-front investment in the infrastructure needed to enable the collection and dissemination of richer, smarter, more real-time data that is relevant to the purposes of the Government Entity, but which otherwise could not be obtained by the Entity. In such cases, business models might involve:
    - Allowing the PPP to charge end users for dgital data to deliver a revenue stream into the PPP on a time-limited basis before the data is eventually made fully Open
    - Providing the PPP with exclusive rights to utilise the data commercially in ways other than direct charging of end users, again on a time-limited basis before the digital data is eventually made fully open.
- Such cases will be exceptional, not routine, and should be approved in advance by the Entity Overseeing Digital Data.
- Approval to charge for access to raw Open Data will only be given when this is clearly in the public interest, and where it is not feasible for the Government Entity to collect and publish the data without charging access fees.
- In making this determination, the Entity Overseeing Digital Data will take into account that all Government Entities are expected - as part of their routine operations and investment planning – to continually improve the methods, quality and timeliness of the data they collect without seeking to charge data users for this.
- Whenever, in such exceptional cases, a Government Entity does charge fees for access to raw Open Data it should follow these principles:

Raw Digital Data Commercialization Principles	Description
1. **Public interest**	• Charged-for Open Data should be accompanied by a clear published explanation on the Government Entity's electronic portal of why such charging furthers the goals of UAE Digital Government and is in the public interest.
2. **Fair pricing and conditions**	• Open Data should be available to all users on a fair, reasonable and non-discriminatory basis.
3. **Account-ability**	• The Entity should establish and publicise effective complaints and redress mechanisms for third parties who believe that it is failing to comply with the above principles.

Developing and marketing commercial value-added digital data services

- In general, the Government believes that the private sector is best placed to create commercial data services, and will not seek to do so itself.

- There may be cases however where a Government Entity may provide such commercial services in the public interest – for example, in order to demonstrate the commercial opportunity, as part of its efforts to foster the market for re-use of its digital data.
- Such cases should be approved in advance by the Entity Overseeing Digital Data, following receipt from the Government Entity of an evidence-based submission showing how the proposed service complies with the principles set out below.
- Whenever, in such exceptional cases, a Government Entity does develop and market value-added digital data services for commercial gain, it should publish and follow these principles.

Value-added Data Commercialization Principles	**Description**
1. Public interest	• The provision of commercial data services should be accompanied by a clear published explanation on the Government Entity's digital portal of why such commercial services further the goals of UAE Digital Government and are in the public interest.
2. Fair competition	• The Government Entity should ensure that it does not have an unfair advantage over third parties who might also wish to market similar services. In particular, this means: – **Publication of the underlying Open Data:** the Government Entity should publish as Open Data on its electronic portal the underlying Open Data that it is using to create value-added services. This publication should be undertaken at the same time as, or before, launch of the Government Entity's value-added data service. – **No use of Shared Digital Data:** for the purposes of developing a value-added digital data service, the Government Entity should only use digital data that has been classified as Open Data. – **No use of public funds:** in order that Private Entities may compete on a fair basis in the provision of commercial services using Open Data, the Government Entity should set fees for any value-added data services in ways that at least recover the full costs of providing those services, including a reasonable return on investment, and should ensure that the provision of these value-added data services is not based upon anti-competitive support or funding from other Government Entities.
3. Fair pricing and conditions	• The Government Entity should make the value-added digital data services available to all users on a fair, reasonable and non-discriminatory basis
4. Account-ability	• The Entity should establish and publicise effective complaints and redress mechanisms for third parties who believe that it is failing to comply with the above principles.

APPENDIX A: UAE FEDERAL OPEN DATA LICENSE

The Federal Open Data License is shown below in two forms:

- A user-friendly, plain language summary for publication on the web pages of Open Data Portals
- The detailed License terms which support this and which should be available for download from Open Data Portals and linked from the summary.

Federal open data license: summary

This is a summary of (and not a substitute for) the license that applies to Your use of Information accessed via [name of Open Data Portal] ("**License**"). A copy of that license may be accessed <**here**>.

1. Overview

We grant you a worldwide, royalty-free, perpetual, non-exclusive license to use and re-use the Information that is available under this license freely and flexibly, subject to the conditions below.

2. You are free to:

✓ copy, reproduce and communicate to the public the Information in any format

✓ adapt or modify the Information

✓ exploit the Information for both commercial and non-commercial purposes

✓ permit third parties to use the Information

3. You must comply with the following terms:

! not use the Information in any way that is unlawful and/or misleading to the general public

! Include the name or identification of the author and retain any copyright notice featured in the original material

! Where possible, include a URL or hyperlink to the Licensed Material

These are important conditions of the License and if you fail to comply with them the rights granted to you under this license may be withdrawn.

4. Exemptions:

The License does not permit the use of:

X any trademarks associated with a Database or with the Open Data Platform

X any images (including logos, graphics or photographs) which appear in a Database

Federal open data license: full text

When you access and use the Information, you accept and agree to be bound by the terms and conditions of this License in connection with your use of the information provided by the License issuer.

Article (1) Definitions:

Federal Government Entity	Means any Ministry, authority, department, public body, independent body, public institution, federal government council or any other governmental or public institution of the federal government of the United Arab Emirates;
license	Means the general license agreement as updated or amended by the license grantor; the legal terms under which the original materials are made available to disclose;
Original Materials	Means all the contents of any database (or any part thereof) that includes any digital data, content, work products or other materials that have been collected in the database and made available to disclose by the license grantor to users under the terms of this general license; or derived materials;
Modified Materials	Means any work in any medium (whether currently produced or to be created in the future) created by entity or created by any other recipient that incorporates or uses any original information or material either alone or in conjunction with materials from another source and as an independent product;
Derivatives Materials	Means any work in any medium (Currently known or to be created in future) created by entity or created by any other recipient incorporating, using or quoting any original information or material that is subject to copyright and similar rights in which the copyright, Conversion or other modification of the original articles in such a manner as to require authorization under copyright and similar rights;
You or the conscience of the addressee	Means the individual or entity using the original material to develop modified or derivative materials under this general license;
License grantor	Means the person/persons or federal entity/entities that granting rights under this general license;
Copyright and disseminating	Means the rights granted by copyright and / or similar or related rights closely related to copyright including performance, broadcasting and recording of sound and rights in databases or literary collections;

The Use	Means copying, reproducing, and making copies or disclosing material that you do through a medium or process requiring permission under this License, communicating with various users, modifying, adjusting and preparing modified or derivative works, and any other work that is Confidential or may become confidential in the future under copyright whether in original or other material
The User	Means any user of information other than you;
User License	Means the license you apply to the user of the modified materials or derivative materials in accordance with the terms and conditions of this general license;
Participation	Means the disclosing of material to the public through any means or process requiring permission under this license, such as copying, public disclosing, public performance, distribution, dissemination, media or import, and making material available to the public, including by means of access to materials In the place and time of their choice;
Exceptions and Limitations	Means any exclusion or other restriction on copyright and similar rights applied to your use of the Original Materials;

Article (2) Scope of license rights:

2.1 Grant of license

2.1.1 Take into consideration the terms and conditions of this General License,

2.1.2 The license grantor grants you a free and non-exclusive license to download information over the Internet and technology media

2.1.3 The License grantor authorizes you to exercise the licensed rights in all media and formats currently produced and known or to be created later, and to apply the necessary technical modifications on.

2.1.4 This General License may not be sublicensed and irrevocable for the exercise of rights under this License in order to copy and share original materials, in whole or in part, or to produce, reproduce and share modified or derivative materials.

2.2 The License grantor waives and/or agrees not to endorse any right or authority or to prohibit any entity or individual from making the technical modifications necessary to exercise the licensed privileges, including necessary technical modifications and performing any authorized variations in this section and does not result in any modified or derivative materials.

Article (3) Terms and Conditions of Use

3.1 Terms and conditions of Use:

3.1.1 You must ensure use is not contrary to UAE or international laws.

3.1.3 Reference should be made to the source of the original material

3.2 You may allow users to use the original materials and, if you do so, you must comply with the terms of this license and you are prohibited from displaying or imposing any additional or different terms or conditions on the use of that information by any other user.

3.3 This license does not cover the use of:

3.3.1 Personal data within information such as identity documents such as passport number or national identity;

3.3.2 Information that not disseminated and not disclosed with the consent of the license grantor;

3.3.3 Rights of third party that license grantor have no authority to disclose;

3.3.4 Any images (including logos, drawings or photographs) that disclose within the original materials;

3.3.5 Information subject to other intellectual property rights, including patents, trademarks and design rights

Article (4) the terms of the license

Your employment of the licensed rights is subject to the following conditions:

4.1 Attribution:

4.1.1 The user may share the original materials by considering the following:

4.1.2 Retain any copyright notice contained in the original material;

4.1.3 Inclusion of an electronic link or hyperlink to the original material in reasonable form,

4.2 Sharing of modified and derivatives materials:

4.2.1 This license grant users the right to redistribute, modify, change, and quote from your materials whether for commercial or non-commercial purposes as long as they associate/attribute your original material to your name.

4.2.2 This license grant users the right to modify, improve, and create new derivative materials from previously modified or derivative materials, whether for commercial or non-commercial purposes, as long as they authorize their new derivative works with the user's license under the terms of this license.

4.2.3 You must comply with the requirements stated in Section 3 and include them in the User License if the contents of the entire database or portion of the original material are shared.

Article (5) Disclaimer of warranties and limitation of liability

5.1 The License grantor is not responsible for any damage or misuse suffered by third parties as a result of the use of such data and does not guarantee the continuity of the availability of such data or part thereof, nor shall it be liable to users of such data and any damage or loss they may suffer due to reuse.

5.2 It is prohibited to sell or resell any original information have been used in accordance with this License for any fee or amount of money or for any form of compensation or reimbursement.

Article (6) - Duration and Termination

6.1 This general license shall apply throughout the period of copyright and similar rights licensed therein. If you do not fully adhere to the terms of this license, your right to use the original materials under this license shall be revoked and the terms of this license will remain in force notwithstanding such cancellation.

6.2 The license grantor has rights to disclose the original materials under separate terms or conditions or discontinue the disclosing of the original materials at any time; however, the terms of this license shall remain in force notwithstanding such cancellation.

6.3 The terms of the License shall remain in force after termination of this General License.

Article (7) - Other terms and conditions

7.1 The license grantor shall not be bounded by any additional or dissimilar terms or conditions unless explicitly agreed to do.

7.2 Any arrangements, considerations or agreements relevant to the original materials not mentioned in this License shall be considered separate and independent from the terms and conditions of the General License.

7.3 This General License shall not be construed as derogation, restriction, prohibition or imposition of conditions on any use of the original materials which may be made legally without permission under this General License.

7.4 It is not permitted to disclaimer of any condition or provision in this General License and neglecting compliance to unless expressly agreed to by the license grantor

7.5 Nothing in this General License shall constitute or be construed as a restriction or waiver of any privileges or immunities applicable to the license grantor or to you, including immunity from legal procedure in any jurisdiction or authority.

7.6 This License is governed by and construed in accordance with the laws of the United Arab Emirates

End of license

APPENDIX B: DIGITAL DATA QUALITY MATURITY MATRIX – ASSESSMENT TOOL

Quality Principle	1 = Initial	2 = partially conformant	3 = Conformant	4 = Improving	5 = Optimizing
Ownership and authority	The dataset has no clear accountable owner within the Entity. Multiple data users keep and manage duplicate versions of the data.	A named Data Custodian takes personal responsibility for the quality of the data. The Data Custodian has undertaken a baseline assessment of current data quality, documenting known quality issues.	As at Level 2. In addition: – The Data Custodian has engaged with current and potential future users of the data to understand and document their Data Quality Requirements, and is managing a plan to close any gaps between current and required quality levels – For a data set used by multiple organizations, systems and processes have been established to ensure that it can be managed as a Primary Registry (ie able to provide the data as a service to all relevant users).	As at Level 3. In addition: – Feedback mechanisms have been established to allow data users to request quality improvements – In the case of a Primary Registry, the dataset is now widely used as the single authoritative source of data. There are no duplicate versions of the data managed elsewhere.	As at Level 4. In addition: – There is clear evidence that effective processes are in place to enable user-driven continuous improvement. – In the case of a Primary Registry, the dataset is now accompanied by clear Service Level Agreements for data users
Accessibility[4]	The data is inaccessible by third parties because it is: – Not published on the web or currently shared with other Entities.	The data is at least one of: – Published on the web or via an API – Available to external users in an open machine-readable format.	The data is accessible through both: – Publication on the web or via an API – Publication in an open machine-readable format.	As at Level 3. In addition, – Published data is available for bulk download – The data uses URIs / URLs to enable others easily to link their data to it.	As Level 4. In addition, the dataset is linked to other relevant data to provide context.

[4] The maturity levels for accessibility draw on the Five Star deployment scheme for open data developed by Sir Tim Berners-Lee, but expanded to cover shared data as well as open data. An open data set scoring 1 to 5 on the Five Star model would score the same on the accessibility dimension of the UAE Data Quality Maturity Model.

Accuracy		Accuracy issues in the dataset (errors, gaps, limitations) are either unknown, or known to be very significant.	There are significant accuracy problems with the data, but these are documented and explained to data users.	Known accuracy problems are documented and explained to data users. Accuracy level is adequate for current use or purpose.	As at Level 3. In addition, the Entity is actively reaching out to potential data users to understand how accuracy improvements could support new use cases for the data.	Data accuracy is fit for purpose for both existing and potential uses of the data, based on clear, documented user-research and feedback.
Descriptiveness		The dataset has no metadata or schema.	The dataset has some metadata or a schema describing the data.	The dataset has all mandatory metadata. In the case of a Primary Registry, it has a schema.	The dataset has all mandatory metadata and it has a schema.	As at Level 4. In addition, the dataset has all the additional recommended metadata.
Time-liness	*For updated datasets*	The dataset is out of date.	The dataset is regularly updated, on a timescale that meets the needs of current users.	As at Level 2. In addition, the dataset has a publishing schedule which is being met in practice and which is included in metadata for publishing frequency.	As at Level 3. In addition, guarantees exist that the most up to date data will be available in future over a specified period.	As at Level 4. In addition, data updates are managed in real-time, with publication or exchange occurring at the same time for internal data users and external data re-users.
	For one-off datasets	The dataset is out of date, to the point where it has no useful value.	The dataset is out of date, but still has some use value for users.	The dataset is recent enough to meet all the needs of its users.	[No Level 3 for one-off datasets]	[No Level 4 for one-off datasets]
Completeness		There is significant missing data in the dataset or coverage is poor and this is not documented.	There is missing data in the dataset or coverage is poor and this is documented and explained to data users.	Data completeness is fit for current purposes. It makes sense as a dataset, can be used by itself or in combination with reference data and any missing records or gaps are documented and explained.	As at Level 3. In addition, the Entity is actively reaching out to potential data users to understand how more complete data coverage could support new use cases for the data.	Data completeness is fit for purpose for both existing and potential uses of the data, based on clear, documented user-research and feedback.
Validation		No validation of data.	For creating/recording the data, use is made of vocabularies or validated fields (e.g. ensuring phone numbers are in a conformant agreed format).	As at Level 2. In addition, all fields which do not need to be free text are now validated (e.g. address lookup from postcode, checking a number is entered when a number is expected).	As at Level 3. In addition, this is encoded into a schema against which the data can automatically be validated.	As level 4. In addition, regular data cleaning is carried out to remove duplicate records and errors across data systems.