The UAE Smart Data Framework

Part 1: Smart Data Principles and Standards

Version 2.0

Document Date: February 2019

TABLE OF CONTENTS

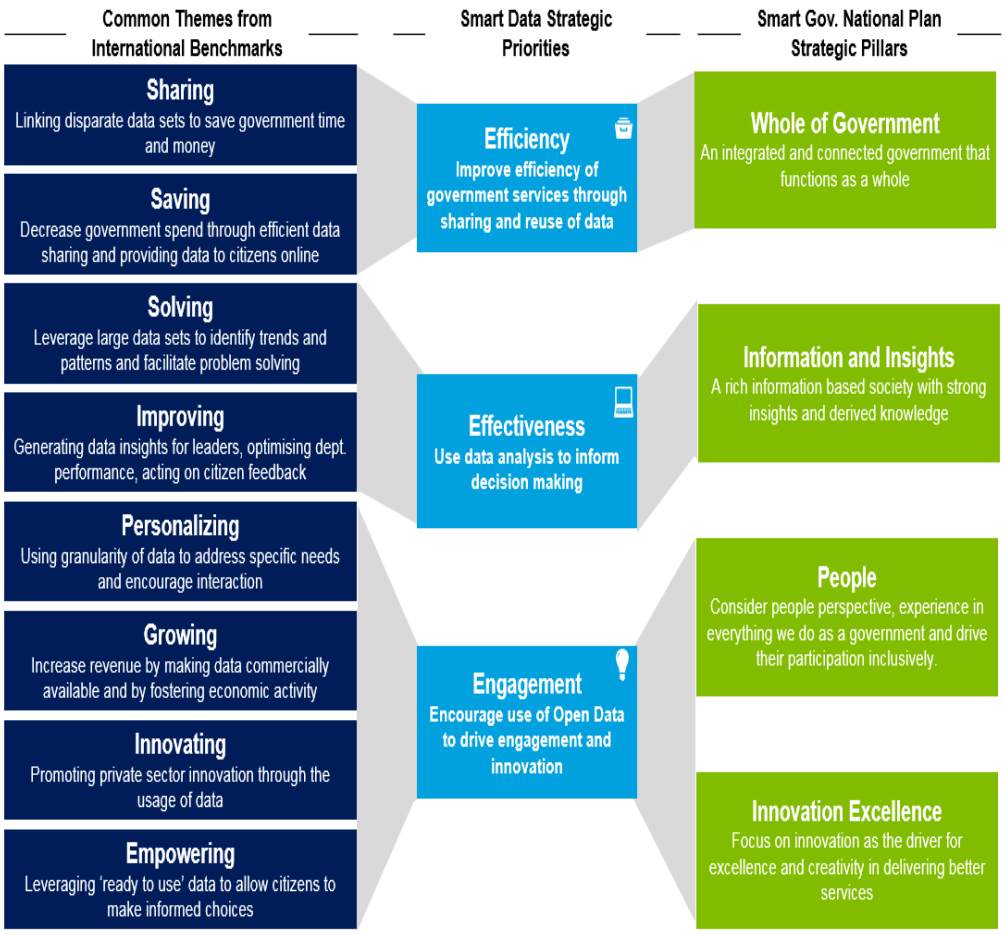
1. INTRODUCTION

This introduction to the United Arab Emirates Smart Data Framework describes:

- The **purpose** of the Smart Data Framework
- The **approach** which the Smart Data Framework takes – combining global and national best practices on data management within a flexible framework that each Entity can tailor to its own needs
- The **scope and applicability** of the Framework, and how it is intended to be used by Government Entities, Semi-Government Entities and by Private-sector Entities that re-use or integrate with United Arab Emirates government data
- The **structure** of the Smart Data Framework, which is divided into:
  - **Smart Data Principles**
  - **Smart Data Standards**
  - **Smart Data Implementation Guide**

1.1 Purpose of the Smart Data Framework

Smart Data is central to the UAE National Plan for Smart Government. As the diagram below illustrates, Smart Data – that is, data which can be used ***efficiently*** and ***effectively*** within government and which encourages use of open data to drive ***engagement and innovation*** - is a key enabler for each of the four pillars for the National Plan.



Against this context, the purpose of this Smart Data Framework is to establish the common standards and best practices needed to deliver this vision of smart government enabled by smart data.

Specifically, the Smart Data Framework has the following goals:

1. Improve data quality nationally, benefitting constituents as well as the government itself
2. Ensure efficient data sharing between government entities
3. Adopt common classification of data, based on openness, confidentiality and secrecy as appropriate
4. Provide a common basis for government data use, reuse and exchange
5. Increase the efficiency of government service delivery
6. Encourage open data sharing with the public.

1.2 Approach of the Smart Data Framework

This Framework outlines a common basis for managing data that enables interoperability and exchange among entities.

Development of the Smart Data Framework has been driven by five imperatives:

1. **Start with user needs**: Data standards only have value if they are used and they will only be used if they meet the requirements of potential users, providing them with practical tools to help address their business needs.
2. **Take a principles-based approach, and don't be prescriptive about process**: At the core of the Smart Data Framework are a set of principles for the management and use of data, described in Section 2 of this document. All Government Entities are expected to follow these principles, but with flexibility on how best to tailor them to the needs of their Entity. Where the Framework does specify mandatory requirements for Government Entities to deliver, these are limited to:
   - Specifying the outcomes that each Entity should achieve, not the specific steps they should go through
   - Specifying requirements that are critical to achieving the federal smart data goals and which any well-managed Entity could reasonably be expected to comply with, given a reasonable transition period.
3. **Build on international best practices for smart data:** The development of this Smart Data Framework has been informed by the relevant international open standards on:
   - **Building a new data-enabled operating model.** In particular, the Smart Data Framework draws on the best practice approaches to data governance, business processes and benefit realization that are set out in the global open standard "The Transformational Government Framework" [1], and in the Smart City version of that framework published by the British Standards Institute and ISO[2]. These provide clear governance frameworks to ensure that information is managed as an asset – with clear accountabilities for maintaining and exploiting data sets, supported by clear, principle-based rules for promoting re-use and innovation.

1 Published by international open standards consortium OASIS. References are to V2 of the standard published in 2014.

2 PAS181: The Smart City Framework – guide to establishing strategies for smart cities and communities. Published by the British Standards Institute in 2014, this applies the OASIS Transformational Government Framework to the specific circumstances of a city. The ISO version, ISO 37106, is to be published in April 2018.

- **International open standards for data interoperability and metadata:** at the more technical level, the Framework draws on the guidance on how to use open standards to drive data interoperability that is set out in the European Interoperability Framework[3], and on relevant open standards, including those developed by ISO and W3C, and international government experience of implementing these, including in the UK and US.

4. **Contextualize those international best practices for the UAE**. The Smart Data Framework leverages the work done on data standards by United Arab Emirates federal and local government entities, and ensures that international best practices are applied in ways that fully meet the needs of the United Arab Emirates.
5. **Technology neutrality.** The Smart Data Framework does not specify physical system details. The underlying IT infrastructures which hold and deliver data can be configured in many ways, and the principles and standards set out in the Framework are independent of this.

1.3 Scope and applicability

The Smart Data Framework is a national resource, intended for use by any Entity wishing to use and share data that originates in the United Arab Emirates. Specifically, it provides good practices and tools for use by:

- Federal Government Entities (FGEs)
- Local Government Entities
- Semi-Government Entities
- Private Sector Entities exchanging data with government bodies or re-using government data.

All Entities should develop plans for aligning their data management practices with the Smart Data Framework.

3 European Interoperability Framework for European Public Services

1.4 Structure of the Smart Data Framework

The Smart Data Framework covers three levels, as illustrated on the following page:

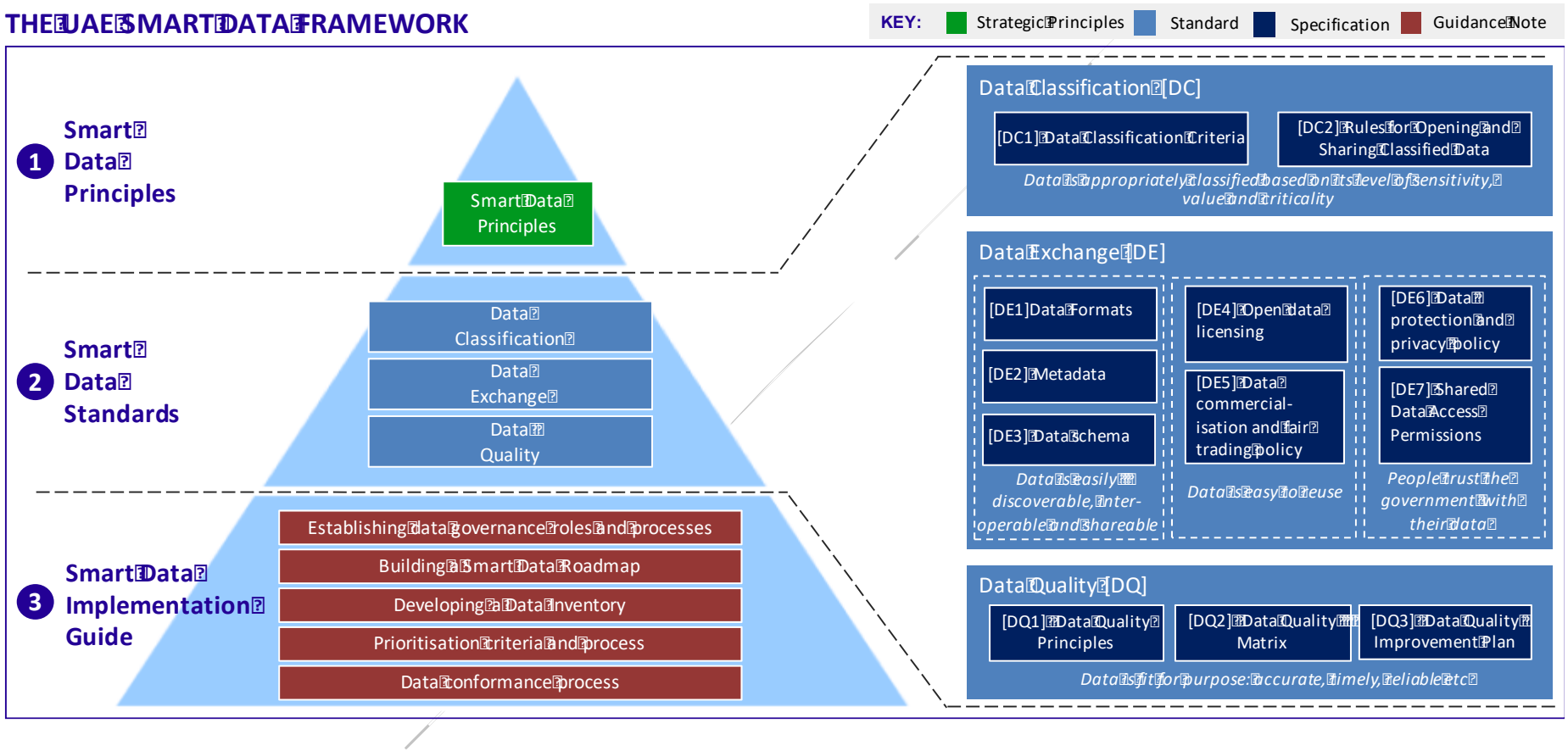
1. **Smart Data Principles:** a clear set of strategic principles to govern the creation, management, use and reuse of data in the United Arab Emirates

2. **Smart Data Standards:** three core standards required to facilitate Data Classification, Data Exchange, and Data Quality, while allowing flexibility to implement the Smart Data Principles. Each of these standards contains a set of documented specifications, that fall into two types:
   - **Dataset Processing Specifications**: these apply at the level of an individual dataset, specifying how that dataset should be classified, formatted and described in order to conform with the Smart Data Standards.
   - **Data Management Specifications**: these specify the business rules and operating principles that Entities should follow as they manage data.

3. **The Smart Data Implementation Guide:** a set of supporting Guidance Notes that Entities may find helpful when implementing the Smart Data Principles and Smart Data Standards. This will be expanded and enriched over time. In this first edition of the Smart Data Framework, the Implementation Guide is focused on meeting the needs of Government Entities seeking to align their data management processes with the requirements of the Smart Data Framework.

The rest of this document describes the Smart Data Principles in Section 2 and the Smart Data Standards in Section 3. The Smart Data Implementation Guide is in a separate document. A glossary of terms used in both documents is included as Appendix A to this document.

THE UAE SMART DATA FRAMEWORK



2. SMART DATA PRINCIPLES

2.1 Introduction

The United Arab Emirates government has developed ten key principles which guide the work of all Entities in the way they manage and use data. All stakeholders are encouraged to commit to these principles. Implementing them in full will require leadership and effort over a sustained period of time.

2.2 The Smart Data Principles

The principles for smart data that every Entity should embed within its own governance systems and business processes cover the following topics:

1. Data as an asset
2. Sharing and re-use of data
3. Duplication of data
4. Open Data publication
5. Privacy, confidentiality, and Intellectual Property Rights
6. Open standards
7. Data quality
8. Data insights
9. Collaborative governance
10. Continuous improvement

For each of these Smart Data principles, the following sections of this document provide a more detailed set of supporting principles.

Principle 1: 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 data as a collective national asset, acting as custodians of that data on behalf of the United Arab Emirates. This means that:

1.1 Entities should seek to maximize the value that the United Arab Emirates as a whole, not just their own Entity, can create from the data they collect and store.

1.2 Entities should ensure that all their datasets are explicitly identified, owned and managed as distinct assets, by following the guidelines set forth in 'Guidance Note 1 – Establishing Data Governance Roles and Processes' and 'Guidance Note 3 – Developing a Data Inventory' within the Smart Data Implementation Guide.

Principle 2: Sharing and re-use of 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 data. This means that:

2.1 Entities should identify current and potential future users of their data – across the public sector and private sector – and pro-actively respond to user needs.

2.2 Entities should encourage and promote the development of private-sector applications that use their open data.

2.3 Entities should respond rapidly and effectively to requests from other Entities and individuals to enrich and extend their open and shared data.

Principle 3: Duplication of data

In order to improve customer-centric government services, Entities should collaborate to avoid duplication and inconsistencies in their data, employing the concept of a 'single source of truth'. This means that:

3.1 Entities should collaborate to establish accurate Primary Registries that are the reliable and authoratitive source of data, and are available for use by other Entities.

3.2 Custodians of Primary Registries should manage this data in full compliance with the Smart Data Standards, making the data available as a high-quality, trusted service for use by other Entities.

3.3 Entities should not maintain duplicate versions of datasets for which a Primary Registry has been established, but instead ensure that their data management systems integrate with and pull from the Primary Registry.

3.4 Customers of United Arab Emirates government entities (at all levels) should only be requested to provide the same data to the government once. Data that has already been provided to one Government Entity – with the consent of the customer – should not have to be provided again.

Principle 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 data openly whenever possible. This means that:

4.1 Entities should, as the default position, seek to publish all non-personal data – both on their website and the United Arab Emirates Open Data Portal.

4.2 Exceptions to open publishing require a compelling case linked to clear criteria, which will generally involve protection of privacy and commercial rights or of safety and security as described in Smart Data Principle 5 below.

4.3 Whenever a dataset cannot be published as Open Data, Entities should:

- Develop a derivative version of the data set (where data is aggregated or anonymized) which can be published openly instead
- Include the dataset in their published Data Inventory, thus potential users are made aware of its existence and are able to question the Entity's rationale for not classifying the data as Open

4.5 Entities should publish their Open Data using the United Arab Emirates Open Data License, setting out clearly the rights of others to reuse the data on an unrestricted basis.

Principle 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. This means that:

5.1 The privacy of the individual should be respected and generally prevail over a desire to classify a specific set of data as Open Data.

5.2 Confidential information relating to a Private Sector Entity should be respected and generally prevail over a desire to classify a specific set of data as Open Data.

5.3 Intellectual Property Rights should be respected and should always prevail over a desire to classify a specific set of data as Open Data.

Principle 6: Open standards

In order to empower government service automation through the sharing and re-use of data, Entities should utilize open standards to make it easy for others to discover, interoperate with, and consume their 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 . This means that:

6.1 Entities should ensure that their data can be re-used by others, by following the guidelines set forth in 'Guidance Note 5.2 – Formatting Data,' Guidance Note 5.4 – Adding Metadata and Schema,' and 'Guidance Note 5.5 – Managing Data Quality' within the Smart Data Implementation Guide.

6.2 Entities should build compliance with the Smart Data Standards into the specifications and contracts for all systems they build or procure.

Principle 7: 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. This means that:

7.1 Entities should measure, monitor and manage the quality of their data in order to ensure it is fit-for-purpose to support both the initial intended use and also potential re-use.

7.2 Entities should commit to continuous improvement in data quality, prioritizing quality improvements that are important for data users.

Principle 8: 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. This means that:

8.1 Entities should implement systems that give real-time, event-level data about what is happening across all of their systems, assets, and customer interactions.

8.2 Entities should deploy tools that enable rich visualization of data in order to facilitate more intuitive data analysis.

8.3 Entities should ensure that their staff have the skills and tools needed to analyze and interpret data to ensure that their decision-making and policy development is evidence-driven and that their services can be improved continuously.

Principle 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 smart data. This means that:

9.1 Entities should manage data more efficiently by taking full advantage of the shared services offered by the Federal and Emirate level data platforms.

9.2 Entities should participate actively in the collaborative governance arrangements for data publishing and data exchange established by Federal and Emirate level data authorities.

Principle 10: Continuous improvement

In order to ensure full implementation of the Smart 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. This means that:

10.1 Entities should actively manage a process of change, seeking to move their Entity from one in which data is locked in silos to one in which it is shared and managed for the benefit of all – with accountability for this change managed at the most senior level in the Entity.

10.2 Entities should develop a Roadmap showing how they will manage the transition to smart data in a phased and prioritized way, with prioritization driven by demand from data users.

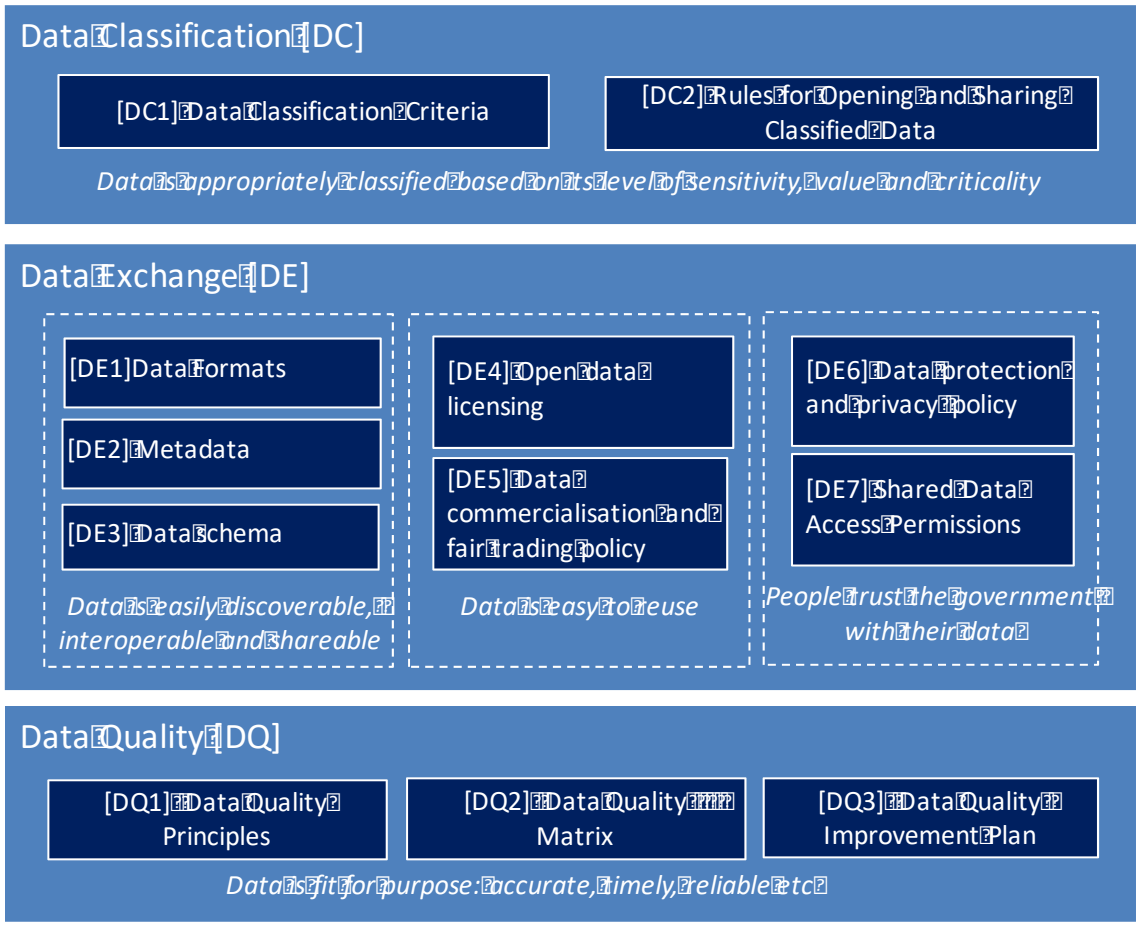
3. SMART DATA STANDARDS

3.1 Overview

The United Arab Emirates Smart Data Standards support implementation of the Smart Data Principles set out in Section 2 of this document, by specifying common requirements to enable data classification, data exchange, and management of data quality. Mandatory requirements are clearly identified and keep to the minimum levels necessary to secure the goals of the Smart Data Framework, while allowing each Entity maximum flexibility on implementation.

The Smart Data Standards are grouped into three, as illustrated below:

- Data Classification Standard
- Data Exchange Standard
- Data Quality Standard.



Each of the three standards contain a set of Smart Data Specifications: a documented specification that includes mandatory and recommended requirements for Entities. These are of two types:

- **Dataset Processing Specifications**: these apply at the level of an individual dataset, specifying how that dataset should be classified, formatted, and described in order to conform with the Smart Data Standards.

- **Data Management Specifications**: these specify the business rules and operating principles that Entities should follow as they manage the data

These standards and specifications cover both structured data and unstructured data. Where the requirements of a standard are intended to apply to unstructured data, this is explicitly stated in the text of the requirement.

3.2 Structure of specifications

Individual specifications within each of the three standards are presented within a common format, as illustrated below.

Specification Number (eg DQ1, DE3)	Name: the title used to refer to this specification within other documents	
Specification type	☑ Dataset Processing Specification	☐ Data Management Specification
Purpose	What this specification should help Entities achieve	
When to use	The point in an Entity's Smart Data uptake and maturity stage that this specification should be used	
Responsibility	Lead role within an Entity responsible for overseeing conformance to this specification	
Requirements		
Mandatory	Mandatory Requirement Code	A list of the mandatory requirements for conforming to this specification
Recommended	Recommended Requirement Code	A list of the recommended best practices for comforming to this specification
Specification Inter-dependencies	List of other specifications within the UAE Smart Data Standards that this specification depends upon	
References to Smart Data Implementation Guide	A reference to tools and guidance to support delivery of this specification that are contained within the Smart Data Implementation Guide.	
External References	Reference to sources and documents that are external to the UAE Smart Data Framework	
Version History	Version history description	

3.3 Data Classification Standard

Introduction to the Data Classification Standard

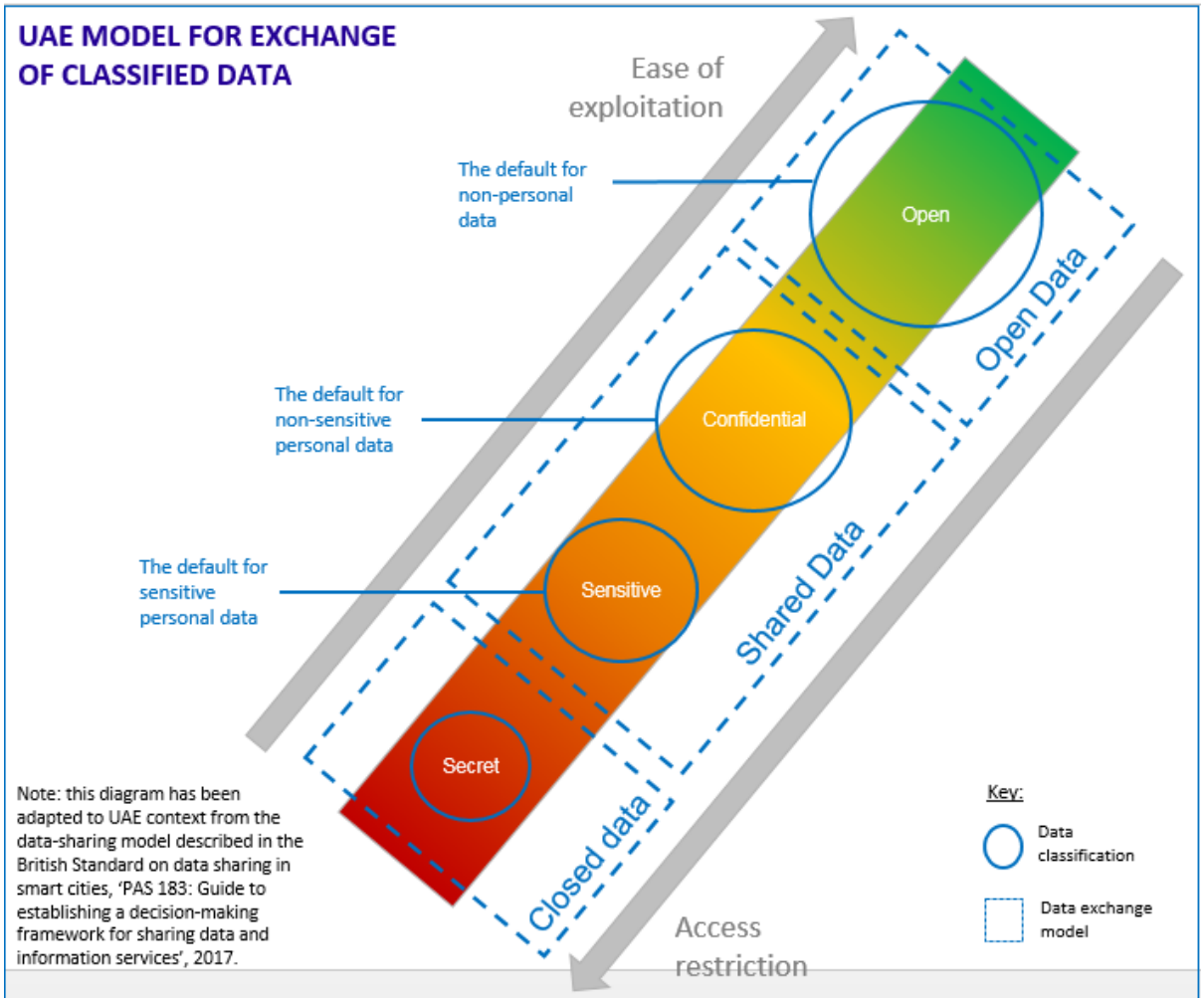
The purpose of the Data Classification Standard is to enable significantly greater levels of open data publication and data exchange between Entities, while at the same time preserving high levels of privacy and security.

Federal legislation requires UAE Government Entities to classify data into four different classes: *Open*, *Confidential, Senstive* and *Secret*. As illustrated below, this Standard supports that policy by setting out more detailed specifications on:

- The criteria Entities should use when classifying data into these four classes – as set out in **DC1 Data Classification Criteria**
- The consequences of that classification for the model that Entities should use when publishing or sharing data across digital channels – **[DC2] Rules for Opening and Sharing Classified Data.**

A visual summary of the data exchange model that is supported by this Data Classification Standard is shown on the following page. This illustrates how the four classes of data (*Open*, *Confednetial, Sesntive and Secret*) map against the modes of digitally-enabled data exchange which should typically be associated with each classification:

- **Open data:** data that is publicly shared and published online with minimal restrictions
- **Shared data**: data that is shared digitally with other government entities, for example through the Smart Data Electronic Platform or Government Service Bus, and potentially also (subject to appropriate privacy protection and consent mechanisms) with private-sector entities
- **Closed data:** data that cannot be shared digitally with anyone else on or off the Smart Data Electronic Platform.


UAE MODEL FOR EXCHANGE OF CLASSIFIED DATA
Ease of exploitation
The default for non-personal data
The default for non-sensitive personal data
The default for sensitive personal data
Open
Confidential
Sensitive
Secret
Open Data
Shared Data
Closed data
Access restriction
Note: this diagram has been adapted to UAE context from the data-sharing model described in the British Standard on data sharing in smart cities, 'PAS 183: Guide to establishing a decision-making framework for sharing data and information services', 2017.
Key:
Data classification
Data exchange model

Data Classification

DC1	Data Classification Criteria	
Specification type	☑ Dataset Processing Specification	☐ Data Management Specification
Purpose	This Standard sets out the duty of Government Entities to classify structured and unstructured data against one of four classes: *Open, Confedential, Sensitive,* and *Secret.*	
When to use	This Standard should be used: • Before any dataset is published as open data • Before exchanging any shared data with another Entity • In a phased and prioritised way over time for existing datasets owned or managed by the Entity, to ensure that the classification of all Entity data is conformant with this Standard. • Whenever an Entity creates a new dataset.	
Responsibility	The responsibility for ensuring that an individual dataset is conformant with this Standard lies with the Data Custodian of that dataset. The Data Management Officer is accountable for ensuring that the Entity as a whole conforms with this Standard.	
Requirements		
Mandatory	DC1.1	Government Entities should classify each dataset they manage as one of four classes (*Open, Confedential, Sensitive,* and *Secret)* based on the most sensitive (most highly classified) item in the dataset, and include that classification in the metadata for the dataset.
	DC1.2	Where data has been categorized as *Confidential or Sensitive*, the Government Entity should consider the scope for creating a summary, redacted version, extract, or other derivative of the data, which would have value as open data but avoid the negative effects identified. This new data set should then be classified as Public Data.
	DC1.3	Entities should review and evaluate dataset classifications on a regular basis to ensure that the assigned classification levels remain appropriate in light of changes to legal and contractual obligations or other relevant changes.
Recommended	DC1.4	It is recommended that Entities classify a dataset's Metadata. The potential sensitivity of the metadata itself should be considered in order to determine whether to disclose various characteristics of the original dataset.
Standard Inter-dependencies	The rules that should apply to the publication of data once classified are set out in **[DC2] Rules for Opening and Sharing Classified Data**.	
References to Implementation Guide	The classification criteria and the process for classifying data in accordance with this Specification are described in **Guidance Note: 5.1 Classifying data** of the Smart Data Implementation Guide.	
External References	Once classified as *Confidential, Sensitive,* or *Secret*, datasets should be managed in accordance with the requirements for physical security and IT security set out in	

	'Regulation of Information Security at the Federal Entities of UAE Cabinet Resolution' No. (21), 2013.
Version History	V1.0

Rules for Opening and Sharing Classified Data

DC2	Rules for Opening and Sharing Classified Data	
Specification type	☐ Dataset Processing Specification	☑ Data Management Specification
Purpose	This Standard sets out the consequences that the classificaton of a dataset has for Entities' ability to publish (for Open) or exchange (for Shared) that data.	
When to use	This Standard should be used: • Before any dataset is published as open data • Before exchanging any shared data with another Entity • In a phased and prioritised way over time for existing datasets owned or managed by the Entity, to ensure that the classification of all Entity data is conformant with this Standard.	
Responsibility	The responsibility for ensuring that an individual dataset is managed in conformance with this Standard lies with the Data Custodian of that dataset. The Data Management Officer is accountable for ensuring that the Entity as a whole conforms with this Standard.	
Requirements		
Mandatory	DC2.1	Each Entity should develop a plan for ensuring that all of their datasets are correctly classified against the **[DC1] Data Classification Criteria** specification and share this plan with the Federal Data Management Office.
	DC2.2	Before any dataset is published or exchanged it should be classified as one of *Open, Confedential, Sensitive,* and *Secret*: • To publish data as open data, it should be classified as *Open* Data • To digitally share or exchange data with another Entity, it should be classified as either *Open, Confedential and Sensitive* • Data classified as *Secret* should not be published or exchanged digitally on the electronic platform, and should only be exchanged with authorised individuals on a "need to know" basis and subject to confidentiality obligations determined necessary by the Entity responsible for that data.
	DC2.3	Once classified as *Open, Confedential and Sensitive,* , the dataset should be managed in accordance with the requirements for physical security and IT security set out in 'Regulation of Information Security at the Federal Entities of UAE Cabinet Resolution' No. (21), 2013.
	DC2.4	Establish management systems to ensure that, whenever it creates a new dataset, or for the first time either shares an existing data set with another

		Entity or publishes it as Open Data, the dataset is correctly classified against the criteria in this Standard
	DC2.5	Entities receiving shared data should adhere to the requirements of the original classification, unless they anonymize or otherwise modify the data such that a new dataset is created and thus, as the custodian of this new dataset, the recipient entity is required to classify the new dataset
Standard Inter-dependencies	• Datasets classified as *Open* and then published as Open Data should be published under **[DE4] Open data licensing**. • Datasets classified as *Open* and then published for which users are charged a fee will only be permitted under exceptional circumstances and as described in **[DE5] Data Commercialisation and Fair Trading** • Data classified as *Confidential or Sensitive* and then exchanged digitally with another Entity should be subject to documented rules on who can access the data, for what purpose and to what level of access, as required by **[DE6] Shared Data Access Permissions**.	
References to Implementation Guide	• **Guidance Note 3**: **Prioritization** describes how to prioritize which datasets to classify first. • A best practice process for classifying data in accordance with this Specification is described in **Guidance Note: 5.1 Classifying data**	
External References	Once datasets have been classified *Open, Confedential, Sensitive,* and *Secret*, they should be managed in accordance with the requirements for physical security and IT security set out in 'Regulation of Information Security at the Federal Entities of UAE Cabinet Resolution' No. (21), 2013.	
Version History	V1.0	

3.4 Data Exchange Standards

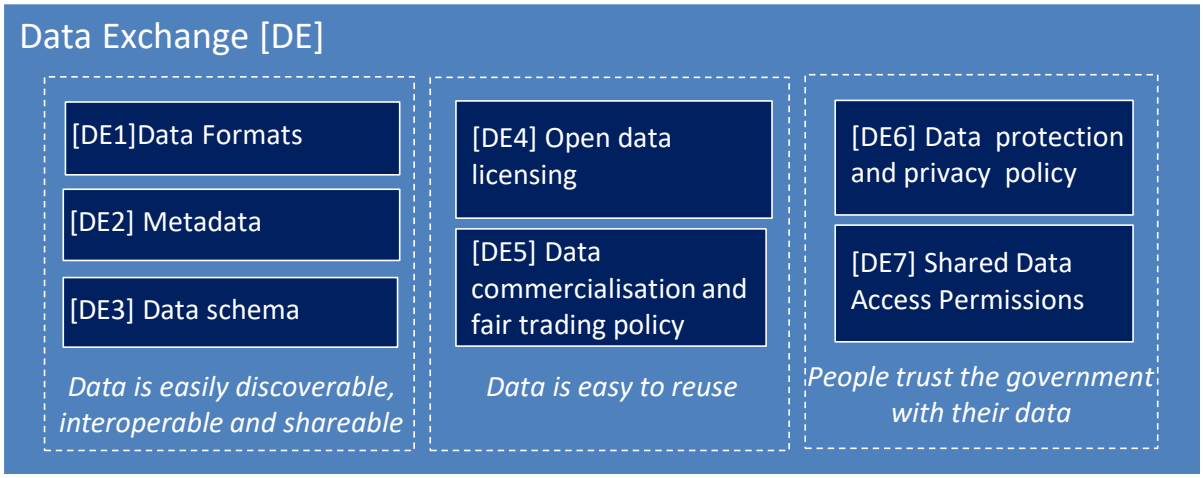
Introduction to the Data Exchange Standard

To exchange data effectively across entities, the data must be discoverable, reliable, and re-usable. To achieve this, Entities should act as publishers of data - not simply producing data for their own internal purposes, but having processes and standards in place to ensure their data is reusable by external Entities by default.

The Data Exchange Standard supports this change by setting out specifications in three areas, as illustrated below, in order to:

- **Ensure data is easily discoverable, interoperable and shareable**
  These include specifications on **metadata** to ensure data can be searched for effectively and to help users understand the content and context of the data, and on **data formats** and **data schema** to facilitate interoperability of Entity data with external data.
- **Ensure that the rights for data re-use are consistent and communicated clearly, in order to make it easy to re-use**
  These include specifications on **licensing** and **commercialization** to give users of data confidence in the terms under which they can utilize the data.

- **Ensure that people trust the government with their data**
  These include specifications on **data protection, privacy, access rights,** and **permissions** to ensure that access to data is appropriate, conformant, and protects individual privacy.



Data formats

DE1	Data formats	
Specification type	☑ Dataset Processing Specification	☐ Data Management Specification
Purpose	This standard sets out the minimum mandatory requirements and best practice recommendations for what formats to use when exchanging or publishing data. A *data format* is a standard way in which information is encoded for storage and transmission by computers. It specifies the way in which data is arranged in such a way that the data can be read by software applications. Using an appropriate open format is particularly relevant for ensuring ease of access by data users, and subsequent interoperability with internal and external data.	
When to use	When creating new datasets and when preparing data for publication or exchange with another Entity.	
Responsibility	Data specialists who have technical oversight and ownership of the data.	
Requirements		
Mandatory	DE1.1	All structured data that is to be published as open data or shared with other Entities should be made available in appropriate (for type of data and intended use), structured, machine readable, open formats. This means that: • Tabular data should be published as CSV • Geospatial data as GeoJSON or KML • Other structured non-tabular data in an open standard where available, for example using: JSON, XML, RDF, GTFS.

		• Real-time data or data being used in real-time services should be made available via a well-documented API[4].
	DE1.2	Each record within a structured dataset should include a unique identifier for the subject of that record.
Recommended	DE1.3	For structured tabular data, it is recommended that in addition to providing a CSV file, Entities also publish the data in a single analytical spreadsheet tool (such as Excel or ODF spreadsheet) containing both data and all descriptive and machine readable Metadata.
	DE1.4	For unstructured data, Entities should assess the type of data contained to see whether it can be turned into structured data.
	DE1.5	High-value[5] reusable unstructured data should be published or exchanged as is, using open formats where these exist.
Standard Inter-dependencies	**[DQ1] Data Quality Standard** gives guidance on how to measure the quality of data formatting.	
References to Implementation Guide	**Guidance Note: 5.2 Formatting Data** of the Smart Data Implementation Guide provides ○ Advice on picking appropriate structured formats. ○ A sample process for choosing a format	
External References	• CSV (Comma Separated Values) guide: https://frictionlessdata.io/guides/csv/ • GeoJSON specification: http://geojson.org • KML specification: http://www.opengeospatial.org/standards/kml/ • JSON general information: http://json.org • JSON data exchange syntax: http://www.ecma-international.org/publications/files/ECMA-ST/ECMA-404.pdf • XML (Extensible Markup Language): https://www.w3.org/XML/ • RDF Primer: https://www.w3.org/TR/rdf11-primer/ • GTFS: https://developers.google.com/transit/gtfs/ • ODF – open document format: http://opendocumentformat.org/aboutODF/	
Version History	V1.0	

[4] Application Programming Interface: a set of definitions of the ways one piece of computer software communicates with another. A web API allows computer programs to dynamically query a dataset using the World Wide Web. For example, a dataset showing the locations of hospitals and doctor's surgeries may be made available for download as a single file (e.g. a CSV), or may be made available to developers through a Web API, such that a computer program could automatically retrieve a list of health addresses for a particular area and display it on an online map alongside other relevant public and private sector data.

[5] Entities can judge the value by using the benefit assessment criteria from **Guidance Note 4: Prioritisation criteria and process**.

Metadata

DE2	Metadata	
Specification type	☑ Dataset Processing Specification	☐ Data Management Specification
Purpose	Metadata is structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage an information resource. Metadata provides valuable context and meaning to data which dramatically increases the usability and discoverability of the data. This Specification sets out the requirement that metadata be added to data when creating, publishing, or exchanging datasets.	
When to use	For existing datasets, use this standard as part of the data compliance process. (in a prioritised order following Inventory and Prioritisation). Whenever creating a new dataset and before publication as open data or exchange as shared data with other Entities.	
Responsibility	Data Custodian	
Requirements		
Mandatory	DE2.1	Datasets should contain all mandatory metadata as specified in the UAE Smart Data Framework Implementation Guide – specifically 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.
	DE2.2	Metadata should be kept up to date along with the data associated with the dataset.
	DE2.3	If expressing the specified metadata fields in an API, electronic platform, or RDF representation of the dataset, Entities should use the dcat vocabulary (tools like CKAN can do this automatically, or follow online guides).
Recommended	DE2.4	It is recommended that Entities add all defined metadata to datasets including 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 data or intellectual property and associated terms of use.
	DE2.5	It is recommended that Entities include sector or topic-specific metadata and vocabularies that are not relevant to all government datasets, but serve the needs of specific data-using communities (e.g. health or transport sectors).
	DE2.6	It is recommended that Entities monitor and generate reports based on the metadata across all Entity datasets to track and review alignment with **[DQ1] Data Quality Principles** and strategic goals.
Standard Inter-dependencies	It is recommended that the metadata pull together or reference requirements and guidance found in **[DE1] Formats**, **[DE3] Schema,** and **[DQ1] Data Quality standards**.	

References to Implementation Guide	**Guidance Note 5.4** provides advice on how to apply metadata in conformance with this Specification.
External references	• Data Catalog Vocabulary (DCAT) specification: https://www.w3.org/TR/vocab-dcat/ • CKAN: https://ckan.org/ • Open Data Institute guide for marking up for dataset with DCAT: https://theodi.org/guides/marking-up-your-dataset-with-dcat
Version History	V1.0

Schema

DE3	Schema	
Specification type	☑ Dataset Processing Specification	☐ Data Management Specification
Purpose	This standard outlines the requirements for publishing a data schema. A schema is a formal description of the format of structured data, as well as a guarantee that future data releases will use the same format.	
When to use	When creating a new structured dataset and as part of **[DQ3] Data Quality** improvement over time.	
Responsibility	Data Specialist	
Requirements		
Mandatory	DE3.1	Structured datasets indicated as a high priority for publication or exchange by the Federal Data Management Office and requiring higher quality requirements should be published with a schema.
	DE3.2	Schemas should be published in a machine readable format (usually JSON).
	DE3.3	Primary Registry datasets should have a published schema and be validated against it.
Recommended	DE3.4	It is recommended that high value, structured and regularly updated datasets have a published schema.
	DE3.5	It is recommended that Entities publish the schemas for data which already have a schema in the metadata with the dataset. Most databases, KML or sector specific formats (like GTFS) will already have a schema.
	DE3.6	Where possible, data should use, and then reference in the schema, any international or local standard vocabularies (such as the ISO-3166-alpha-2 country codes or a Primary Registry of government entities).
Standard Inter-dependencies	The schema should be added to the dataset's **[DE2] Metadata** and align with the **[DE1] Format** standard.	

References to Implementation Guide	**Guidance Note 4: Prioritisation criteria and process** for establishing high value datasets **Guidance Note 5.4** provides advice on how to develop schema in conformance with this Specification.
Version History	V1.0

Open Data Licensing

DE4	Open Data Licensing	
Specification type	☑ Dataset Processing Specification	☐ Data Management Specification
Purpose	This specification sets out the requirements for licensing open data to ensure that the strategic goals of driving engagement and innovation around Open Data are realised. Data or information is open "if anyone is free to access, use, modify, and share it — subject, at most, to measures that preserve provenance and openness."[6]	
When to use	Whenever publishing open data.	
Responsibility	Director of Data	
Requirements		
Mandatory	DE4.1	All open data, both structured and unstructured, should have a clear open data license associated with it in the metadata. This license should: • Allow unrestricted access to the data • Allow the data to be adapted, modified, combined with other data and re-published or shared – free of charge and subject at most to the requirement for attribution • Explicitly allow the commercial use of data • Be published online, within the Entity's website or through a link to the Federal Open Data Licence.
	DE4.2	Government Entities should use the UAE Federal Open Data License or a UAE issued license which conforms to the requirements in [DE.4.1]
Recommended	DE4.3	It is recommended that the open data license be user friendly, clear, simple and visual –as the purpose of the license is to make clear the rights of the re-users and remove barriers to re-use, not to protect the rights of the publisher (although it is common for open data licenses to state that the Entity publishing the data accepts no liability over incorrect data).

[6] As per the Open Definition: http://opendefinition.org/od/2.1/en/

Standard Inter-dependencies	• **[DC1] Data Classification** describes the basis for classifying a dataset as Public Data. • All Public Data classified in accordance with that standard should be published as Open Data in the ways described in this Specification, unless a clear case for commercial publishing can be made in compliance with **[DE5] Data commercialization and fair trading**. • Conformant open datasets need to link to a conformant Open Data License in their **[DE2] Metadata.**
References to Implementation Guide	• **Guidance Note 5.6** provides advice on how to publish open data in conformance with this Specification. • The UAE Federal Open Data License is available at [link] and at **Appendix A to the UAE Smart Data Implementation Guide**.
External References	Open Definition and specification for an open license: http://opendefinition.org/od/2.1/en/
Version History	V1.0

Data commercialization and fair trading

DE5	Data commercialization and fair trading	
Specification type	☐ Dataset Processing Specification	☑ Data Management Specification
Purpose	The UAE Government is committed to publishing non-personal, non-sensitive data openly wherever possible, and as a general rule with the data provided for free under the terms of **[DE4] Open data licensing**. However, there may be limited circumstances in which it is in the public interest to make exceptions to this general rule, and to permit Government Entities to charge fees for either raw public data or value-added data services. The purpose of this Specification is therefore to set out the requirements Entities should meet in such cases, to ensure that any charges are set on a fair competitive basis with the private sector that encourages rather than crowds out private-sector investment in the UAE market for data services.	
When to use	Before taking a decision to charge fees either for raw data or for value-added data services.	
Responsibility	The Director of Data is responsible for implementation of this Standard across the Entity. The Data Custodian within the Entity who is accountable for a specific dataset will normally take the lead in developing the business case for any commercialisation of that dataset.	
Requirements		

Mandatory	DE5.1	All Government Entities proposing to charge a fee for data or data services should first seek approval from the Federal Data Management Office, setting out clearly why this is justified in line with the principles set out in this Standard: 1. Public interest 2. Fair competition 3. Fair pricing and conditions 4. Accountability.
	DE5.2	Public 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 for access to data that the Entity does not currently collect and/or that would require significant additional action and investment by the Entity to provide, then there may be a case for charging fees to data users in order to help finance this investment
Recommended	DE5.3	Semi-government Entities should also seek to follow the advice in this standard, which aims to maximise growth of the United Arab Emirates data economy through an integrated, consistent and pro-competitive approach to Public Data.
Standard Inter-dependencies	**[DC1] Data Classification** describes the basis for classifying a dataset as Public Data. Only Public Data classified in accordance with that standard may be commercialised in the ways described in this specification.	
References to Implementation Guide	**Guidance Note 5.6** provides advice on a best practice process to follow when seeking permission to apply data charges in conformance with this Specification.	
External references	• This Specification implements, within the specific national context of the UAE, the core principles on charging for Public Sector Information which were agreed by the 32 countries of the OECD in 2008 and regularly re-committed to since. • These principles, and resources to support their delivery – including case studies and evaluation evidence on the benefits countries are achieving through implementation of such an open and pro-competitive approach to public data – are set out here: OECD Recommendation on Public Sector Information (PSI).	
Version History	V1.0	

Data protection and privacy

<table>
<tr><th>DE6</th><th colspan="3">Data protection and privacy</th></tr>
<tr><td>Specification type</td><td colspan="2">☐ Dataset Processing Specification</td><td>☑ Data Management Specification</td></tr>
<tr><td>Purpose</td><td colspan="3">The purpose of this Specification is to:
Ensure that people and businesses in the UAE have trust and confidence that their data is ethically used and enjoys strong levels of protection and privacy
Build a culture of privacy awareness and responsibility within officials dealing with data
Ensure personal data management and infrastructure is resilient and secure
Ensure data is only used in ways that meet documented ethical standards
Enable uniformity and consistency in decision making in relation to data protection and privacy.</td></tr>
<tr><td>When to use</td><td colspan="3">Across all stages of the data management lifecycle: creating, processing, analysing, storing, exchanging and re-using data.</td></tr>
<tr><td>Responsibility</td><td colspan="3">The Director of Data has responsiblity for ensuring that the Entity has the systems, infrastucture and controls necessary to comply with this Standard specification, and that these operate effectively and consistently.
The Data Custodian within the Entity who is accountable for a specific dataset is responsible for ensuring that the requirements of this specification are met in relation to that dataset.</td></tr>
<tr><td colspan="4">Requirements</td></tr>
<tr><td rowspan="3">Mandatory</td><td>DE6.1</td><td colspan="2">All Entities should work towards achieving, across all personal and commercial datasets for which they are responsible, full compliance with the Data Privacy Principles set out in this specification:
Consent
Transparency
Purpose
Proportionality
Personal access and control
Security
Sectoral compliance
Documentation
Awareness
Accountability</td></tr>
<tr><td>DE6.2</td><td colspan="2">Government Entities should publish these Data Privacy Principles on their websites, and provide complaints and redress mechanisms for data subjects who believe they are failing to manage their data in accordance with the above principles</td></tr>
<tr><td>DE6.3</td><td colspan="2">Government Entities should assess where there are gaps in how their current data management practices conform with these Data Privacy Principles, develop plans to close these, and share these plans with the Federal Data Management Office.</td></tr>
</table>

Recommended	DE6.4	Semi-government and Private Sector Entities are also recommended to embed the UAE Privacy Principles within their own data management practices in order to build a cohesive national system of trusted data exchange within a strong framework of data protection and privacy.
Standard Inter-dependencies	• **[DC1] Data Classification** sets out the criteria that Entities should apply when determining whether a dataset contains Personal Information or Commercial Information of the sort that is covered by the privacy requirements of this specification. • The access rights that should be attached to a dataset after application of the principles in this Standard should be documented through either **[DE5] Open Data License** or **[DE7] Shared Data Access Permissions**, and through **[DE2] Metadata**.	
References to Implementation Guide	**Guidance Note 5.3** provides a more detailed description of the principles in this Standard, together with advice on a best practice process to follow when applying these to an individual dataset.	
Version History	V1.0	

Shared data access permissions

DE7	Shared data access permissions	
Specification type	☐ Dataset Processing Specification	☑ Data Management Specification
Purpose	This Specification describes principles and practices for permitting access to Confidential and Sensitive Data, in a way that facilitates cross-government service integration and complies with the principles of **[DE6] Data protection and privacy**.	
When to use	When preparing *Confidential or Sensitive* data for exchange with another Entity for the first time, Entities should document who is permitted to have what level of access to the data in compliance with this Specification. Entities should then apply this Specification when responding to future requests for additional access permissions.	
Responsibility	The Director of Data has responsiblity for ensuring that the Entity has the systems, infrastucture and controls necessary to comply with this specification, and that these operate effectively and consistently. The Data Custodian within the Entity who is accountable for a specific dataset is responsible for ensuring that the requirements of this specification are met in relation to that dataset.	
Requirements		
Mandatory	DE7.1	Government Entities should follow the five Access Permission Principles described in this specification whenever they share their *Confidential or Sensitive* data with a third party:

		1. Entities should facilitate cross-government sharing of their data 2. Data sharing should protect personal and commercial privacy 3. Use of the Smart Data Electronic Platform 4. Data Sharing Access Permissions should be documented 5. Access to shared data should be secured and audited
	DE7.2	Government Entities should assess where there are gaps in how their current data management practices conform with these UAE Access Permissions Principles, develop plans to close these, and share these plans with the Federal Data Management Office.
	DE7.3	Government Entities should respond promptly in writing to requests for data sharing from other Entities, and notify the Federal Data Management Office of all such requests.
Recommended	DE7.4	Entities are also recommended to make the audit functionality that is required under Access Permission Principle [6] openly available for use by individual data subjects.
Standard Inter-dependencies	• **[DC1] Data Classification** sets out the criteria that Entities should apply when determining whether a dataset should be classified as *Confidential or Sensitive*, and is thus subject to this specification. • The access rights that are set out in the Shared Data Access Permissions required by this standard should comply with the requirements of **[DE6] Data protection and privacy**.	
References to Implementation Guide	**Guidance Note 5.3** provides advice on a best practice process to follow when a) documenting an initial set of Shared Access Data Permissions for a dataset and b) responding to requests from other Entities for additional Shared Data Access Permissions .	
Version History	V1.0	

3.5 Data Quality Standard

Introduction to the Data Quality Standard

Data Quality is defined as the degree to which the characteristics of the data meet and achieve the requirements of being appropriate for purpose for use or reuse in operational delivery, decision making, analytics, planning and knowledge sharing (ISO 9000: 2015 clause 3.6.2).

By improving data quality, the objective of the Data Quality Standard is to deliver increased levels of:

- **Reliability** – meaning that data is accurate and complete, and decisions can be made on the basis of that data without additional checks and reviews. This increases confidence in the data.
- **Effectiveness** – meaning the Entity is better equipped to deliver on its operational and strategic objectives, as well as the objectives of the Smart Data program.
- **Efficiency** – ensuring services are delivered with fewer errors, faster and at a lower cost to the Entity. Users and citizens are empowered to get the service outcomes they desire and the information they seek quickly and easily.

To achieve this, the Data Quality Standard specifies quality requirements in three areas, as illustrated below:

- **[DQ1] Data Quality principles**, which describe seven key principles to increase data quality in the United Arab Emirates, and a set of core minimum quality requirements which Entities should meet as they work towards these principles.

- **[DQ2] Data Quality Maturity Matrix**, which provides a common tool for measuring a dataset's level of quality against these principles.

- **[DQ3] Data Quality Improvement Plan**, which requires Entities to develop and manage plans for progressively improving their data quality in line with the principles.

These three specifications are all Data Management Specifications – that is, they focus on the business rules and operating principles that Entities should apply to the way they manage data quality. Elements of **[DQ1] Data Quality Principles** also describe technical features of quality within the dataset itself. These technical dataset quality features are key to enabling data exchange between Entities, and are covered in more detail within three of the Dataset Processing Specifications set out in the Data Exchange Standard: **[DE1] Data Formats**, **[DE2] Metadata** and **[DE3] Schema**.

Data Quality Principles

DQ1	Data Quality Principles	
Specification type	☐ Dataset Processing Specification	☑ Data Management Specification
Purpose	The purpose of this specification is to ensure that Data Quality in the UAE is 'appropriate for purpose' as defined by ISO[7]. It sets out good practice principles on data quality for use by all entities in the UAE.	
When to use	Across all stages of the data management lifecycle: creating, processing, analysing, storing, exchanging, and re-using data.	
Responsibility	The responsibility for ensuring that an individual dataset is conformant with this Standard lies with the Data Custodian for that dataset. Overall accountability for ensuring that the Entity as a whole complies with the Standard lies with the Data Management Officer, reporting to the Entity's Director of Data.	
Requirements		
Mandatory	DQ1.1	Government Entities should embed the following Data Quality Principles in their data management practices, and in those of third parties contracted to manage data and services on their behalf: 1. Ownership and authority 2. Accessibility 3. Accuracy 4. Descriptiveness 5. Timeliness 6. Completeness 7. Validation
	DQ1.2	Government Entities should meet a minimum set of Core Data Quality Requirements detailed as mandatory requirements in **[DE1] Data Formats**, **[DE2] Metadata,** and **[DE3] Schema.**
Recommended	DQ1.3	Private-sector Entities are also recommended to embed the Data Quality Principles within their own data management practices, in order to help build a cohesive national system of trusted data exchange with high levels of quality.
Standard Inter-dependencies	• The extent to which a dataset currently conforms with the Data Quality Principles can be assessed using [**DQ2] Data Quality Maturity Matrix**. • Actions Entities should take to conform with the Data Quality Principles are set out in **[DQ3] Data Quality Improvement Plan**.	
References to Implementation Guide	**Guidance Note 5.5** provides advice on a best practice process to ensure that a dataset conforms with the Data Quality Principles set out in this Specification.	

[7] ISO 9000, 2015, clause 3.6.2: Data quality is defined as the degree to which the characteristics of the data meet and achieve the requirements of being appropriate for purpose for use or reuse in operational delivery, decision making, analytics, planning and knowledge sharing.

External references	ISO 9000, 2015, clause 3.6.2: Data quality is defined as the degree to which the characteristics of the data meet and achieve the requirements of being appropriate for purpose for use or reuse in operational delivery, decision making, analytics, planning and knowledge sharing.
Version History	V1.0

Data Quality Maturity Matrix

DQ2	Data Quality Maturity Matrix	
Specification type	☐ Dataset Processing Specification	☑ Data Management Specification
Purpose	The Standard provides a common basis for measuring and comparing the quality of datasets across all seven of the **[DQ1] Data Quality Principles**.	
When to use	When undertaking a data quality audit of an individual dataset. For use in providing reports and analytics and benchmarking of data quality between entities.	
Responsibility	Data Management Officer and Data Custodians.	
Rrequirements		
Mandatory	DQ2.1	Data Custodians should use this standard to assess current levels of quality in their data, and as the basis for agreeing target levels of future quality with data users utilising the Data Quality Maturity Matrix that defines, for each of the seven Data Quality Principles, five levels of maturity: • **Level 1: Initial** – unmanaged data, no owner, no open format, no metadata, etc. • **Level 2: Partially conformant** – the dataset has an identified owner and is making progress towards conformance with the Data Quality Standard • **Level 3: Conformant** – the dataset meets all core requirements of the Data Quality Standard • **Level 4: Improving** – the dataset meets all core requirements and is also implementing additional good practices • **Level 5: Optimizing** – data quality fully meets the needs of current and potential future users, with clear systems for driving continuous improvement.
	DQ2.2	The Data Management Officer should draw together Data Quality Assessments using this matrix from data across the Entity, to give an overall assessment of data quality throughout the Entity.
Recommended	DQ2.3	Private-sector Entities are recommended to use this specification.

Standard Inter-dependencies	The Data Quality Maturity Matrix measures conformance of a dataset with the **[DQ1] Data Quality Principles**.
References to Implementation Guide	**Guidance Note 5.5** provides advice on how to use the Data Quality Matrix.
Version History	V1.0

Data Quality Improvement Plan

DQ3	Data Quality Improvement Plan	
Specification type	☐ Dataset Processing Specification	☑ Data Management Specification
Purpose	This Standard sets out the requirements that Entities should meet as they develop and manage plans for progressively improving their data quality in line with the **[DQ1] Data Quality Principles**.	
When to use	Across all stages of the data management lifecycle: creating, processing, analysing, storing, exchanging and re-using data.	
Responsibility	The Data Management Officer is responsible for conforming with this Specification when developing a Data Quality Improvement Plan for the Entity as a whole. Data Custodians are responsible for applying this Specification to the datasets for which they are accountable.	
Requirements		
Mandatory	DQ3.1	Each Entity should develop an **Entity-level Data Quality Improvement Plan** setting out how the Entity will implement the **[DQ1] Data Quality Principles**, and share their Plan with the Federal Data Management Office. These Plans should be **prioritized**, **baselined**, **user-focused**, **SMART**, **managed,** and **reported**.
	DQ3.2	For all priority datasets, Entities should undertake Data Quality Audits, using the **[DQ2] Data Quality Maturity Matrix**
	DQ3.3	For all priority datasets, the Entity should develop clear statements of **Data Quality Requirements** that: • Are evidence-based • Reflect the documented quality needs of users • Set measurable targets for quality improvement.
	DQ3.4	For all priority datasets, Entities should document a **Dataset-level Data Quality Improvement Plan** for that dataset, including measurable targets for quality improvement.

	DQ3.5	Government Entities should establish **systems to track and report** on data quality status, and how this is performing against targets in the Data Quality Improvement Plan.
Recommended	DQ3.6	Government Entities are recommended, when establishing systems in conformance with [DQ3.5], to establish reporting which is **automated and managed in real-time.**
	DQ3.7	Private-sector Entities are recommended to use this specification.
Standard Inter-dependencies	The Data Quality Improvement Level should improve performance against the **[DQ1] Data Quality Principles**, as measured by the **[DQ2] Data Quality Maturity Matrix.**	
References to Implementation Guide	**Guidance Note 5.5** provides advice on a best practice process for reviewing a dataset against the Data Quality Principles, ensuring minimum quality standards are met and then improving quality over time.	
Version History	V1.0	

APPENDIX A - GLOSSARY

Term	Definition
API	Application Programming Interface: a set of definitions of the ways one piece of computer software communicates with another. A web API allows computer programs to dynamically query a dataset using the World Wide Web. For example, a dataset showing the locations of hospitals and doctors surgeries may be made available for download as a single file (e.g. a CSV), or may be made available to developers through a Web API, such that a computer program could automatically retrieve a list of addresses for a particular area and display it on an online map alongside other relevant public and private sector data.
Cataloguing	The process of adding metadata to datasets listed in an Entity's Data Inventory. For the UAE Smart Data Framework this includes classification, finalizing format, adding metadata and a schema, and reviewing data quality.
Closed data	Government data that is highly secured, confidential and cannot be shared outside a government body or shared electronically.
Sensitive Data	Within the UAE Model for Exchange of Classified Data, Sensitive Data is a type of Shared data, and the second highest level of classification overall. It is less highly classified than Secret Data (which is the only level of Closed data), yet more highly classified than Confidential Data (which is the other level of Shared Data).
Conformance	Fulfilment of a requirement specified in a documented standard or specification.
Country	the United Arab Emirates (UAE)
Data	A structured or unstructured set of datum, facts, concepts, instructions, information, observations or measurements that shall be in the form of numbers, letters, symbols, images, maps or any other form, in a manner that allows interpretation, exchange or manipulation by individuals or computers
Data access permission	The permit and its conditions under which the shared data may be accessed by any authorized entity or person.
Data Classification Standard	The standard by which datasets and unstructured data can be classified into *Open, Confedential, Sensitive,* and *Secret* data which then impacts whether the data can be published as Open Data, exchanged between Entities as Shared Data or should be fully Confidential as Closed Data.

Data Custodian	A Data Custodian has business responsibility over their data. They should understand the value and risks associated with their data so that they can effectively prioritize, classify, and catalogue it. They will be responsible for determining whether the data should be Open or Shared and setting out the access permission rules.
Data exchange	Sharing or providing data access to a different entity than the one producing and initially using the data.
Data Governance	is a system of decision rights and accountabilities for information-related processes, executed according to agreed-upon models which describe who can take what actions with what information, and when, under what circumstances, using what methods.
Data inventory	An inventory or list of the datasets controlled or owned by an Entity.
Data Management	Refers to the disciplines and techniques to manage data as an asset.
Data Management Officer	The Data Management Officer (or DMO) is the delivery and operational lead for an Entity's data management activities. They could report to and deputize for the Director of Data and lead on coordinating the required change management processes to ensure conformance with Smart Data Framework standards.
Data Modelling	The creation of a model or overall description of the data in a system or used in a business process.
Data prioritization	The process of deciding which datasets should be prepared for publication or exchange, and in what order, within an Entity. It is recommended this is done according to a series of criteria which assess each dataset against the value and benefit of publication and readiness for publication described in the Implementation Guide.
Data provider	Any governmental, semi-governmental, or private sector entity, or any natural person who offers the data in any form, in a way that does not conflict with the laws in force in the United Arab Emirates.
Data publication	The process of making data available to others, through publication on the web, electronic platform, Government Service Bus or via an API.
Data Quality Audit	An assessment of the quality level of a specific dataset against the Data Quality Maturity Framework or an Entity-wide assessment of data quality practices against the Data Quality Principles found in the UAE Smart Data Standards.
Data Quality Requirement level	A specification for the data quality requirements of a dataset (as compared to the Data Quality Principles directly or by using the Data Quality Maturity Matrix), relevant to the current use of the data or potential use of the data. Should be informed by consultation with existing and potential users.
Dataset	A collection of data that it makes sense to group together, along with the metadata and schema that describes it. Each Entity identifies the datasets specific to supporting the needs of their respective mission or business functions. Note that a dataset is a deliberately flexible concept. A given dataset may represent an entire database consisting of multiple distinct entity classes, or may represent a single table in a database, or a map.

Data Specialist	A role with technical responsibility over data, typically within IT or database administrator teams. They will facilitate between the IT and business teams and ensure that the data for which they are responsible meets the format and quality requirements in the UAE Smart Data Standards.
Data sprints	A program-managed sequence of data publishing cycles, in which batches of data (starting with the highest priority data) are catalogued by adding metadata, an appropriate format, a schema if applicable, and any data quality related changes or responsibilities put in place.
Data Subject	Any person whose personal data is being collected, held or processed
Data User	Any entity or person wishing to take advantage of and use open or shared data in accordance with the terms and conditions on which such data are made available.
Decisions	Includes instructions, directives and regulations issued by the Committee or the Office regarding the performance of tasks and responsibilities.
Director of Data	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 that strategic goals are realized. Ideally, the Director of Data should be a member of the Entity's management board; as a minimum, they must 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.
Entity	Any organization or body defined as any of the following within this document: • Federal Government Entities (FGEs) • Local Government Entities • Semi-Government Entities • Private Sector Entities
Entity Roadmap	A time-based plan for the Entity as a whole to mature its data management practices and meet the requirements and recommendations of the UAE Smart Data Framework.
Federal Government	Government of the United Arab Emirates.
Federal Government Entity	Any ministry, authority, directorate, public body, independent body, public entity, federal government council, or any other government or public entity of the United Arab Emirates federal government
Format	A standard way in which information is encoded for storage and transmission by computers. It specifies the way in which data is arranged in such a way that the data can be read by software applications.
Government Entities	Ministries, bodies and entities of the Federal Government as well as directorates, bodies and institutions of the Local Government.
Government data	Electronic or non-electronic data or information of or belonging to the Federal Government or local governments of the Emirates of the United Arab Emirates, the public or federal bodies, or the local public institutions.

Identifiable information	Any information or personal data that reveals the identity of a living individual or natural person
Individual	For purposes of jurisprudence, an individual is a human entity or natural person
Information and Knowledge	Any useful results that are derived from data processing, that are also used for purposes of the strategy and the policy.
Intellectual Property Rights	Innovations by the mind, such as inventions, literary, artistic, or scientific works, designs, logos, names and images used in trading. Intellectual property is also legally protected by such rights, e.g. patents, copyright and trademarks that enable people to gain recognition or financial benefit from their innovation or invention. Through establishing a proper balance between the interests of innovators and the those of the general public, the system of intellectual property rights aims at creating an environment that promotes the prosperity of creativity and innovation.
Legal Person	A non-human entity recognized as a legal entity having distinct identity, legal personality, and duties and rights. In other words, an entity created by law that is treated as a person for limited legal purposes - corporations, for example. Legal persons can sue and be sued, own property, and enter into contracts. Also called artificial person, juridical entity, or juristic person.
Local Government	Governments of the member Emirates in the Federation.
Local Government Entity	Any entity that is administratively and financially appended to the local government of the Emirate.
Machine readable format	A format which can be read and correctly understood by machines. This means that it uses the symbols, rules, or conventions correctly and unambiguously and conforms to an existing standard.
Metadata	Structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage a data resource.
Natural Person	A natural person is a person (in legal meaning, i.e., one who has its own legal personality) that is an individual human being, as opposed to a legal person, which may be a private or public organization
Open by Default	The concept of non-personal data being available for general dissemination, unless justification is in place for this to be prevented.
Open data	Data published by Entities to be shared with the public freely or with minimal restrictions in order to maximize public participation and stimulate creativity, innovation, and economic growth.
Open format	Generally, this refers to an open standard format (where the specification for the format is accessible to all and openly licensed to be used by anyone). Consequently, it is a format which does not require the purchase of proprietary software to use or access the data.
Open Data License	A license which grants the user the right to use, modify and distribute the licensed data for any purpose, limited at most to the requirement for attribution.

Open Standard	An open standard or specification is one where: - All stakeholders have the same possibility of contributing to the development of the specification and public review is part of the decision-making process - The specification is available for everybody to study - Intellectual property rights related to the specification are licensed on fair, reasonable and non-discriminatory (FRAND) terms or on a royalty-free basis in a way that allows implementation in both proprietary and open source software
Personal data	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
Primary Registries	It is a register or electronic registers or databases containing data on the rights, transactions or status of individuals or companies. It is also any regulatory or administrative data to be referenced or relied upon as an accurate and reliable data source or necessary for the implementation of procedures and services.
Principles	The UAE Smart Data Framework is principle-based, setting out a number of principles to inform data management in the UAE. The term principles in this sense use the definition set out in the OASIS Transformational Government Framework: *"An enduring statement of values which can used on a consistent basis to steer decision-making by multiple stakeholders over the long term, and which are:* • *used to inform and underpin strategy;* • *understood, agreed and owned by stakeholders."*
Private Information	Information that is confidential and that relates to a natural person that would not be expected to be made publicly available without that person's choice or express consent, including but not limited to information that can identify the person, information regarding the person's family, information relating to the person's health, age, marital status, address, financial standing, religion, ethnic origin, political affiliations or opinions, criminal records, trade union memberships.
Private Sector Entities	Any entity or body that is not classified as a governmental, federal, local governmental or semi-governmental entity, including companies and institutions owned by individuals and private sector entities in the Emirate, including the authorities of the free zones in the UAE.
Public /Open Data	Within the UAE Model for Exchange of Classified Data, Open data, and the lowest level of classification overall.
Reference Data	Data that is the set of controlled values to be used in other specified areas. It is unlikely to be affected by the user's business or systems, but changes should be reflected in the system. A list of countries is an example of Reference Data.

Confidential Data	Within the UAE Model for Exchange of Classified Data, Confidential Data is a type of Shared data, and the third highest level of classification overall. It is less highly classified than Confidential Data (which is the other level of Shared data), and more highly classified than Public Data (which is the only level of Open Data).
Schema	A formal description for how something should look and behave. Includes the rules for what counts as conforming to the schema. In the context of data this could be a description and example of column headings and the type of data allowed to be in the rows underneath those headings and any validation rules which should be applied (for example, check it's a number from 0 – 100 with no spaces).
Semi-Government Entity	Any body, organization, bank or company to which the Government contributes.
Personal Data Examples	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.
Shared data	Government data that is shared digitally with other government entities, for example through the Smart Data Electronic Platform, or with Private-sector Entities. According to the UAE Model for Exchange of Classified Data within this document, data classified as Confidential or Sensitive falls into the category of Shared data.
Smart Data	Data that conforms with the requirements for data classification, data exchange, and data quality set out in the UAE Smart Data Standards.
Smart Data Electronic platform	The electronic data systems that allow electronic connectivity of services and/or collection, storage, analysis, exchange and/or availability of data from multiple sources between the connected parties according to given and defined privileges after being authenticated by a data provider in a secure network system. e.g. The Government Service Bus and UAE Open Data Portal are examples of systems within the Federal Smart Data Electronic Platform.
Structured Data	Structured data refers to data that is organized and constrained by a pre-defined model describing it. Structured data is often machine encoded but can equally be human readable. The structured nature of the data enables the data to be indexed and searched, and makes it more widely available, greatly increasing its potential value.
UAE Smart Data Framework	A suite of interrelated documents and document parts (UAE Smart Data Principles; UAE Smart Data Standards; UAE Smart Data Implementation Guide) which together provide a common basis for individual UAE Entities to manage data, in ways that provide maximum flexibility for each Entity to respond to their own business needs yet which also enable a common approach to data classification, exchange of data, and data quality.
UAE Smart Data Principles	The part of the UAE Smart Data Framework that sets out a clear set of principles to govern the creation, management, use and reuse of data in the UAE.

UAE Smart Data Standards	The part of the UAE Smart Data Framework that sets out the core standards around data classification, exchange and data quality to ensure UAE data is reliable, interoperable and fit-for-purpose.
UAE Smart Data Implementation Guide	The part of the UAE Smart Data Framework that sets out a set of supporting guidance and tools to help entities manage their data and implement the Smart Data Standards and Principles.
Unique identifier	With reference to a given (possibly implicit) set of objects, a unique identifier (UI) is any identifier which is guaranteed to be unique among all identifiers used for those objects and for a specific purpose. For example: serial numbers, URLs (domain addresses), codes from a registry, etc.
Unstructured data	Unstructured data refers to data that is not organized or constrained by a pre-defined model describing it. Unstructured data is often free text in documents, graphs and tables in spreadsheets, or video and audio files.
User-focused	An approach to the design and delivery of government data and services that is driven by the needs of their users rather than the government's organizational structures. Also known as customer-centric.
Secret data	Within the UAE Model for Exchange of Classified Data, Secret Data is Closed data, the highest level of classification overall).
Vocabulary	A vocabulary, is a classification system used to name or refer to things in a standardized way. For example, a standard way to classify educational establishments into types.