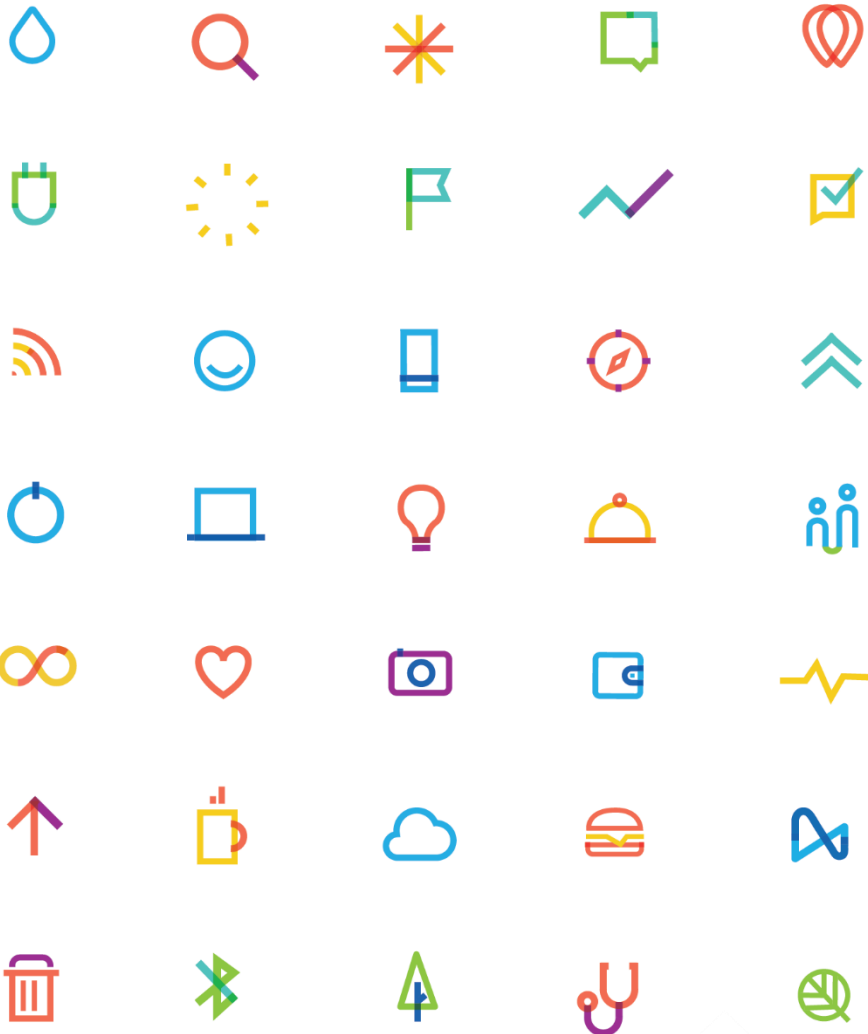


DUBAI DATA

DATA STANDARDIZATION GUIDEBOOK

Version 0.1 (OCT 2023)

DATA STANDARDIZATION GUIDEBOOK



Purpose:	To provide guidance and a clear details for data entry standards to be followed and implemented across Dubai government services.
Lead responsibility for implementing this standard in each Government Entity:	Data Management team or unit, reporting to the Data Leader.
When to use this guidebook:	We recommend using this guidebook at service form design or design update to decide the technical characteristics of the attributes.
Document owner:	Dubai Data Establishment
Licensing:	This document is published under the terms of a Creative Commons Attribution 4.0 International Licence in order to facilitate its re-use by other governments and private sector organisations. In summary this means you are free to share and adapt the material, including for commercial purposes, provided that you give appropriate credit to the Dubai Data Establishment as its owner and do not suggest the Dubai Data Establishment endorses your use.
Version	Version 0.1, dated 12 October 2023.



INTRODUCTION

Dubai Data Establishment set out in this document prescribe the requirements for data in the Government of Dubai. This guidebook is designed to provide government entities in Dubai with a set of standard data entry requirements for collecting data.

The purpose is to ensure that all government forms and systems are consistently collecting and storing data in a way that is accurate, consistent, and compliant with Dubai Data Law.



GUIDEBOOK OVERVIEW

This guidebook is created to detail the standard data entry requirements for each attribute used in the service forms of Dubai government entities. Examples are provided to illustrate the correct way to enter this data.

The standard data entry requirements detailed in this guidebook are in line with international data collection standardization and storage and can be considered as a recommendation by Dubai Data Establishment.



GUIDEBOOK BENEFITS

Having and implementing such a guidebook has several benefits for government entities in Dubai:

UNIFIED DATA STRUCTURE

set a standardized format for collecting, storing, and sharing data. This allows streamlining data management processes and reducing duplication, government entities can save time and resources and eliminate manual data entry and management. This can help to improve the overall efficiency of government operations and increase the quality of services provided to citizens.

IMPROVING DATA CONSISTENCY

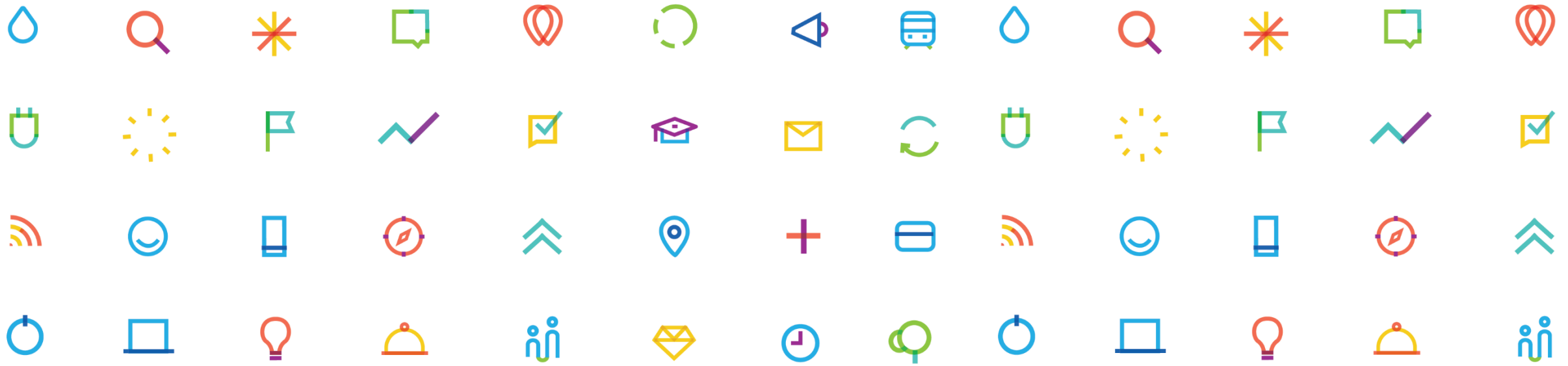
Consistent data entry across all forms and systems ensures that data can be easily compared and analyzed. This is essential for making informed decisions and identifying trends or patterns.

ENSURING DATA ACCURACY

By standardizing the way that data is collected and entered, the risk of errors and inaccuracies is greatly reduced. This is especially important for personal information, as incorrect or incomplete data can have serious consequences for individuals

ACHIEVING COMPLIANCE WITH DUBAI DATA LAW

Adhering to international and local standards helps to ensure that data is collected, stored, and processed in a way that is secure and compliant with best practices



DATA STANDARIZATION

IDENTITY INFORMATION

ATTRIBUTE	DATA TYPE	DESCRIPTION	NUMBER OF CHARACTERS	VALIDATION RULE	EXAMPLE	AVAILABILITY
FIRST NAME	Alphanumeric	The first name of the individual	Up to 50 characters	Must contain only letters (A-Z, a-z),), digits (0-9) and/or spaces	Ahmad	NA
LAST NAME	Alphanumeric	The last name of an individual	Up to 50 characters	Must contain only letters (A-Z, a-z),), digits (0-9), and/or spaces	Al Shamsi	NA
FULL NAME	Alphanumeric	The full name of the individual including 7 placeholders for holding the names according to the above rules of first and last name	Up to 500 characters	Must contain only letters (A-Z, a-z),), digits (0-9), and/or spaces	Ahmad Ali	NA
UAE NATIONAL ID	Numeric	A unique identification number assigned to an individual by the UAE government	15	Must contain only digits (0-9).	123456789123	NA
PASSPORT NUMBER	Alphanumeric	The number of an individual's passport	20	Must be a valid passport number format, can contain letters and numbers, and should match passport number in the ID or ID card	A12345678	NA
PASSPORT EXIPRY DATE	Date	The date on which an individual's passport expires	10	Must be a valid date in the format of DD/MM/YYYY and be in the future	31/12/2030	NA
PASSPORT ISSUE	string	The place where an individual's passport was issued	50	Must be Drop Down List	Dubai, UAE	NA
NATIONALITY	string	The nationality of an individual	50	Must be Drop Down List	UAE National	Available

INTERNATIONAL ADDRESS INFORMATION

ATTRIBUTE	DATA TYPE	DESCRIPTION	NUMBER OF CHARACTERS	VALIDATION RULE	EXAMPLE	AVAILABILITY
DETAILED ADDRESS	Alphanumeric	The mailing address of an individual	1-500	Must contain only letters (A-Z, a-z), digits (0-9), and/or spaces	123 Main St.	NA
CITY	string	The name of a city	50	Must be Drop Down List	Manchester City	Not Available
COUNTRY	string	The name of a country	100	Must be Drop Down List	United Kingdom	Available
COUNTRY CODE	string	The country code of a phone number	4	Must be Drop Down List	+44	Not Available

LOCAL ADDRESS INFORMATION

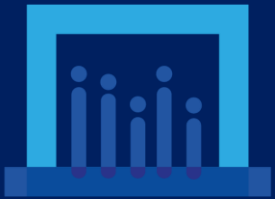
ATTRIBUTE	DATA TYPE	DESCRIPTION	NUMBER OF CHARACTERS	VALIDATION RULE	EXAMPLE	AVAILABILITY
DETAILED ADDRESS	Alphanumeric	The mailing address of an individual, Street number and Villa or Apartment number	1-500	Must contain only letters (A-Z, a-z), digits (0-9), and/or spaces	Street 12 Villa 25	Not Available
AREA	varchar	The name of a area	50	Must be Drop Down List	Khawaneej 1	Not Available
MAIN AREA	string	The name of a area	50	Must be Drop Down List	Deira/ Bur Dubai	Not Available
CITY	string	The name of a city	100	Must be Drop Down List	Dubai / Hatta	Not Available
EMIRATE	string	The name of a city	100	Must be Drop Down List	Dubai / Sharjah	Available
LONGITUDE	Numeric	Longitude	50	Map pin	53.2734	NA
LATITUDE	Numeric	Latitude	50	Map pin	-7.77832031	NA

CONTACT INFORMATION

ATTRIBUTE	DATA TYPE	DESCRIPTION	NUMBER OF CHARACTERS	VALIDATION RULE	EXAMPLE	AVAILABILITY
EMAIL	Alphanumeric	The email address of an individual	1-255	Must be a valid email address format	john.smith@example.com	NA
MOBILE NUMBER	Varchar	Phone number in the format of (country phone code) area code-local number	20	Must be a valid phone number format, should start with a '+' followed by the country code, area code and local number	+971 (555) 555-5555	NA
WHATSAPP NUMBER	Varchar	Phone number in the format of (country phone code) area code-local number	20	Must be a valid phone number format, should start with a '+' followed by the country code, area code and local number	+971 (555) 555-5555	NA

DEMOGRAPHIC INFORMATION

ATTRIBUTE	DATA TYPE	DESCRIPTION	NUMBER OF CHARACTERS	VALIDATION RULE	EXAMPLE	AVAILABILITY
DATE OF BIRTH	Date	The birthdate of an individual	10	DD/MM/YYYY Read from Calendar plugin	Must be a valid date format (DD/MM/YYYY)	NA
GENDER	Alphabetic	The gender of an individual (e.g. Male, Female)	1-255	Male/Female Drop down	Must be "Male" or "Female".	Available
MARTIAL STATUS	Alphabetic	The marital status of an individual	8	Marital status register from Dubai Court	Single/Married/divorced	Available
PLACE OF BIRTH	string	The city where an individual was born	50	Drop down	Dubai	Not Available
COUNTRY OF BIRTH	string	The country where an individual was born	50	Drop down	UAE	Available
RELIGION	string	The religion of an individual	50	Drop down add option others	Islam	Not Available



Statistical Classification

ECONOMIC CLASSIFICATIONS

ATTRIBUTE	DATA TYPE	DESCRIPTION	NUMBER OF CHARACTERS	VALIDATION RULE	EXAMPLE	AVAILABILITY
LEGAL ENTITY OF THE ESTABLISHMENT	Alphabetic	Legal entity or form of the establishment	Maximum 100 characters	Required	Government Sector, Sole proprietorship	Available
ECONOMIC SECTOR OF THE ESTABLISHMENT	Alphabetic	Economic sector or industry to which the establishment belongs	Maximum 100 characters	Required	Public Sector, Private Sector	Available
CLASSIFICATION OF SMALL, MEDIUM AND LARGE-SIZED ESTABLISHMENTS	Alphabetic	Classification of establishments based on size	Maximum 100 characters	Required	Small Establishments, Medium-sized Establishments	Available
ESTABLISHMENT'S CHARACTERISTIC	Alphabetic	Specific characteristic or feature of the establishment in terms of being an individual, head office, branch or foreign company.	Maximum 100 characters	Optional	Headquarter with Branches, Foreign Establishment branch	Available
ESTABLISHMENT'S STATUS	Alphabetic	Current status or condition of the establishment	Maximum 100 characters	Required	Working Continuously, Under Equipping	Available
STATUS OF COLLECTING THE ESTABLISHMENT DATA (VISIT RESULT)	Alphabetic	Status or outcome of the visit or data collection for the establishment	Maximum 100 characters	Optional	Completed, Temporary closed	Available

POPULATION, DEMOGRAPHIC, AND SOCIAL CLASSIFICATIONS

ATTRIBUTE	DATA TYPE	DESCRIPTION	NUMBER OF CHARACTERS	VALIDATION RULE	EXAMPLE	AVAILABILITY
HOUSEHOLD TYPES	Available	Type or category of household	Maximum 100 characters	Required	Emirati Households, Non-Emirati Households	Available
POPULATION CLUSTERS	Available	Specific location to live in temporarily, such as labor accommodations and public residence	Maximum 100 characters	Optional	Labor Accommodation, Public Residence	Available
GENDER	Alphabetic	The gender of an individual (e.g. Male, Female)	1-255	Male/Female Drop down	Must be "Male" or "Female".	Available
RELATIONSHIP WITH THE HEAD OF THE HOUSEHOLD	Available	Relationship of individuals to the head of the household	Maximum 100 characters	Required	Son / Daughter, Brother / Sister	Available
MARITAL STATUS	Available	Current marital status	Maximum 50 characters	Required	Married, Single	Available
REASONS FOR DIVORCE	Available	The main reasons for divorce or separation	Maximum 100 characters	Optional	Not spending, Family reasons	Available
STATUS OF PRESENCE	Available	Current presence status (e.g., present, absent)	Maximum 50 characters	Required	Present, Absent	Available
RESIDENCE STATUS	Available	Current residence status of the individual within the household	Maximum 50 characters	Required	Usual resident of the household, Usual Resident of other Emirates	Available
SOURCE OF HEALTH INSURANCE	Available	Source of health insurance coverage	Maximum 100 characters	Optional	Personal, Government Sector	Available
NATIONALITY	Available	Nationality or citizenship	Maximum 100 characters	Required	American, British	Available

CLASSIFICATION OF BUILDINGS

ATTRIBUTE	DATA TYPE	DESCRIPTION	NUMBER OF CHARACTERS	VALIDATION RULE	EXAMPLE	AVAILABILITY
CLASSIFICATION OF TYPE (FORM) OF BUILDING	Alphabetic	Type or form of the building	Maximum 50 characters	Drop down	Villa, Mosque	Available
BUILDING STATUS	Alphabetic	Current status of the building ,whether they have never Existing / Completed, under construction, or under demolition.	Maximum 50 characters	Drop down	Under Construction	Available
BUILDING CHARACTERISTICS	Alphabetic	The status of freestanding constructed buildings, whether established permanently or temporarily	Maximum 100 characters	Drop down	Main Building	Available
TYPE OF PREDOMINANT SUBSTANCE OF THE BUILDING	Alphabetic	Type of substance predominant in the building (e.g., concrete, steel)	Maximum 50 characters	Drop down	Concrete	Available

CLASSIFICATIONS OF RESIDENTIAL UNITS AND REAL ESTATE

ATTRIBUTE	DATA TYPE	DESCRIPTION	NUMBER OF CHARACTERS	VALIDATION RULE	EXAMPLE	AVAILABILITY
RESIDENTIAL UNIT TYPE	Alphabetic	Type of residential unit (e.g., house, apartment)	50	Drop down	Apartment	Available
UNIT STATUS	Alphabetic	Current status of the unit (e.g., occupied, vacant)	20	Drop down	Existing / Complete	Available
UNIT USAGE	Alphabetic	Intended usage of the unit (e.g., residential, commercial)	50	Drop down	Residence	Available
TENURE TYPE	Alphabetic	Type of ownership or tenancy (e.g., owned, rented)	20	Drop down	Owned	Available
SOURCES OF WATER FOR ESTABLISHMENT	Alphabetic	Water sources used for the establishment	100	Drop down	Sea water, Groundwater	Available
SOURCES OF WATER FOR DOMESTIC USE	Alphabetic	Water sources used for domestic purposes	100	Drop down	Public Network, Private Well	Available
SOURCE OF DRINKING WATER	Alphabetic	Primary source of drinking water for the unit	50	Drop down	Mineral water, Distributed water	Available
PRAYER PLACES & MOSQUES	Alphabetic	Presence of prayer places and mosques nearby	100	Drop down	Public Mosque, Eid Mosque	Available
TYPE OF AIR CONDITIONING	Alphabetic	Type of air conditioning system used in the unit	50	Drop down	Split units	Available

EDUCATIONAL CLASSIFICATIONS

ATTRIBUTE	DATA TYPE	DESCRIPTION	NUMBER OF CHARACTERS	VALIDATION RULE	EXAMPLE	AVAILABILITY
EDUCATION STAGE	Alphabetic	Stage or level of education (e.g., primary, secondary)	Maximum 50 characters	Required	Secondary level	Available
EDUCATIONAL LEVEL	Alphabetic	The educational level that an individual reached an age of 10 years or older	Maximum 50 characters	Optional	Master	Available
EDUCATIONAL STATUS	Alphabetic	Current educational status (e.g., illiterate, Educated)	Maximum 50 characters	Optional	Educated	Available
SCHOOL ENROLLMENT	Alphabetic	Enrollment status in a school or educational institution	Maximum 50 characters	Required	Currently attending	Available

LABOR CLASSIFICATIONS

ATTRIBUTE	DATA TYPE	DESCRIPTION	NUMBER OF CHARACTERS	VALIDATION RULE	EXAMPLE	AVAILABILITY
EMPLOYMENT STATUS	Alphabetic	The classification aims to determine the individual status in terms of the workforce by identifying the individual's relationship with the establishment in which he works	Maximum 50 characters	Required	Paid Employee	Available
RELATIONSHIP TO WORKFORCE	Alphabetic	Relationship of the individual whose age is 15 and above to work, and identify its types	Maximum 50 characters	Optional	Employed, Unemployed	Available

ENVIRONMENTAL CLASSIFICATIONS

ATTRIBUTE	DATA TYPE	DESCRIPTION	NUMBER OF CHARACTERS	VALIDATION RULE	EXAMPLE	AVAILABILITY
HEALTH AND SAFETY/POLLUTANTS	Alphabetic	Health and safety concerns related to pollutants	Maximum 100 characters	Optional	Smoke – car exhaust fumes, Dust	Available
HEALTH AND SAFETY AT THE HOUSE/DWELLING	Alphabetic	Health and safety considerations specific to the house or dwelling	Maximum 100 characters	Optional	Moisture and mold, High temperature	Available
OCCUPATIONAL SAFETY	Alphabetic	Safety measures and concerns related to occupational activities	Maximum 100 characters	Optional	Number of first aid injuries	Available
COST INCURRED BY ENVIRONMENT PROTECTION ACTIVITIES IN THE ESTABLISHMENT	Alphabetic	The practices carried out by the establishment for the purposes of protecting the environment from pollution emanating from production processes	Maximum 100 characters	Optional	Wastewater management, Environmental charges	Available
REMOVAL (DISPOSAL) OF WASTEWATER/SEWAGE	Alphabetic	Methods or systems used for wastewater or sewage removal and disposal	Maximum 100 characters	Optional	Contract with a company, Septic tank	Available

ENVIRONMENTAL CLASSIFICATIONS

ATTRIBUTE	DATA TYPE	DESCRIPTION	NUMBER OF CHARACTERS	VALIDATION RULE	EXAMPLE	AVAILABILITY
SOURCE OF ENERGY AND FUEL	Alphabetic	Sources of energy and fuel used for various activities	Maximum 100 characters	Optional	Electricity, Central gas	Available
SOURCE OF ELECTRICITY	Alphabetic	Sources of electricity for the establishment	Maximum 100 characters	Optional	Public power network, Private generator	Available
SOLID HAZARDOUS WASTE	Alphabetic	Management or disposal of solid hazardous waste	Maximum 100 characters	Optional	Chemical waste, Medical waste	Available
SOLID NON-HAZARDOUS WASTE	Alphabetic	Management or disposal of solid non-hazardous waste	Maximum 100 characters	Optional	Paper and cardboard, Plastic	Available
SOLID WASTE DISPOSAL METHODS	Alphabetic	Methods or systems used for solid waste disposal	Maximum 100 characters	Optional	Waste container, Recycling	Available
LIQUID WASTE DISPOSAL METHODS	Alphabetic	Methods or systems used for liquid waste disposal	Maximum 100 characters	Optional	Onsite treatment, Discharge into the sea	Available
HOUSEHOLD WASTE DISPOSAL METHODS	Alphabetic	Methods or systems used for household waste disposal	Maximum 100 characters	Optional	Burned, Collected by collectors	Available

GUIDEBOOK GLOSSARY

TERM	DEFINITION
DDE (Dubai Data Establishment)	The Dubai Data Establishment was established in 2016 is the governing body that is charged with driving forward the strategy and implementation roadmap for Dubai Data.
DDE CEO (Dubai Data Establishment Chief Executive Officer)	The Dubai Data Establishment is led by its Chief Executive Officer (CEO) who is in effect the Data Leader of Dubai.
Dubai Data	The Dubai Data Law defines Dubai Data as “The Data which is available to Data Providers and is related to the Emirate”. This includes any data relating to any aspect of the Government, economy, culture and life within the Emirate of Dubai.
Dubai Data Law	A law issued in 2015 by the Dubai Government that governs the Classification and use of Dubai Data within the Emirate of Dubai.
Government Entities	Ministries, bodies and entities of the Federal Government as well as directorates, bodies and institutions of the Local Government.
Dubai Paperless Government	An initiative launched by Smart Dubai which aims in fully transforming and digitizing internal processes and customer facing services from 2021. That means government will no longer issue or ask for paper documents across all its operations.
Data Collection	The process of gathering and storing data. May include sorting, analysis, transformation, summarisation, etc.
Data Controller	‘Data Controller’ refers to the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data. (GDPR Article 4)

GUIDEBOOK GLOSSARY

TERM	DEFINITION
Data	An organized or non-organized set of inputs, facts, concepts, instructions, observations or measurements, in the form of numbers, letters, symbols, images, or any other form, allowing its interpretation, exchange or processing, by individuals or computers.
Dataset	A collection of data that it makes sense to group together, along with the metadata 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 collection of paper records, etc.
Data Base	A collection of information that is organized so that it can be easily accessed, managed and updated.
Data Management	Refers to the disciplines and techniques to manage data as an asset.
Data Classification	The process of classifying data as either Open or Shared using the Dubai Data Classification Framework and enabling review and approval of those classification decisions by DDE.
Data Register	A digital list which serves as the single source of truth in Dubai for its subject and has been officially recognized by the Dubai data establishment as a reference register under the Dubai Data Law.
Data Sharing Agreement	A data-sharing agreement is a formal contract that clearly documents what data are being shared, how the data will be shared and how the data can be used.
Data Governance	Data governance encompasses the people, processes, and information technology required to create a consistent and proper handling of an organization's data across the business enterprise.
Data Modelling	Data modelling is the process of creating a data model for the data to be stored in a database. This data model is a conceptual representation of data objects, the associations between different data objects, and the rules.

GUIDEBOOK GLOSSARY

TERM	DEFINITION
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 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.
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 Maintenance	Data maintenance is the process of organizing and curating data
Data Analytics	Data analytics is the science of analyzing raw data in order to make conclusions about that information.
Data Architecture	Data architecture is composed of models, policies, rules or standards that govern which data is collected, and how it is stored, arranged, integrated, and use in data systems and in organizations.
Data Infrastructure	A data infrastructure is a collection of data assets, the bodies that maintain them and guides that explain how to use the collected data.
Data Quality	Data quality indicates how reliable a given dataset is.
Data Ingestion	Data ingestion is the transportation of data from assorted sources to a storage medium (Dubai Pulse) where it can be accessed and used.

