
Data



Data

City Planning Labs

The last decade has witnessed an urban data revolution, as cities internationally have started mobilizing geospatial data to harness the potential of urbanization and address its challenges.

City Planning Labs (CPL) is a Technical Assistance program of the World Bank, which aims to enhance the technical and institutional capacity of municipal governments to produce, share, and utilize geospatial data for evidence-led urban planning. The foundation of CPL includes the establishment of a robust and innovative **Municipal Spatial Data Infrastructure (MSDI)** platform. An MSDI ensures the sustainability of geospatial innovations being introduced, and is being pioneered by partner cities in Indonesia. In addition, agile and adaptable **Urban Planning Tools** are aimed at empowering cities to make informed decisions to improve the quality of life of their residents.

CPL's approach is aligned with the **World Bank's Build, Boost and Broker framework** as it Builds critical municipal spatial data foundations and institutionalizes them; Boosts their capacity to utilize information for evidence drive planning, and; Brokers the relationships between city governments and the private sector to leverage innovation in technology.

Municipal Spatial Data Infrastructure

Municipal Spatial Data Infrastructure (MSDI) is the platform that facilitates the organization, sharing and utilization of geospatial information to tackle the challenges to achieve sustainable urban development. It is the cornerstone of any strategy for cities aspiring to embrace digital transformation and for the long-term success of smart city initiatives.

Clear action plans for MSDI operationalization i.e., **MSDI Roadmaps**, have been developed by CPL partner cities using a **four-pillar IPDS framework**. IPDS pillars correspond to **Institutions** (local regulations and data governance protocols), **People** (competency frameworks and skill development), **Data** (collection, processing, management, utilization, and data driven tools), and **Systems** (Integrated Data Platform i.e., Geoportal).

This comprehensive IPDS framework enables **governance and regulatory innovations** to go hand in hand with technological solutions, and also takes into account the importance of human resources to strengthen data foundations. The IPDS framework is **agile, flexible, and scalable to cities with varying levels of capacity within Indonesia and globally**.



WORLD BANK GROUP



Swiss Confederation



Korea
Green Growth
Trust Fund

This work received financial support from the Swiss State Secretariat for Economic Affairs (SECO) through the Indonesia Sustainable Urbanization Multi-donor Trust Fund (IDSUN MDTF).

Gayatri Singh, Task Team Leader – gsingh9@worldbank.org **Seetha Raghupathy** – sraghupathy@worldbank.org
Champaka Rajagopal – crajagopal@worldbank.org **Aurora Dias Lokita** – alokita@worldbank.org **Tony Hartanto Widjarnarso** – twidjarnarso@worldbank.org **Irene Arifin** – ikarifin@gmail.com **Rebekka Hutabarat** – rhutabarat@worldbank.org;
Editors: **Gauri Uday Gadgil**, **Christopher Paul Crow**



“Data is to this century what oil was to the last one: a driver of growth and change. Flows of data have created new infrastructure, new businesses, new monopolies, new politics and — crucially — new economics.”

-The Economist (2017)

CPL- IPDS Framework:

The Data Pillar

A data revolution is underway with geospatial information being used ubiquitously by institutions and communities for day-to-day decision-making. Cities across the globe are at the forefront of this movement, increasingly leveraging the use of geospatial data and technology to improve urban planning and service delivery.

A data-driven planning approach not only helps city governments address existing gaps in service delivery more effectively, but also offers powerful inputs to determine appropriate policy interventions and prioritize infrastructure investments for sustainable and inclusive growth-- thus harnessing the full potential of urbanization.

Spatial information, such as maps, satellite imagery, and aerial photography describe the connection between a place, its people and their activities, providing a holistic view of the cityscape that can be used to identify and analyze relationships and trends over time. It can also be used to identify additional value of proposed infrastructure. For example, data on travel time and accessibility to employment nodes from public transit stations can translate into information on fulfilled or

unfulfilled economic potential of a city.

While the benefits of a data-driven approach to urban planning and management are clear and widely accepted, significant challenges hamper its full adoption, especially in low and middle-income countries. Data (especially geospatial data) available at the municipal level can be incomplete, out-of-date, missing authoritative sources or lack accuracy in representing features on the ground.

In addition, the absence of clear data governance frameworks and centralized data platforms prevent effective and sustained data sharing across sectors. Most importantly, inadequate investment in human resource capacity to convert data into information is a significant roadblock.

To address these barriers, CPL has undertaken an ecosystem approach to operationalize a robust Municipal Spatial Data Infrastructure (MSDI) in partner cities through its holistic IPDS framework ►

This document details the 'Data' component, which refers to the current state of affairs on data availability,

quality, and related policies regarding formats, analysis, and sharing of geospatial information. The data diagnostic also considers the business case for investment in data and the extent to which the use of geospatial information can add value to existing line department Key Performance Indicators.

I
Institutional Arrangements

P
People

D
Data

S
Systems

Understanding Data Challenges and their Implications

Challenges in using geospatial data stem primarily from the absence of standardized core datasets in cities, despite common data needs and utilization across municipal agencies. This fuels data duplication, inaccuracies, and ambiguous data custodianship, consequently impeding evidence-led planning for inclusive service delivery. Such issues have been identified in city governments across the globe. CPL's in-depth examination of Indonesian cities of varying capacities forms the basis of the diagnostics and toolkits developed under this pillar.

Some of the data challenges and their implications for urban planning and management are as follows:



Discussion on Data Challenges



CAPACITY GAPS PROMOTE RELIANCE ON EXTERNAL CONSULTANTS to produce analytical reports, including data and maps as deliverables. The outputs submitted to the government often do not include the raw datasets or base information used in their production. As a result, city staff are unable to verify the analysis while reviewing the reports. In some cases, this results in a repeated creation of data, which is financially inefficient. It also hinders capacity development as city officials are not required to develop skills for data management, and serves to further a cycle of dependency on external consultants.

Contracting analytical work to external consultants has become an accepted practice globally and is followed even by high-capacity cities. For this process to be effective, however, city staff must have the capacity to define the parameters of the analytical inputs and track the quality of outputs delivered.



AGENCIES OFTEN OPERATE IN SILOS since mechanisms and protocols for inter-departmental coordination are absent or weak. The result is data incompatibility since the same data is produced by different agencies, without reconciliation of results or alignment between sources.

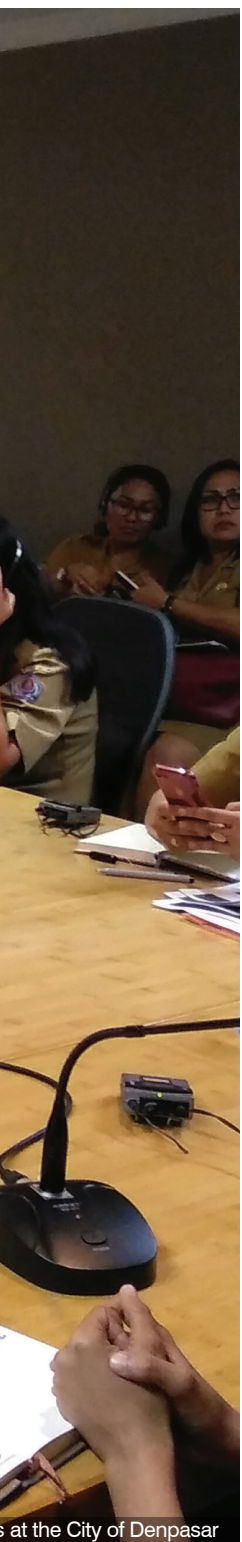


DATA INTEGRATION IS CHALLENGING WHEN GEOSPATIAL DATASETS EXIST IN DIFFERENT FORMATS. This issue is exacerbated by both the inadequate streamlining of processes for data updating and maintenance by line departments and absence of protocols and processes for data sharing. Therefore, the same datasets may be created multiple times, leading to inefficient use of constrained budgets.

The CPL Approach



Discussion on Fundamental Data Sets



Drawing from various international examples, and building on existing national and local mandates, CPL has adopted a multi-pronged approach to address the data-related challenges of partner cities as illustrated below:

1

Data Inventory (DI)

The move toward evidence-based decision making requires cities to possess core datasets with clear documentation of how the data was collected, geocoded, and processed. Data inventories are useful to understand the prevailing data management practices of different agencies. Data inventories also allow governments to more easily review existing datasets and assess needs for developing new datasets.

Recognizing this, CPL has developed a DI tool to gather standard metadata¹ on spatial datasets. Metadata standards are a component of data standards that specify how datasets should be cataloged and what information about the datasets should be included. The use of standardized descriptions eases access to data across line departments as well as public and private stakeholders and helps determine the accuracy of data produced. Standardized descriptions also enable search engines to process queries more efficiently, ensuring that users and applications conducting searches are presented with relevant results. Custodians of geospatial data assets will benefit as their information resources become discoverable by a wider range of potential users at negligible cost. Overall, benefits will accrue not only to the organizations but to the wider population.

As part of the DI process, line departments must complete and return the questionnaire to the MSDI Diagnostic consultant who will use the information to identify any irregularities and address them with the agencies. The DI questionnaire includes questions such as:

- The general name of the dataset in the local language (in the case of Indonesia, Bahasa Indonesia)
- Data title in English
- Name of the file, when the document was created
- Last update and update frequency
- Data format
- Data access category etc.

This exercise also helps inform line departments as to which fields of information are critical to fill in when assembling datasets to upload onto the Geoportal.

2

Support to Optimize the Utility of Existing Datasets

A review of the DI provides a clear mapping of the type and number of datasets cities possess. However, while cities may already possess spatial datasets, these datasets are often underutilized due to constraints posed by institutional arrangements, people, and systems. While addressing all of these challenges will take a multi-faceted effort, existing data can be cleaned and standardized to make them analysis-ready in the short term.

CPL enables this data analysis by providing technical assistance to cities, allowing them to undertake analytical studies e.g., perform infrastructure accessibility analysis based on location of schools, hospitals; identify high-poverty areas with inadequate access to urban services etc.

3

Introduce and Enable the Adoption of Data Standards

Data standards are an important component to enable MSDI, since their adoption can ensure that datasets can be integrated with one another. This can mean something as simple as having datasets with common attribute information, allowing them to be combined and subsequently used for analysis together. In the context of data sharing, having a common standard can provide users of a Geoportal with a consistent and structured method to search for datasets, based on their attributes. CPL compiles relevant data standards to introduce them to data users within city line departments, enhancing the quality of analysis, while also improving their spatial data production chain and interoperability.

FUNDAMENTAL DATASETS (FDS)

Fundamental Datasets are defined as the authoritative core datasets that provide spatial reference, enable inter-operability of geospatial data and applications, and form the common base map that is essential to a city's economic, social, and environmental management and development. (GSDI, 2012) (UNECA, 2007) (UN-GGIM, 2016) (ANZLIC, 2018) (PAIGH, 2017). They represent the lowest common denominator of data requirements for all sectors and are important for the realization of all planning goals. FDS should be generic, and provide basic reference for a wide range of stakeholders to achieve critical planning objectives.



Why are FDS Important

There are three key reasons that FDS are critical to MSDI:

- FDS offer a framework for prioritizing the time, budget, and manpower needed to collect and standardize geospatial data for planning and decision-making across sectors.
- Standardizing FDS enables widespread inter-operability and helps break down information silos, as it ensures that key datasets representing each sector are integrated.
- The FDS framework also helps build accountability for data management throughout the lifecycle of datasets by assigning responsibility to specific line departments for aspects such as data collection, standards, quality, etc.

DETERMINING A FUNDAMENTAL DATASET

Based on review of international literature, the following points can be considered for the definition of FDS:

- The data has a spatial component (with x, y coordinate, latitude/longitude or addresses)
- The data is non-derivative, but forms a base layer based on primary collection
- The data relates to factors that are essential for city-wide development, its economy or public infrastructure
- The data is critical for multiple stakeholders within the municipal government

Furthermore, FDS should be Authoritative, Accurate and Accessible, consistent with general information management principles.

Recognizing the importance of having complete FDS, Indonesia's Badan Informasi Geospasial (BIG) has determined a list of FDS as the national standard

- BIG's FDS standards consists of 59 data sets
- These are further grouped into 14 data themes

In the context of Indonesia, the list of FDS thematic categories, as outlined in Book 1, Principles of Katalog Unsur Geografi Indonesia (KUGI) (BIG, 2018) are ►

The acquisition of these datasets are expected to ease the process of spatial data production, and contribute to resolving various data-related issues, such as availability, referencing, and inter-operability.

No.	Thematic Group Name
	Definition
1	Spatial References
	Datasets on the coordinate reference system used in defining and locating the position of a geospatial entity, including horizontal and vertical positions, its weight and changes over time
2	Administrative Boundaries
	Datasets on objects that reflect the management of regional/ administrative boundaries
3	Transportation
	Datasets that describe transportation facilities and infrastructure
4	Hydrography
	Datasets that describe physical representations of seas, lakes, rivers, and their relation with coastal regions, as well as utility for navigational needs
5	Hypsography
	Datasets that describe the relief of the earth's surface, following a certain height reference
6	Vegetation
	Datasets on vegetation objects that cover the surface of the earth
7	Built Environment
	Datasets that describe the physical representation of built environments on the surface of the earth
8	Utilities
	Datasets that describe particular-use facilities and infrastructures
9	Geology
	Datasets that describe the composition, structure, and physical nature of the earth's surface
10	Land/ Soil
	Datasets on objects related to soil (characteristics)
11	Toponymy
	Datasets on objects that are used to describe the name of places
12	Cadaster
	Datasets on objects that are used to describe land parcels
13	Disaster
	Datasets on events or chain of events that threatens or disrupts lives and livelihoods that are caused by either natural, or man-made factors
14	Special Dataset
	Datasets on objects with special utilities

CROSS-REFERENCE THE DI TOOL FINDINGS WITH FDS

Developing FDS is a long and complex process that will entail several years of investment in data by municipal governments. Recognizing this, CPL works with partner cities to commence the development of FDS and guide them through the process. As a first step, the DI questionnaire responses are cross-referenced with BIG's official list of Indonesian FDS. City officials can use the results of this analysis to check their current level of spatial data availability against the pre-determined list of FDS and strategize accordingly to develop the missing datasets.

Given that FDS are data that can be used as a reference or base to produce other spatial datasets, the identification, prioritization, and production of FDS will help accelerate the production of other spatial datasets as well.

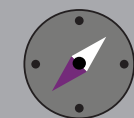
FDS ACQUISITION PRIORITIZATION:

CPL used the following steps in partner cities to determine the priority FDS to be acquired:

- Consider international examples and, the national requirements (e.g. the BIG KUGI list in the Indonesian case) of FDS to help determine the city's FDS needs. **Spotlight 2** illustrates the FDS adopted by BIG in comparison to UN-GGIM (United Nations Committee of Experts on Global Geospatial Information Management).
- Cross-reference the DI with the FDS list to determine the FDS that cities already possess and those that cities are yet to possess.
- Assess the priority challenges identified by the cities through a consultative approach (e.g., surveys/ inter-agency MSDI working group discussions etc.) to derive the list of FDS that cities most urgently need to address these challenges.



UN-GGIM's Minimum List of Global Fundamental Geospatial Data Themes



Global Geodetic Reference Frame



Addresses



Buildings and Settlements



Elevation and Depth



Land Parcels



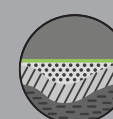
Population Distribution



Functional Areas



Geographical Names



Geology and Soils



Physical Infrastructure



Land Cover and Land Use



Transport Networks



Water



Orthoimagery

Source:
UN-GGIM 8th
session report on
the determination
of global
fundamental
geospatial data
themes

*Adopted by BIG

Leverage Innovative Approaches to Address Unavailability of Other Critical Datasets

While FDS takes time to develop, some types of data are crucial to help cities achieve short and medium-term goals. To acquire such datasets CPL has collaborated with the European Space Agency (ESA). Expert consultants, in partnership with ESA, used high resolution and very high-resolution satellite imagery to map urban growth patterns and trace urban expansion across four points in time (up to 2015). This effort provided essential processed data to understand forces driving spatial growth in the city.

Specific analytical products included:

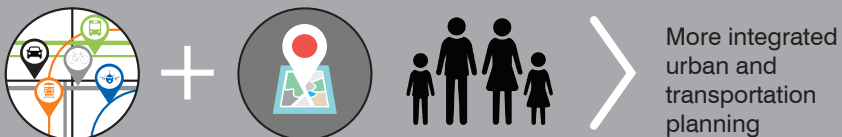
- Urban Land Use/Land Cover (current and historic)
- Extent of urbanized area (for the four points in time)
- Transport infrastructure – road network
- Terrain motion analysis using radar data
- Assessment of change of green areas
- Flood history analysis

CPL is also exploring the possibility of expanding the list of data types that it can provide. A few ongoing initiatives are described in the Coming Up section.

FDS General Use Case Sample

Road Network

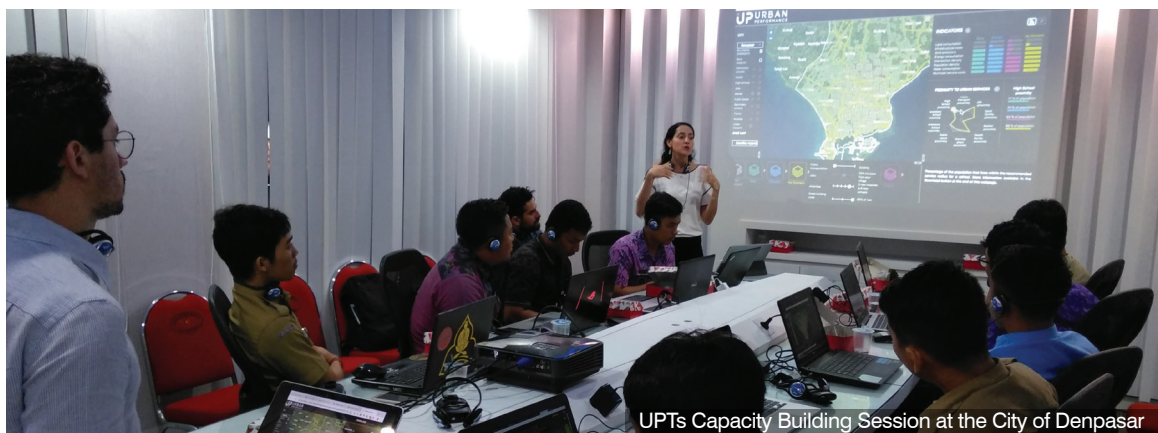
The road network dataset can be overlaid with land-use plans and demographic data to analyze possible optimal locations for road infrastructure



Cadastral Land Parcel

The cadastral land parcel dataset can be overlaid with groundwater movement and rental price data to assess the impact of land subsidence on property value





UPTs Capacity Building Session at the City of Denpasar

No.4

Using FDS in the CPL Partner City of Balikpapan

ROAD NETWORK The road network dataset can be overlaid with land use plans and demographic data to analyze possible locations where roads can be built to serve the population optimally

The City of Balikpapan identified urban planning and disaster management (flooding) as two priority concerns during their workshops with CPL. Subsequently, CPL helped the city to articulate the spatial data they would need to address these challenges, including data on drainage networks, housing and settlement plans, and building outlines. Superimposing these layers on top of each other, the city was able to identify settlement areas that are affected by flooding, and begin to address these challenges through their drainage and land-use plans. This approach enabled the city to leverage datasets that they already possessed (drainage networks), while also improving their stock of FDS (building outlines), which can be used to advance their other spatial planning priorities in the future.

5

Use Spatial Data Analytics for Urban Planning

Two agile and user-friendly urban planning analytical tools, the Suitability Tool (ST) and the Urban Performance (UP) tool, have been developed as part of CPL to support evidence-driven urban planning and effectively lower the knowledge barriers that often prevent cities from reaping the full benefits of spatial data. Using GIS and multi-criteria analysis, the ST identifies optimal locations for a specific activity within a city. It does so by evaluating the current status of neighborhoods across the city, based on the availability of infrastructure and distance to services and amenities. This tool has multiple possible applications including:

- Mapping infrastructure hotspots
- More efficiently prioritizing investment
- Assessing planning decisions to optimize for the interests of multiple stakeholders
- Identifying optimal locations for social housing

The UP tool facilitates the assessment of a city's possible future conditions in relation to urban policies. Modeled scenarios provide a quantitative forecast of the benefits and drawbacks of policy decisions on the environmental, social, and economic dimensions of the city. By modeling and comparing the effect of policies on such factors as land consumption, infrastructure costs, walkability, and proximity to vital public facilities, cities are able to weigh the benefits of urban policies prior to enacting them.

Outcomes and Impact in CPL cities:

	Activity	Product/Outcome	Cities	Impact on Cities
1	Develop a Data Inventory (DI) Tool	Completed DI responses	(S) (D) (B)	Awareness across all line departments about existing and missing datasets
2	Optimize the Utility of Existing Datasets	<p>Analytical Reports on Urban Growth and Access to Public Infrastructure</p> <p>Urban Planning Tools (UPT): the ST and UP tools</p>	(S) (D)	<p>Knowledge and capacity development on using datasets to inform analytics and spatial planning; enhanced understanding of the use case of the IDP (please see the Systems component of the CPL toolkit for further details)</p> <p>Scaling up of UPT to about 150 cities by 2020</p>
3	Introduce and Enable the Adoption of Data Standards	Document on International and Indonesian Data Standards to Adopt	(S) (D) (B)	Adoption of standards to be followed and awareness of benefits
3a	Fundamental Data Sets (FDS)	Document on FDS- purpose, data types etc.	(S) (D) (B)	Increased knowledge on FDS
3b	Cross-reference the DI Tool with FDS	Report on FDS required vs. currently available	(S) (D) (B)	Acceleration of production of missing datasets
4	Address Unavailability of Other Critical Datasets	Analytical Products:		Informing the Spatial Planning processes; in particular, the Urban Terrain/ Infrastructure Motion product has been used by the City of Semarang to help conduct an optimal-location analysis for affordable housing development, among many other applications
		• Urban Land Use/Land Cover (current and historic)	(S) (D)	
		• Extent of urbanized area (for the four points in time)	(S) (D)	
		• Transport infrastructure – road network	(D)	
		• Terrain Motion Analysis using radar data	(S)	
		• Assessment of change of green areas	(S) (D)	
		• Flood history analysis	(D)	
5	Leverage Spatial Data Analytics for Urban Planning: UPT	Agile and User-Friendly UPT	(S) (D)	Adoption of Urban Data Diagnostics to inform Urban Planning: Data-driven approach to infrastructure provision, optimization of investment and adoption of policies with environmental, social and economic impacts on the cities



Semarang



Denpasar



Balikpapan

Coming Up



Discussion between CPL Team and Representatives from the CPL Partner Cities

As highlighted in the table, CPL has initiated several geospatial innovations which have begun to catalyze a shift to a strategic and evidence-led spatial planning approach.

The CPL collaboration with the ESA described in the previous section will enable the cities to use satellite image interpretation to produce building footprints and built-area datasets such as roads, pavement, and sidewalks at the optimal level of detail. The use of this technique represents a significant shift from the expensive and time-consuming process of conducting land surveys. This dataset could then be used as a reference for the production of a fit-for-purpose cadastral dataset, enabling cities to conduct just-in-time planning, as opposed to having to wait for land survey data to be collected and compiled. The use of high-resolution satellite imagery and image interpretation is also being explored to map the built environment and its relationship with ecologically sensitive areas in Balikpapan. For a city faced with ecological hazards on an annual basis, the creation of this data will help the city target interventions where most essential.

Current activities in CPL's partner cities will be accelerated alongside new activities. In addition, the initiative will be scaled-up to other cities in Indonesia. CPL is working on collaborating with

several other partners, such as Digital Globe and the World Bank's Geospatial Operations Support Team (GOST), to help Indonesian cities procure and develop geospatial datasets for the purposes of data-driven planning. The collaboration with GOST, in particular, will help the CPL team procure high-resolution satellite imagery, as well as Green Cover geospatial datasets for Indonesia's top 20 priority cities. This initiative is further strengthened by collaboration with volunteers from organizations and communities such as Open Street Map (OSM), for purposes of data collection and validation. Options are also currently being explored to expand the reach of this work to more cities.

Further, the Ministry of Spatial Planning has expressed an interest in utilizing the Suitability Tool for expediting the approval process of multiple detailed spatial plans (RDTR) and the Urban Performance Tool for evaluating the city level spatial plans (e.g. RTRW).

While CPL's products have been developed in Indonesia, the program's toolkit which guides development of MSDI's 'Data' dimension is scalable and replicable given its agile and modular approach. Efforts are currently underway to explore the adoption of these toolkits globally.

Data Inventory (DI) Survey

Survey instrument

Allows quick assessment of the overall data availability, quality, and custodianship in cities. Completing this survey helps raise awareness on what data is available, enabling organizations to identify and strategize the collection of any unavailable datasets

Baseline Fundamental Data Set (FDS) List

Baseline FDS reference list

Compiled from several international standards, a baseline FDS list could be used as reference for organizations or municipalities seeking to develop their own FDS list in the absence of one, or benchmark their existing FDS.

FDS Assessment Toolkit

An algorithm, accompanied by a markdown document/ codebook that explains how to use the tool, and the methodology behind the algorithm

The FDS assessment toolkit is a simple string-matching algorithm that enables a rapid assessment on the availability of FDS, based on the list of FDS and the result of the DI survey. Using this tool, users can consistently replicate an FDS assessment, and get a quick summary of availability of FDS in the target organization.

*These products have
been developed
by CPL and are
available for use
by any interested
stakeholders.*

