



DFA Implementation Strategy

18 April 2018
r1

COMMUNITY SYSTEMS FOUNDATION

205 East 42nd Street, Suite 15050
New York, NY 10017 USA

+1 212 500 1335
www.communitysystemsfoundation.org

Section 1. **Background**

The global SDG indicator framework¹ is the basis for reviewing the implementation of the 2030 Agenda for Sustainable Development. Through this framework, the General Assembly urges the United Nations Statistical Commission to ensure the harmonization and consistency of data and statistics for the indicators used to follow up and review the Sustainable Development Goals and targets.

The High-Level Group for Partnership, Coordination and Capacity-building for Statistics (HLG-PCCB) has been tasked by the Statistical Commission to promote system-wide coherence and coordination for data and statistics in support of the 2030 Agenda for Sustainable Development, as the implementation of the SDG indicator framework requires national statistical systems to store, analyze, share and disseminate a large volume of complex SDG-related data and indicators, disaggregated by characteristics such as age, sex, residence, income, and disability at the national and sub-national levels.

The main framework for responding to the data needs of the 2030 Agenda is provided by the *Cape Town Global Action Plan for Sustainable Development Data* (2017), adopted by the UN Statistical Commission at its 48th session and endorsed by the UN General Assembly in its resolution A/RES/71/313. Strategic Area 4 of the Cape Town Global Action Plan on “Dissemination and use of sustainable development data” and its Objective 4.1 call to “develop and promote innovative strategies to ensure proper dissemination and use of data for sustainable development”.

Based on the achievements and lessons learned from monitoring the Millennium Development Goals (MDGs), national statistical offices have stressed the need for national SDG indicator reporting and dissemination platforms to improve access to data on key national development priorities and the transparency of official statistics, in compliance with the Fundamental Principles of Official Statistics.

¹ Developed by the Inter-agency and Expert Group on SDG indicators, and adopted by the General Assembly in 2017, in its resolution on Work of the Statistical Commission pertaining to the 2030 Agenda for Sustainable Development (A/RES/71/313)

Section 2. Approach

The DFA toolkit is a platform of open source software tools designed to produce reliable and timely data for programme planning, service delivery, transparency, and accountability. DFA is contributing to the realization of the SDGs and the 2030 Agenda by assisting Member States to achieve sustainable development through strengthening data and statistical systems for evidence-based policy formulation.



DFA is a platform of open source software tools for monitoring national development priorities including the Sustainable Development Goals.

DFA promotes interoperability within the ecosystem of socioeconomic statistics by complying with international standards. It is easy to exchange data between the loosely-coupled modules of the platform.



CAPTURE

Tools to capture data from surveys, administrative systems and real-time SMS polling.



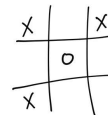
MANAGE

Tools to transform micro-data to indicators, standardise indicator metadata, and load data into a data warehouse.



ANALYZE

Tools to visualise data in graphs, tables and maps. Options to disseminate data in public dashboards and data profiles at the subnational level.



STRATEGIZE

Tools to plan and monitor development strategies to reach the most vulnerable with high quality, sustainable services.

www.dataforall.org


☰
DFA TOOLKIT
🔍

CAPTURE

- Admin Records
- Survey
- SMS
- Logistics
- Checklist
- WiFi Connect
- Smart Devices
- Big Data Capture
- Geo Data Capture

MANAGE

- Data Integrator
- Data Manager
- Indicator Registry
- Digital Map Library
- Data Entry
- Data Quality Checker
- Data Exchange
- My Data
- Data APIs
- R Studio Stats



ANALYZE

- Visualizer
- Report Builder
- Dashboard Builder
- Profile Builder
- Infographics Studio
- Data Story Builder
- Mapper
- Gallery Manager
- Doc Repository
- Portal Manager
- Mobile App
- Advocacy Designer
- Learning Studio

STRATEGIZE

- Monitoring
- Modelling
- Risk Mitigator
- Human Rights Monitor
- Data Games

www.dataforall.org



Section 3. **Architecture**

DFA promotes greater interoperability within the ecosystem of socioeconomic statistics. The DFA approach complies with the Common Statistical Production Architecture (CSPA)². CSPA is a set of common principles and standards designed to promote greater interoperability among the different stakeholders that make up the ecosystem of official statistics. The CSPA provides guidelines for defining, developing and deploying statistical services more efficiently.

Business Processes Layer

Generic Statistical Business Process Model. Business Architecture covers all the activities undertaken by a statistical organization, including those undertaken to conceptualize, design, build and maintain information and application assets used in the production of statistical outputs. Business Architecture drives the Information, Application and Technology architectures for a statistical organization.

Information Layer

Generic Statistical Information Model. Information Architecture classifies the information and knowledge assets gathered, produced and used within the Business Architecture. It also describes the information standards and frameworks that underpin the statistical information. This layer facilitates discoverability and accessibility, leading to greater reuse and sharing.

Software Applications Layer

DFA Artefacts Catalogue. This catalogue is a description of the major logical grouping of capabilities that manage the data objects necessary to process the data and support the business. This layer details the structure of components, their inter-relationships, and the principles and guidelines governing their design and evolution over time.

Technology Layer

DFA Technology Stack. This describes the IT infrastructure required to support the deployment of business services, data services and applications services, including hardware, middleware and networks.

² See <https://statswiki.unecce.org/display/CSPA/Common+Statistical+Production+Architecture>



Section 4. Principles

The DFA approach complies with the following principles reviewed by the 49th Session of the United Nations Statistical Commission (New York, 2018):

- 1. Coordination.* The NSO will be responsible for the coordination of the SDG reporting platform within the national statistical system. The implementation of the system within national institutions will be mandated by law and guided by standard operating procedures. The NSO will be responsible for the coordination, testing, and validation of the platform.
- 2. Fundamental Principles of Official Statistics.* The SDG data warehouse and reporting platform will comply with these fundamental principles and address the priority requirements for collection and dissemination of socioeconomic statistics. The purpose, scope and features of the platform will be clearly established in cooperation and consultation with all stakeholders and users.
- 3. Sustainability.* The development and administration of the SDG data warehouse and reporting platform will be based on the resources and capacities of the NSO to collect reliable data on key development priorities, and then to ensure that the data are put to good use.
- 4. Interoperability and Statistical Standards.* The components of the SDG data warehouse and data dashboard reporting platform will follow international statistical standards and best practices. Standards will be supported that aim to facilitate data harmonization and exchange across the different stages of statistical production and across institutions, including common data structure definitions and code lists.

Section 5. Guidelines

These guidelines are recommended for the successful implementation of an SDG reporting platform built using the DFA toolkit:

- 1. Institutional Ownership.* The NSO will have the ability to maintain, adapt, transform and customize the implementation of the platform to address their own needs, such as the management of national and subnational administrative boundaries, ethnic and language groups, and additional indicator definitions related to national development priorities. Reference principles: 1, 2, 3.
- 2. Collaboration.* The platform will be designed, developed, improved and maintained on the basis of a collaborative approach that leverages learning between technology developers, donors, policy makers, subject-matter experts, business partners, advocacy groups and both institutional and civil society users. Reference principles: 1, 2, 3.
- 3. Indicator Definitions and Metadata.* The platform will support indicator definitions and metadata at the appropriate level of customization of international indicator frameworks for monitoring national development priorities. The design of the system will allow for indicators to be linked to multiple monitoring frameworks, such as, national development plans, the SDG indicator framework, and sectoral development plans. Reference principles: 2.
- 4. User-centered Design.* The platform will be designed for and with users, including both data producers and data consumers. The design process will be supported with constant feedback from users that will result in an iterative process of continuous improvement. Reference principles: 2, 3.
- 5. Data Disaggregation.* The platform will support improved access to, and use of, disaggregated data to focus on all segments of the population, including the most vulnerable. The system will allow the



management and dissemination of data disaggregated by subnational geographic areas, sex, age group, residence, wealth and income group, disability, ethnicity, migrant status, and other important socioeconomic characteristics. Reference principles: 2.

6. *Accessibility.* The platform will promote the use of data for policy and decision making at the local level. The system will support national languages. The user interface of the system will assist persons with disabilities. The system will be deployed across multiple types of devices, including notebooks, tablets, and mobile phones. The system will provide access on and off line. Reference principles: 2.

7. *Data Communication.* The platform will implement innovative strategies to improve the presentation, communication and use of data for sustainable development. The platform will support multiple ways to explore, represent and communicate data on statistical indicators, and address the needs and priorities of diverse groups of users, including policy-makers, private sector, the media, academia, and the public. The system will provide innovative data visualization and data story-telling modules. Reference principles: 1, 2, 3.

8. *Modularity and Extensibility.* The platform will be composed of interoperable modules that are loosely coupled. The data that these modules consume as inputs and produce as outputs will comply with open data standards and protocols. The platform modules will comply with standards, such as Statistical Data and Metadata Exchange (SDMX) and Common Statistical Production Architecture (CSPA). The platform will support extensibility through addition of modules or components, upstream or downstream. Reference principles: 2, 3, 4.

9. *Application Programming Interfaces.* The platform will provide standardized Application Programming Interfaces (APIs) in compliance with the OpenAPI Specification to facilitate data exchange. Reference principles: 3, 4.

10. *Scalability.* The platform will be able to support needs beyond the immediate or 'pilot' scope of application. By taking into account the needs and resources of different stakeholders, the design of the platform will facilitate adoption across regional and national statistical systems. Reference principles: 2, 3.

11. *Open Data.* The platform will comply with the Open Data principles to provide data to the public that can be freely used by anyone. Reference principles: 2, 3.

12. *Open Source Software Technology Stack.* The platform will comply with an Agile open source software development approach to encourage a community of developers to contribute to the platform. Standard operating procedures will be established and strictly followed for updating the technology stack. Reference principles: 2, 3.

13. *User Experience and Style Guide.* The modules of the platform should be developed in compliance with the DFA style guide. Standard operating procedures will be established for brainstorming about enhancements to the style guide, followed by an approval process to make changes. Reference principles: 2, 3.

14. *Component Library.* The modules of the platform will be developed from the DFA component library and in compliance with the DFA style guide. Standard operating procedures will be established and followed for updating the component library, including brainstorming about any new components, design, development, quality assurance, and integration. Reference principles: 2, 3.

15. *Back-end Development.* All platform modules will be developed using a common approach to database technology that decouples the back-end from the front-end to provide a database-agnostic way of interacting with the application. Reference principles: 2, 3



Section 6. Methodology

DFA provides a leading-edge information technology platform to leverage the role of the national statistical system to support the overall normative and analytical work required for successful monitoring of the SDGs by Member States. DFA promotes compliance with emerging international statistical standards endorsed by the UN Statistical Commission. These standards advance the technical processes for data and metadata storage and exchange, particularly in areas where the UN is providing global leadership and technical support, such as: the SDG indicator framework management, SDG indicator metadata, customization of SDG indicators to regional and national priorities, and SDG progress reporting.

DFA broadens the scope of data sources for effective SDG monitoring. The toolkit uses traditional reliable data sources (household surveys, population censuses, civil registration and vital statistics systems) and strengthened administrative management information systems (health, education). DFA also helps advance the use of new innovative data sources, such as geospatial data and, where relevant, crowdsourcing (disaster response planning), mobile phone data collection, smart devices, big data, and SMS polling.

The DFA method for development embraces a collaborative strategy using the DFA open source software toolkit. Under this software development methodology, the source code is made available under open source license to analyze, change and improve its design over time. The open source software license:

- Ensures that all contributions to the development of the tools are consistent with the overall requirements of regional and national development needs and UN global monitoring requirements, particularly the need to monitor disaggregated SDG indicator data;
- Ensures that investments made in software development and enhancements are shared with all stakeholders using these tools;
- Provides full ownership of the software to Member States for SDG monitoring;
- Results in an integrated suite of loosely-coupled reporting platform modules, each with its own purpose (data capture, curation, storage, analytics, visualization) and designed to be interoperable with other existing systems;
- Provides a mechanism for oversight and guidance of the software design process with the flexibility for NSO authorities to add customized modules while maintaining global standards for common data reporting.

This strategy for the development of an SDG data monitoring platform provides an efficient model for Member States to share the costs of the development of the modules of the system among all stakeholders of the system. When one country or one UN agency develops a new feature or module for the system, the new software will be reviewed, and if appropriate, made available to all countries using the system. In this way, Member States can benefit from the global DFA community of users.

Section 7. Validation Checklist

The DFA platform will be validated with reference to a checklist of desirable features and best practices for the development of national statistical systems, as described below.

Development Methodology

- Is the software developed with a collaborative approach? Are users driving the design process with persistent user feedback for continuous improvement?
- Is the development methodology cost effective development with shared code?
- Is the application scalable from prototype to nationwide use?
- Is the application developed with open source code?

National Ownership

- Is there support for complete government ownership with features to independently customize, adapt, and maintain the platform?
- Is there support to add, customize indicators for national development priorities?
- Is there support to add, customize indicator metadata to meet national requirements?
- Is there support to add, modify, maintain subnational administrative areas?
- Is there support to brand the platform without any reference to the solution provider?
- Can the platform be easily customized to fit specific needs of a country/organization?

Platform Architecture

- Does the platform conform with the Common Statistical Production Architecture (CSPA)?
- Is the system designed with loosely-coupled, interchangeable components?
- Is there support for adding new components?

Statistical Standards

- Does the platform comply with the Fundamental Principles for Statistics?
- Support for Statistical Data and Metadata eXchange (SDMX)?
- Support for Data Documentation Initiative (DDI Alliance) for metadata?
- Support for Open Geospatial Consortium standard?
- Support for International Aid Transparency Initiative (IATI) standard?

Indicator Frameworks

- Support for SDG indicator framework (linked to SDMX code lists)?
- Support for national development plan indicator frameworks?

Open Data

- Data available on the web (in any format) under an open license?
- Data available as structure data (CSV instead of scanned image of a table)?
- Data available in a non-proprietary open format (CSV instead of Excel)?

Data Management

- Support for Extract-Transform-Load (ETL) tools for data migration?
- Compliance with international standards for data security?
- Support for standardized data structure definitions for data harmonization?
- Support for data disaggregation? By subnational geographic areas? By residence? By sex, age group, wealth and income group, disability, ethnicity, migrant status? Support for user-defined data disaggregation?

Data Searching and Filtering

- Support for free-text search of the database?
- Support for search by indicators, geographic areas, time periods, data sources?
- Support for search by goals, targets, or other elements of indicator frameworks?
- Support to filter by data disaggregation?

Data Visualization and Communication

- Support for data visualization in tables, graphs, maps?
- Support for building data dashboards with interactive data visualizations?
- Support for user-generated data profiles (with drill down by geographic area)?
- Support for user-generated data storytelling?

Monitoring

- Support for user-defined results-based matrix (baselines/targets, planned/achieved)
- Support for longitudinal monitoring of baselines/targets?
- Support for user-defined indicator tags (gender, environment, human rights)

Interoperability

- Does the system allow exchange of data with other systems?
- Are data exchange protocols and formats documented?
- Does the platform have an API on which other

DATA FOR ALL

applications can rely? Are these APIs well documented? Can the platform incorporate data from other APIs?

- Does it lock in users through any proprietary protocols and formats?
- What relevant open standards does the system support?
- Does the system support integration with other upstream or downstream products? If it does, what are the mechanisms for such integration?

Licensing and Pricing

- Is there a free version of the system available?
- Are there any restrictions on the right to copy, install, modify and re-distribute the system?
- Is there any warranty or legal recourse?
- Do all the libraries used by the system have compatible licenses?
- Is free technical support and maintenance available? Or is technical support available through an annual paid subscription?

Deployment

- Cloud? (such as Amazon, Azure)
- Microsoft (Windows) – offline?
- Apple (MacOS) – offline?
- Mobile (iOS)?
- Mobile (Android)?

Technical Support

- Help desk (paid)?
- Help desk (free)?
- Community of users?
- Support for installation and configuration?
- Support for troubleshooting?
- User manuals?
- Administrator manuals?
- Quick reference guides?
- What arrangements are in place for the provision of support over the medium to long term?

Training

- On-site?
- E-learning courseware?
- Live online?
- Webinars?
- How-to videos?

Accessibility, Performance and Scalability

- Compliance with the Open Accessibility Framework?
- Accessible offline for use with no/limited internet access? Can the system operate offline (in certain locations) and upload

seamlessly once online connectivity is achieved?

- Support for smartphone users? Support for basic phone users (SMS, voice)?
- How intuitive (i.e., "user friendly") is the user interface? Is it easy to use without technical training? Is the platform easy to use by the general public?
- Does the system support multiple languages: English, French, Spanish, Chinese, Arabic, Russian, user-defined languages?
- How good is the quality of the user interface? Does the system conform to any user interface guidelines?
- What is the largest number of concurrent system users that can be supported?
- Are there any load and stress testing metrics? (e.g., server response time and render response time under specific scenarios)
- Can the system be deployed over small capacity infrastructures?

Reliability

- Are there any reports/metrics on how often does the system work/crash?
- Are there any output accuracy metrics?

Adoption

- How popular / widespread is the system compared to others?
- How popular / widespread is the underlying technology stack?

Maturity

- How long have the system and its developers been around?
- How mature is the underlying technology?
- Is the system actively maintained? How often are new versions and patches being released?
- Are there any communities of users and/or developers? Are there user feedback mechanisms in place?
- Are there multiple developers and service providers?

Security

- Does the system comply with security standards, policies, laws and regulations?
- Has the system state-of-the-art anti-virus and malware software installed and functioning properly?
- Does the system comply with specific security needs (e.g., registration, authentication, data disclosure prevention mechanisms)