Trainer guidance: Data publication

Learning materials provided by the Guide on Open Government Data (OGD) Action Planning for Sustainable Development can be used for sensitization and training of government officials and other OGD actors. This document aims to provide guidance to trainers who will use the learning materials on *Data publication* as a training module on this topic.

Module Overview		Notes
Type	Group work with short introductory lecture/presentation	Followed by Q&A moderated by facilitator
Time	120 minutes	This module contains two sessions, each: ○ 10 min introductory lecture/presentation followed by group work. ○ 10 min briefing on group work methodology ○ 70 min group work. Each group has one facilitator and one rapporteur ○ 10 min reporting back to plenary
Target Participants	All groups B to F (see below)	Max 30 participants, in groups of 5 to 6 per table
Prerequisites	None	
Materials	○ Selected and prepared datasets ○ Template repository for data publication ○ Annex 1: Sample data evaluation sheet	○ One laptop per group ○ One flipchart per group ○ Sticky notes, stickers and colored markers
Presentation	Slides English	Data publication
Further reading	OGDCE Guidelines	OGDCE Guidelines

Learning objective

By completing this training module, participants will learn about:
- Best practice for publishing data on the Internet
- Options for publishing data on the Internet
- Key features of data catalogues
- Using the key features of a data catalogue
- Adding an external dataset to the catalogue
- Uploading a dataset to the catalogue
- Updating the metadata of a dataset in the catalogue
- Removing data from a data catalogue

Group work setup

10 min	Lecture presentation on the topic - by Expert
10 min	Group work introduction

70 min   Group work - facilitated by Facilitators (see "methodology" below)
30 min   Presentation and discussion of result - by Rapporteurs (see "methodology" below)

Group work flow

Exercise 1. Data portal (35 min)
- Register a user at www.datahub.io
- Join Organization www.datahub.io/organization/undesa
- Explore features in the data catalogue using the CKAN feature tour

Exercise 2. Data publication (35 min)
- Upload a prepared dataset
- Add a new dataset using external data source
- Edit the metadata information of both datasets
- Remove the datasets from the data catalogue

Metadata exercise

Context, which is usually contained in metadata, will give re-users a better understanding of a dataset. To this end, the following can be included in the metadata:

- Title
- Description
- Keywords
- Legal information
- License and attribution required
- Geographical and temporal coverage
- Granularity and completeness
- Date of publication, last modification and update frequency

All these points will help users know what a dataset contains, how to interpret it and if it should be trusted. In some cases, more detailed documentation on the data might be desirable, such as specific methodologies or methods for its collection, explanations on the way it might have been aggregated, methods used for anonymization, specific recommendations for interpretation of the data, etc.

Guiding Questions

By completing this module, participants will be able to answer the following guiding questions:
- What are the options for publishing data on the Internet?
- What are the key features of data catalogues?
- How to add, upload and update a dataset to the catalogue?
- Why is metadata important, and how to use it?

Expected outputs

Through this module, participants will be able to:
- Use the key features of data catalogues
- Add, upload and update datasets to a data catalogue
- Understand the importance of metadata and how to use it

Methodology

This module will be conducted by one expert on OGD for sustainable development and a group of facilitators at a ratio of one facilitator per ten participants. The expert will introduce the module and provide an overview on the methodology, learning objectives, materials to be used, expected outcomes, and required documentation and reporting. The expert will then deliver a 10-minute presentation on the subject.

After the presentation, participants will split in small groups of 5 to 6 people for hands on group work for 45 to 70 minutes. Each group will have a facilitator who will facilitate the discussions, help the group meet the learning objectives, use the materials and produce the expected outputs. Each group will have one laptop with Internet access and standard office software. Each group will nominate one group member as rapporteur. The rapporteur will take notes and capture the main discussions and outputs of group work. After the group work, the rapporteurs from each group will summarize the discussions and present the outputs to the plenary (5 minutes for each group). The presentations from the rapporteurs will be followed by a Q&A session moderated by a dedicated facilitator.

Target Participants

A. Government leaders (top-level, policy/decision makers)
B. Government specialists and technicians (mid-level, ICT department, statistical office)
C. Civil Society Organizations (CSO) representatives
D. Independent journalists and representatives from media organizations
E. Academia (public administration, informatics, statistics)
F. Private Sector (ICT companies and start-ups)