Statistical capacity development in a changing environment

RALF BECKER

UNITED NATIONS STATISTICS DIVISION

10th Session of OIC-Statcom

6 Oct 2021

UNSD activities

- UNSD capacity development activities have always focused on:
  - Support of institutional setup and functioning of NSOs and NSSs
  - Support in specific statistical domains

- New training needs
  - Support to respond to the COVID-19 pandemic
    - Support to maintain statistical production activities under new conditions / restrictions
    - Support not only focused on SDG-related areas

How are training activities changing? (1)

- New forms
  - Online training courses
    - Not as result of COVID alone
      - Advances in underlying technology and access
  - UNSD uses an e-learning platform
    - Various methods of e-learning are employed
  - Not appropriate for all trainings
  - Requires a change in training styles
    - No simple transfer of in-person to online training
    - UNSD works on best practices to facilitate common approach

How are training activities changing? (2)

- New forms – Caution
  - Many current online activities are not a full replacement of traditional activities (e.g. for a one-week workshop)
  - Easier for one-directional activities
    - Activities requiring interaction become difficult
  - Positive: large audience
  - Negative: reduced feedback and discussion
  - Technical difficulties in many areas / regions

How are training activities changing? (3)

- New topics
  - Statistics production in new environment
    - E.g. alternatives to traditional in-person data collection
  - New data sources
    - Admin data
      - Several projects underway to establish guidelines
    - Big data
      - Many projects on specific topics and general training programmes

How are training activities changing? (4)

- New approaches
  - How do we develop training courses?
  - For big data (and other fields):
    - Development of competency frameworks
    - Assessment of status and needs via Maturity matrices
    - Identification of skill gaps
    - Deriving of training needs
    - Consistent approach to define course parameters
      - Learning objectives, prerequisites etc.
- (training previously supply-driven, now consider more the demand)
- For big data: https://unstats.un.org/bigdata/task-teams/training/index.cshtml

UNBigDataLearning UNGP LMS English (en)

You are not logged in. (Log in)

UN Global Platform - Learning Management System

Search courses Go

Course categories

Collapse all

- Big Data
  - Privacy preserving techniques (2)
  - Automatic Identification System (AIS) (2)
- System of Environmental Economic Accounting (SEEA) (13)
- Social and Demographics Statistics (1)
- Sustainable Development Goals (SDGs) (2)
- Other courses (1)

New characteristics of training courses (1)

- Traditional face-to-face workshops will continue
- But e-learning courses will play a larger role
  - As stand-alone trainings
  - As preparations for face-to-face meetings
  - As part of hybrid events

New characteristics of training courses (2)

- E-learning courses allow for more nuances in addressing training needs
- Courses will be designed by targeting:
  - More detailed topics than in the past
  - Specific persona (user types)
    - Senior managers, project managers, statisticians, data scientists, programmers etc.
  - Specific skill levels
    - Foundation, Intermediate, Advanced etc.

  - Such distinctions would be impossible for face-to-face meetings due to resource limitations

New characteristics of training courses (3)

- Competency frameworks identify skills needed in different statistical areas
  - Support also individual progression of skills
- Maturity matrices identify gaps at the institutional level
  - Allow identification of training needs
- Feedback from these tools allows for priority setting of courses to be developed

- These tools are being developed for different areas, but are most advanced for big data topics

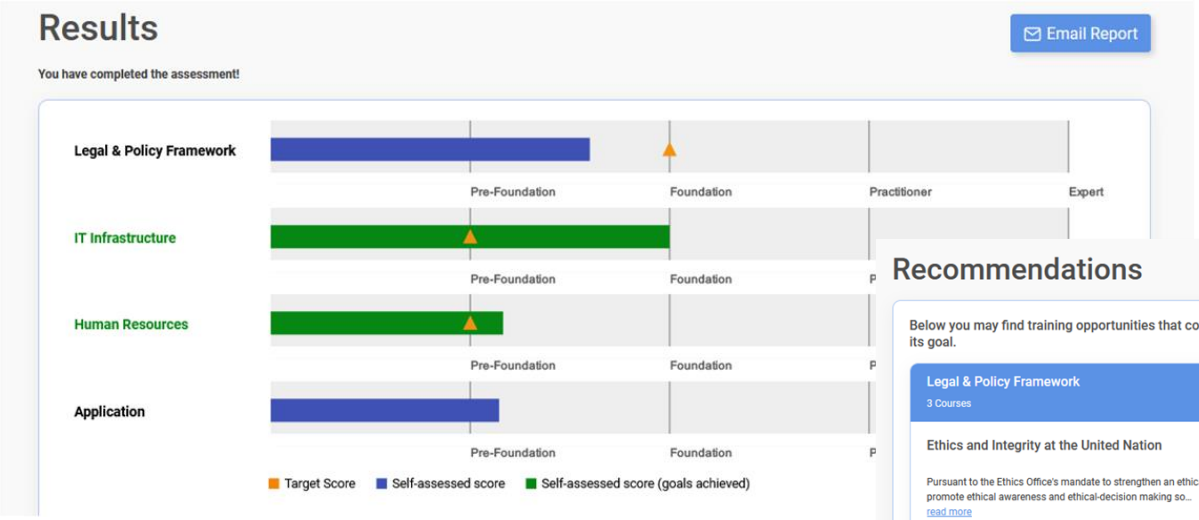
Competency framework (Big Data)

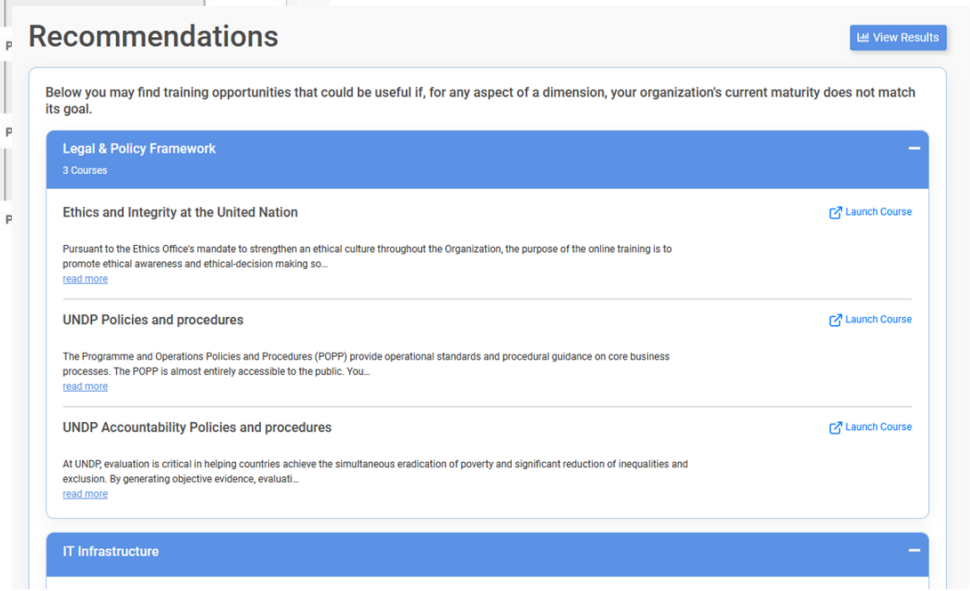
Dimension 1			
Name of the area	**Data management**		
Dimension 2			
Competence title and description	To possess data management knowledge in a range of below-listed issues: 1) Database systems: database management systems, data models – definition and types, entity relationship model, models implementation (pre-relational, relational and object-oriented models) 2) Basics of cryptography: hash function, binary tree 3) Database: relational database, tabular data, data frames and series, shard, on-line analytical processing, data warehousing, data lakes, data vaults, logical multidimensional data model, extract, transform and load (ETL), NoSQL 4) Varied data formats: (Json, shp, XML, csv)		
Dimension 3	A - Foundation	B - Intermediate	C - Advanced
Proficiency levels	Demonstrate knowledge and understanding basic data management skills.	Demonstrate knowledge and understanding of, data base management tools and methods, and ability to apply some of them.	Thorough knowledge of proficiency in data base management and skillfulness in performing operations on varied data sets. Is able to advise others in finding data management solutions.
Dimension 4			
Knowledge examples	• Know the basic concept of SQL and NoSQL databases (such as table, column, field, field type, primary and foreign key) • Understand the consequences of using the hash function • Define functional dependencies occurring among the analyzed data		
Skills examples	• Able to create database structures in selected database management systems (e.g. MySQL, MongoDB, *more in annex*) • Able to present the logical structure of the database using tables and graphical relationships in selected programs • Apply ETL techniques - acquisition, processing (including pre-purification) and loading data from non-statistical sources		
Attitude examples	• Systematically supplement knowledge of new trends in the field of computer science on the subject of computer data storage • Identify data sources and assess their usefulness in complementing studies at hand • Carefully analyze the data and adjust them to the needs of database users		

Competency framework (Big Data)

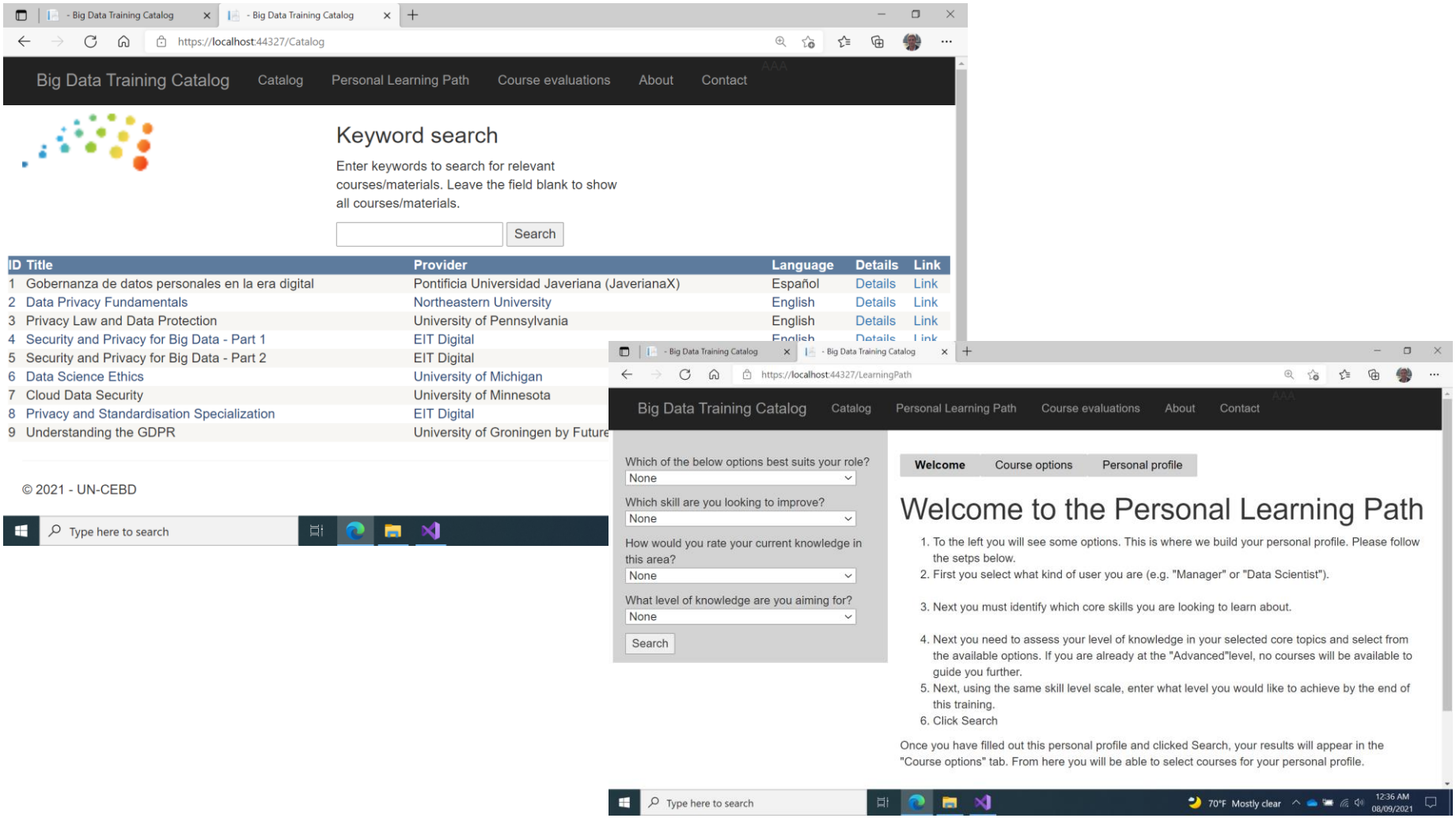
	Data acquisition	Data processing	Data analysis	Data visualization
Core competencies	Ethics and privacy	Ethics and privacy	Ethics and privacy	Ethics and privacy
	Data management	Data management	Mathematics	Statistics
	Machine Learning	Mathematics	Statistics	Programming
	Programming	Programming	Programming	Data visualization
		Machine Learning	Machine Learning	
Generic skills	product understanding	curiosity	curiosity	product understanding
	critical thinking	business acumen	adaptability	business acumen
	business acumen	critical thinking	critical thinking	storytelling
	curiosity	communication	communication	communication
	team player	team player	team player	team player
	agile project management	agile project management	agile project management	agile project management

NSO Maturity Matrix (Big Data)





Training catalog (Big Data)



Thank you