The Competency Framework

• General guidance for the NSOs on building Big Data skills
• Extensive (yet not exhaustive) set of skills and knowledge useful for acquiring and processing big data
• Recognition that:
  • different NSOs run different projects
  • different types of data specialist (e.g. data analyst, data engineer, data scientist, etc.) require different compositions of skills and knowledge
  • not each data specialist must possess all of the skills listed in the framework

• Two approaches:
  • Competency areas approach
  • (Simplified) statistical production process approach
Competency areas approach - core competencies

- Ethics and Privacy
- Mathematics
- Data Management
- Statistics
- Programming
- Machine Learning
- Data Visualization
# Competency areas approach – the format

<table>
<thead>
<tr>
<th>Dimension 1</th>
<th>Name of the area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dimension 2</td>
<td>Competence title and description</td>
</tr>
<tr>
<td>Dimension 3</td>
<td>A - Foundation</td>
</tr>
<tr>
<td>Proficiency levels</td>
<td></td>
</tr>
<tr>
<td>Dimension 4</td>
<td>Knowledge examples</td>
</tr>
<tr>
<td></td>
<td>Skills examples</td>
</tr>
<tr>
<td></td>
<td>Attitude examples</td>
</tr>
<tr>
<td>Dimension 1</td>
<td>Data management</td>
</tr>
<tr>
<td>-------------</td>
<td>----------------</td>
</tr>
<tr>
<td><strong>Name of the area</strong></td>
<td>Data management</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dimension 2</th>
<th>To possess data management knowledge in a range of below-listed issues:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Competence title and description</strong></td>
<td>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)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dimension 3</th>
<th>A - Foundation</th>
<th>B - Intermediate</th>
<th>C - Advanced</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Proficiency levels</strong></td>
<td>Demonstrate knowledge and understanding basic data management skills.</td>
<td>Demonstrate knowledge and understanding of, data base management tools and methods, and ability to apply some of them.</td>
<td>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.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dimension 4</th>
<th>Knowledge examples</th>
<th>Skills examples</th>
<th>Attitude examples</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Knowledge examples</strong></td>
<td>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</td>
<td>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</td>
<td>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</td>
</tr>
</tbody>
</table>

- **Big Data UN Global Working Group**
Generic (soft) skills

Indispensable and inherent part of Big Data competencies catalogue

Span all stages of the statistical production process

*Not the main focus of this Competency Framework
Generic skills

- Agile Project Management
- Adaptability
- Business Acumen
- Communication
- Critical Thinking
- Curiosity
- Product Understanding
- Storytelling
- Team Player
Simplified production process approach

CORE COMPETENCIES

DATA ACQUISITION  DATA PROCESSING  DATA ANALYSIS  DATA VISUALIZATION

GENERIC (SOFT) SKILLS
<table>
<thead>
<tr>
<th>Core competencies</th>
<th>Data acquisition</th>
<th>Data processing</th>
<th>Data analysis</th>
<th>Data visualization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethics and privacy</td>
<td>Ethics and privacy</td>
<td>Ethics and privacy</td>
<td>Ethics and privacy</td>
<td>Ethics and privacy</td>
</tr>
<tr>
<td>Data management</td>
<td>Data management</td>
<td>Data management</td>
<td>Mathematics</td>
<td>Statistics</td>
</tr>
<tr>
<td>Machine Learning</td>
<td>Mathematics</td>
<td>Statistics</td>
<td>Programming</td>
<td>Programming</td>
</tr>
<tr>
<td>Programming</td>
<td>Programming</td>
<td>Programming</td>
<td>Programming</td>
<td>Data visualization</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Generic skills</th>
<th>product understanding</th>
<th>curiosity</th>
<th>business acumen</th>
<th>adaptability</th>
<th>product understanding</th>
<th>business acumen</th>
<th>storytelling</th>
<th>communication</th>
<th>communication</th>
<th>communication</th>
<th>team player</th>
<th>team player</th>
<th>team player</th>
<th>team player</th>
<th>agile project management</th>
</tr>
</thead>
<tbody>
<tr>
<td>agile project management</td>
<td>agile project management</td>
<td>agile project management</td>
<td>agile project management</td>
<td>agile project management</td>
<td>agile project management</td>
<td>agile project management</td>
<td>agile project management</td>
<td>agile project management</td>
<td>agile project management</td>
<td>agile project management</td>
<td>agile project management</td>
<td>agile project management</td>
<td>agile project management</td>
<td>agile project management</td>
<td>agile project management</td>
</tr>
</tbody>
</table>
On-going & future work - training development & provision

- On-line training to be developed by subject matter UN GWG Task Teams (earth observation, AIS, scanner data, mobile phone data)
- Guidelines by TT on Training, Competencies and Capacity Development
- Link of the skills acquired during training to the Competency Framework
- Ensuring global accessibility of the training
- Additional work: a tool to help the learner to select the right training for their needs
Useful Links

• UN Big Data: https://unstats.un.org/bigdata/
• Task Team products: https://unstats.un.org/bigdata/taskteams/training/

Contacts:
• Maturity Matrix test input - Ceri.regan@ons.gov.uk
• Competency Framework feedback - Do.nowak@stat.gov.pl

THANK YOU.