# **UNIVERSITY OF CAPE COAST**

# **COLLEGE OF HUMANITIES AND LEGAL STUDIES**

# **SCHOOL OF ECONOMICS**

# **DEPARTMENT OF DATA SCIENCE AND ECONOMIC POLICY**

# 

# **Course Code: DMA 820S**

# **Course Name: DATA CURATION AND MANAGEMENT PLAN**

# **Lecturer: DR. RAYMOND E. KOFINTI (PhD)**

**CECILIA ACHEAMPOMAAH DWOMOH SE/DMD/24/0026**

**⁠JOSEPH ASARE SE/DMD/24/0019**

**PRISCILLA BEDZRA SE/DMD/24/0005**

**RAZAK BABA MOHAMMED SE/DMD/24/0004**

**Identification of some of the global open data sources and demonstrating how to access data from at least two of the several sources.**

Open data sources are repositories or platforms where data is made freely available for anyone to use, reuse, and share — often under minimal or no restriction. Governments, international organizations, research institutions, or NGOs usually maintain them.

Open data supports transparency, accountability, collaboration, and innovation by providing researchers, policymakers, journalists, and the public with direct access to information that would otherwise be kept siloed.

Collaboration – Global open data fosters partnerships among researchers, institutions, and innovators across national and disciplinary boundaries. By providing a shared foundation of standardized, accessible datasets, it removes informational barriers and facilitates the co-creation of knowledge and technology. This collective approach accelerates problem-solving on complex global challenges such as climate change, health crises, and sustainable development.

Transparency – Open data enhances public insight into critical global issues, including economics, public health, and environmental change. By making accurate and timely information freely available it enables independent verification, promotes informed debate, and fosters trust between institutions and stakeholders. Example, Covid-19

Innovation – The accessibility of global datasets fuels the development of novel research, applications, and technological solutions. Entrepreneurs, scientists, and civic organizations can leverage open data to design predictive models, create data-driven services, and develop tools that address emerging societal needs.

Evidence-Based Decision-Making – Open data strengthens policy formulation and program design by grounding them in verifiable facts. Governments, NGOs, and international agencies can analyze global datasets to identify trends, forecast outcomes, and allocate resources effectively, thereby improving policy relevance and impact.

Accountability – By making progress indicators, budgetary records, and performance metrics openly available, global open data enables stakeholders to track commitments—such as the Sustainable Development Goals (SDGs)—and hold institutions accountable. This reduces opportunities for mismanagement and reinforces the principle of answerability in governance.

Examples of Global Open Data Sources

1. World Bank Open Data – https://data.worldbank.org
   * Global development indicators (economics, health, education, environment).
   * Offers both a web interface and an API for programmatic access.
2. UNdata (United Nations Statistics Division) – https://data.un.org
   * Consolidated UN datasets covering demographics, trade, social statistics, and more.
3. WHO Global Health Observatory – <https://www.who.int/data/gho>
   * Global health indicators and disease statistics with downloadable data and OData API.
4. International Monetary Fund (IMF) Data – https://data.imf.org
   * Macroeconomic statistics such as GDP, inflation, and financial accounts.
5. Organization for Economic Co-operation and Development (OECD) Data – https://data.oecd.org
   * Economic, social, and environmental indicators for OECD and partner countries.
6. Food and Agriculture Organization (FAO) FAOSTAT – <https://www.fao.org/faostat/en>
   * Agricultural production, trade, food security, and climate-related data.
7. Humanitarian Data Exchange (HDX) – https://data.humdata.org
   * Humanitarian and crisis-related datasets, including from UN OCHA and NGOs.
8. Eurostat – https://ec.europa.eu/eurostat
   * Statistical data for European countries across all policy areas.

Step-by-Step Guide for Accessing Data from Global Open Data Systems

This guide demonstrates how to access and download datasets from two UN open data systems: UNData (United Nations Statistics Division) and FAOSTAT (Food and Agriculture Organization). These platforms are suitable for MSc research projects, as they provide internationally recognized indicators, time-series data, and country-level statistics.

Accessing Data from World Bank Open Data

1. Go to https://data.worldbank.org.
2. Search for an indicator — for example, *GDP per capita (current US$)*.
3. Click on the result to open its data page.
4. Select a country (e.g., Ghana) and time range (e.g., 2010–2020).
5. Click Download and choose CSV or Excel to save the dataset.

Accessing Data from UN Data (United Nations Statistics Division)

1. Open the Portal: Go to https://data.un.org
2. Search for the Dataset: Use the search bar to type a keyword (e.g., 'urbanization, Ghana').
3. Open the Dataset: Click the dataset link to open detailed statistics.
4. Download the Data: Click the Download button and select CSV or Excel format.

WHO Global Health Observatory (GHO)

1. Go to <https://www.who.int/data/gho>.
2. Click **Data** and browse *Indicators* or *By Country*.
3. Select an indicator, e.g., *Life expectancy at birth*.
4. Filter by country and year, then click **Download** to get CSV or Excel.

Accessing Data from FAOSTAT (Food and Agriculture Organization)

1. Open the Portal: Go to https://www.fao.org/faostat
2. Select a Database: For example, choose 'Production' → 'Crops and livestock
3. products'.
4. Filter the Data: Country = Ghana, Item = Maize, Element = Production Quantity, Years = 2000–2023.
5. Download the Data: Click Download Data and choose CSV or Excel format

### **Key Risks of Global Open Data**

#### **a. Privacy and Confidentiality Breaches**

Even when data is anonymized, advances in data analytics and cross-referencing with other datasets can lead to re-identification of individuals. This can expose sensitive personal information, such as health records, income levels, or political affiliations.

#### **b. Misinterpretation and Misuse of Data**

Open datasets can be misinterpreted by non-experts, leading to flawed conclusions, misinformation, or policy errors. In extreme cases, data can be intentionally manipulated to support biased narratives.

#### **c. Data Quality and Reliability Issues**

Not all open datasets undergo rigorous quality control. Outdated, incomplete, or inaccurate data can compromise research validity and lead to misguided decisions.

#### **d. Cybersecurity and Data Integrity Risks**

Open data platforms may become targets for cyberattacks, data manipulation, or corruption of files, potentially undermining trust in the dataset.

#### **e. Intellectual Property and Licensing Conflicts**

If data is improperly shared without respecting original copyright, licensing terms, or attribution requirements, it can lead to legal disputes.

#### **f. Economic and Strategic Risks**

In sensitive sectors such as defense, energy, or natural resources, making certain datasets openly available could inadvertently reveal strategic vulnerabilities or be exploited by malicious actors.

### **3. Mitigation Strategies**

#### **a. Strong Data Governance Policies**

Establish clear protocols for what data can be made public, including privacy thresholds, ethical guidelines, and licensing rules.

#### **b. Robust Anonymization and De-identification**

Apply advanced anonymization techniques such as data masking, aggregation, and differential privacy to protect personal identities.

#### **c. Metadata and Documentation Standards**

Publish datasets with comprehensive metadata that explains methodology, scope, limitations, and intended uses to reduce misinterpretation.

#### **d. Data Quality Assurance Mechanisms**

Implement regular validation, peer review, and updates to ensure accuracy and timeliness.

#### **e. Access Controls for Sensitive Data**

Adopt tiered access models where highly sensitive datasets are shared under restricted agreements rather than full public release.

#### **f. Capacity Building and Data Literacy**

Invest in training programs for policymakers, researchers, and the public to interpret and use open data responsibly.

### **Conclusion**

### Global open data holds transformative potential for advancing knowledge, fostering transparency, and stimulating innovation. However, the associated risks — particularly around privacy, quality, and misuse — mean that robust governance, technical safeguards, and ethical considerations must guide its release. A balanced approach, where openness is coupled with protection, ensures that open data can serve the global good without compromising security or individual rights.

References

World Bank. (n.d.). *World Bank Open Data*. [https://data.worldbank.org](https://data.worldbank.org/)

United Nations Statistics Division. (n.d.). *UNdata*. [https://data.un.org](https://data.un.org/)

World Health Organization. (n.d.). *Global Health Observatory (GHO)*

*data*. <https://www.who.int/data/gho>

International Monetary Fund. (n.d.). *IMF Data*. [https://data.imf.org](https://data.imf.org/)

Organisation for Economic Co-operation and Development. (n.d.). *OECD Data*. [https://data.oecd.org](https://data.oecd.org/)