NAVIGATING UNITED NATIONS FREELY AVAILABLE DATA SOURCES

INTRODUCTION

The number of data resources available from the United Nations is vast. It is easy to be consumed by information published by the UN and its agencies, the World Bank, and the Organization of Economic Cooperation & Development (OECD). Academic librarians can assist diplomacy and international relations students to be focused in their search for relevant open access data with an emphasis on proper citation of publically available data. The purpose of this article is to highlight current UN data sources that can be used by academic librarians to enhance international relations, public policy, and political science research. This article describes where these resources exist, how to extract the data students and faculty need from the resources, and how to cite this information properly. These resources are important because they are freely available and contain the latest data from UN agencies and related intergovernmental organizations (IGOs) and non-governmental organizations (NGOs). Additionally, promoting statistical and data literacy is an important part of the research equation for academic librarians.

University Libraries at Seton Hall University supports the School of Diplomacy and International Relations, which has a total of 320 undergraduate and 110 graduate students. Faculty research and courses vary widely and therefore resources for each course need to be customized. To support the School of Diplomacy, University Libraries provides Research Guides for United Nations focused research (library.shu.edu/un), classroom instruction and research consultations. University Libraries also collaborates with the School of Diplomacy for campus events and recruiting speakers. Additionally, providing statistical sources and data visualization training are becoming important focal points for Diplomacy students and faculty.

HIGHER ED LANDSCAPE

In *Research Data Services in Academic Libraries: Data Intensive Roles for the Future?,* Tenopir's research states "academic librarians are now often integrated as partners in all aspects of the research process, from data collection to publication and preservation of research output" (2015, p. 4). She writes that research data services (RDS) are "needed by academic communities is emerging in academic libraries in response to the growth of data intensive research." This data intensive research is supported by the data resources described in this article.

The New Media Consortium's (NMC) Horizon Report: 2016 Higher Education Edition[1] describes technology trends five years out for higher education institutions. The report states that academic libraries facilitate ways for students to build confidence in using technologies for learning. Visualization technologies are identified as a 5-year trend, range from "simple infographics to complex forms of visual data analysis" (2016, p. 35). These technologies "tap the brain's inherent ability to rapidly process visual information, identify patterns, and sense order in complex situations". Data visualization will be discussed later in this article. There are many free tools available from the UN and related agencies. The NMC Horizon Report supports the online trend of free visualization tool that support creation of infographics, interactive maps/timelines, and other digital humanities tools "that generally making the complex simple" (2016, p. 35).

Martin Edwards, Associate Professor and Director, Center for United Nations and Global Governance Studies, School of Diplomacy and International Relations agrees. In an interview on July 18, 2016, Edwards stated, "We are in a world that is increasingly awash in data. Data is frequently no longer proprietary, so students in other countries have the same access to data as

[1] This report is jointly conducted by the New Media Consortium (NMC) and the EDUCAUSE Learning Initiative (ELI), an EDUCAUSE Program.

students in the United States. This creates an important challenge for students regardless of level: they will need to become acclimated to deciphering trends in data to help their organizations make better decisions. This means that becoming statistically literate is an essential skill for the 21st century workforce."

Investigating statistical literacy at UK universities, Carter et al. report that secondary data on the open web "provides both opportunities and challenges for researchers and learners in higher education" (2011, p.223). They also describe that improving statistical literacy in academia and with the public is "recognized by research councils, statistical organizations and other bodies" (2011, p.224).

GETTING STARTED

If you are supporting students and faculty who need data for international relations or diplmacy-centric assigments or research, librarians need to ask questions to determine the most appropriate data source. These data sources may focus on a specific agency, UN initiative, indicator or general topic. This article focuses on freely data available sources. That said, search strategies discussed in this article can be used with paid data sources.

Carter, et al. described commonly used data sources for UK undergradates including macroeconimic indicators, World Development Indicators (WDI), IMF's International Financial Statistics (IFS) and WDI data. For Master's level research, these sources include WDI for dissertations and statistical packaages and Excel for running regressions with larges amounts of data (2011, p.236).

Datasets that are available may have limitations for dates or regions. Librarians can help students to shift student expecations for data that may not have been published for recent years or may not exist for a country but may exist for a region, or vice versa. More specifically, a trade

data source may exist in the UN Comtrade Database (http://comtrade.un.org/), but perhaps only through 2014. Trade data may be available by region instead of country or it may exist for some countries but not all. The UN Statistical Division has data for its branches that range from trade statistics and demographics to energy, gender statistics, and geospatial information. Following the appropriate UN agencies can keep you updated about hen new data is released on feeds that include @UNdata, @UNStats, and @UNLibrary.

There are hundreds of UN Agencies and academic librarians can guide students to the appropriate one from this useful glossary: www.un.org/en/sections/about-un/funds-programmes-specialized-agencies-and-others/index.html. Students will research current UN Initiatives to understand the priorities of the Security Council and General Assembly. Data about UN priorities such as the Sustainable Development Goals (SDGs) are outlined in the SDG Indicators Global Database (http://unstats.un.org/sdgs/indicators/database/). This is part of UNstat.org, the website of the UN Statistical Division.

Global indicators can be found United Nations (for Sustainable Development), World Bank, (World Development Indicators) and OECD (from Agriculture to Society) data websites. On oecd.data.org, Average Wages data is available. Data for 7 countries is available. Results can be downloaded into a *.csv (comma separated value) for viewing and analysis in Excel or other statistical software for further analysis.

Topics are more general search terms and many UN Agency websites help guide students to key issues related to an agency or initiative. This is an excellent way for students to determine what topics are most relevant to their research question.

In information literacy courses and in research consultations, diplomacy students are encouraged to employ current lingo used by UN agencies *and* information technologists. This

would couple search terms (SDGs, sustainable development, poverty) with a combination of these terms: data, dataset, repository, digital library, datasets, data visualization, digital library or infographic. When using these terms with the keywords (agency name, initiative, topic or indicator) students will gain data literacy skills to identify appropriate data sources. UN agencies use a variety of terminology, but most use "Resources" as a major heading on their websites. There is a good chance a data section will be located here.

To find digital collections, search terms may include: archives, repository, finding aid and digital collections. For example, in environmental research, the United Nations Environment Programme Knowledge Repository (http://www.unep.org/publications/) is an excellent resource for global environmental issues including climate change and ecosystem management. While publications are listed, data is not (as of publication). A search for "dataset" within this site links to over 9300 sources for data including Gender-Environment Datasets with Broadest Country Coverage which is available at http://web.unep.org/ggeo/resources/data-sources. This also links to the Food and Agriculture Organization (FAO) of the United Nations – Statistics Division (http://faostat3.fao.org/home/E) which is another repository for production, trade and emissions data related to agriculture and land use. These two sites are examples of what data sources focused searching with relevant terminology can identify for students.

One final example is Country Data for Pakistan available at: http://uneplive.unep.org/country/index/PK#data_tab. There are several datasets available. One option is Resource Efficient Indicators - Natural Resources. When this is selected, a Microsoft Excel window appears. Once the user selects "ok", they instantly have a free spreadsheet that is available for data analysis. The spreadsheet shows "Natural Resource Use" and materials and

indicates whether these materials match up with SDG indicators. As Dr. Edwards stated, this data is freely available to students around the globe. This data can be edited or put into SPSS. Again, students will improve their data literacy by understanding what a *.csv (comma separated value) file format is and using this format to open and save data in Excel for analysis or transfer into other statistical packages. A proper citation of this dataset is included later in this article.

In the Economic Commission for Latin America and the Caribbean's (ECLAC) Digital Repository, there is a section for Data & Statistics (http://www.cepal.org/en/datos-y-estadisticas) that contains the La Comisión Económica para América Latina's CEPALSTAT. CEPALSTAT is the gateway to all the statistical information of Latin America and the Caribbean countries collected, systematized, and published by ECLAC. REDATAM (**RE**trieval of **D**ata for Small **A**reas by **M**icrocomputer) is also included and creates and processes hierarchical databases from censuses, surveys, vital statistics and other sources for local, regional, and national analyses, and produces thematic maps. This is a data rich resource for students and faculty researching Latin American countries. Statistics and indicators, country profiles and infographics are available here.

DATA REPOSITORIES

The broadest view of UN data is found in data repositories. **UNdata** is a statistical portal that allows simultaneous searching of 24 statistical databases from United Nations Statistics and Population Divisions, UN specialized agencies (International Labour Organization, FAO, World Health Organization) and Member Countries (statistics collected by the national statistical agencies of member countries.) It is a meta-search of statistics collected from all those various bodies.

unstats.un.org is the website of the UN Statistical Division, within the Department of

Economic and Social Affairs (DESA) of the UN Secretariat. It contains the statistics collected by the statistical division, publications produced by the statistical division and information about the work of the division. It will not have the statistics collected by the UN specialized agencies or the Member Countries.

UNdata is a database produced by the UN Statistical Division. If researchers are looking for purely statistical data, this would be the database to use. The Statistical Division Website (unstats.un.org) will give you the data (once again, only from what the UN statistical division collects) as well as reports in which this data is used.

Finally, the United Nations Conference on Trade & Development (UNCTAD) has created UNCTADStat, which houses data and country profiles including over 150 indicators and statistical time series at http://unctadstat.unctad.org/EN/

The Humanitarian Data Exchange (HDX) https://data.humdata.org/ is an open data-sharing platform maintained by the UN's Office for the Coordination of Humanitarian Affairs (OCHA). The goal of HDX is to make humanitarian data easy to find and use for analysis. Launched in July 2014, HDX has been accessed by users in over 200 countries and territories. This site can be searched by data, location, organization or specific crisis such as El Nino or ebola. For example, OCHA Afghanistan and the World Food Programme have over 20 data sets each. This is useful for diplomacy and global or public health students. Natural disaster incidents from 1 January to 31 December 2012 are available in an Excel format for download. Students can then manipulate data by disaster or Province. As described earlier, The World Bank's data repository is located at http://data.worldbank.org/. OECD data is found at https://data.oecd.org/.

MANAGING DATA

It is important to note that the repositories and agency websites will allow the users to manage data by country or initiative, or part of an initiative and download the data into an Excel spreadsheet into *.csv format. This allows students to then import data into Excel or SPSS or Stata for further analysis. Data information literacy is important for undergraduates and graduate students and will help them secure internships, placement in graduate programs and ultimately, employment.

In spring 2016, I collaborated with Dr. Nabeela Alam, Assistant Professor of International Economics and Development, for two sections of sustainable development (DIPL 3201). Juniors and seniors were working on group projects where they had to choose a country, identify an issue related to sustainable development for that country, and devise a policy. The students needed an overview of data but did not have a lot of experience working with datasets. I worked with Professor Alam to give students a tour of World Development Indicators (World Bank), OECD, and Database of Political Institutions (World Bank). The results of this collaboration can be seen at library.shu.edu/countryresearch. Organizations, in both policy and business, rely increasingly on data-based decision making. In addition to data management and statistical analysis, the data skills required include interpreting trends and analyses, and translating findings into actionable policy items.

During an interview on July 18, 2016, Dr. Alam stated, “working with data helps students develop a more in-depth knowledge of the countries or phenomena they study. In my sustainable development class, students track key economic indicators to assess progress on 17 Sustainable Development Goals operationalized through 169 specific targets on outcomes such as inequality and carbon dioxide emissions. Building standard and alternate indices demonstrates the importance of looking more deeply at indicators.” Dr. Alam requires students to use economic

databases in undergraduate courses in trade (DIPL 4555) and sustainable development (DIPL 3201). Dr. Alam states, "When coupled with theories from economic development, working with data helps student understand how to use data in crafting policies and also the importance of monitoring and evaluation of the policy outcomes. While we conduct individual exercises through the semester, I expect students to use relevant data in their group projects to identify an urgent policy issue in sustainable development and develop an evidence-based set of policy recommendations. These data analysis skills are transferable to other domains such as business, politics, and national security to name a few."

Historians, public health experts and political scientists make use of the UN Human Development Index, http://hdr.undp.org/en/data, which allows experts to keep up (through 2015 at present) with data relevant to health, education, and standard of living, and to compare a country of expertise with most others in the world.

UNITED NATIONS LIBRARIES

The Dag Hammarskjöld Library at the UN in New York is an excellent resource for academic librarians. Research guides are located at http://research.un.org/en or academic librarians can direct questions to: ask.un.org. The librarians are responsive, subject matter experts and open to helping academic librarians improve their understanding of UN data (and all UN related research).

Research guides from the UN Libraries that will help guide you and your students for general and statistics are available at:

United Nations Library Geneva - http://libraryresources.unog.ch/statistics

Dag Hammarskjöld Library Research Guides - http://research.un.org/en

United Nations Statistical Sources - http://research.un.org/az.php?t=3701

http://research.un.org/c.php?g=98277&p=636126

Statistics Resource Guide - Statistics by Theme -

http://libraryresources.unog.ch/c.php?g=462654&p=3162836

Digitized UN documents are a complement to data sources, so what is the best way to determine what UN documents are digitized? Digital projects are agency dependent. Twitter feeds often publish the most recent updates to UN agency dataset. Additionally, General Assembly and Security Council Minutes have been digitized from 1946-1993 (check again) and updates are available at:

http://www.un.org/depts/dhl/deplib/digitization_program_status_report.pdf.

CITATIONS

It is important to guide students to proper citation practices for data sets. They should point readers to raw data by providing a web address (use "Retrieved from") or a general place that houses data sets on the site (use "Available from"). Here are examples from sources reference in this article.

Example: Quantitative Data Set from UNEP:

United Nations Environment Programme (UNEP). (2016). Pakistan - Resource Efficient Indicators - Natural Resources [Data File]/ Retrieved from http://uneplive.unep.org/country/index/PK#data_tab.

For graphic data (e.g. Interactive Maps and Other Graphic Representations of Data) students should give the name of the researching organization followed by the date. In brackets, provide a brief explanation of what type of data is there and in what form it appears. Finally, they should provide the project name and retrieval information.

Example: CEPAL Infographic.

Latin America and the Caribbean and Millennium Development Goal 1. (2015). Eradicate Extreme Poverty and Hunger. Economic Commission for Latin America and the Caribbean. Retrieved from http://www.cepal.org/sites/default/files/infographic/files/15-00878_odm_fichas_web_english-1.pdf

OUTREACH

Providing data reference or statistical literacy instruction to students can also lead to cross-disciplinary outreach opportunities at your academic library. For example, to celebrate the UN's World Statistics Day in October 2015, University Libraries hosted an informal data visualization event for students and faculty. We had representation from Political Science, Diplomacy, and the Psychology Departments. The faculty came with laptops and showed relevant data visualization examples to (mostly) graduate students. Students rotated between stations to see the visualizations across disciplines. One graduate student said that the event made data seem "less scary" which is certainly a goal with data and statistical literacy. World Statistics Day happens every five years, so start planning for 2020.

"R" is an open source statistical programming language that can be used for data visualization. There may be interest across departments on your campus for new tools such as "R."

DATA VISUALIZATION TOOLS

In addition to raw data, many agencies offer tools that allow students to create maps, timelines or charts. One example of the UN's data visualization tools are at Comtrade's Visualize Data (http://comtrade.un.org/labs/). http://devdata.worldbank.org/DataVisualizer/ OECD's Data Lab (http://www.oecd.org/statistics/datalab/) features 21 different tools including AidFlows, which visualizes how much development aid is provided and received around the

world. Users can select individual donors (providing the aid) and beneficiaries (receiving the aid) to track the sources and uses of aid funding.

Another data visualization example is United Nation's Educational, Scientific and Cultural Organization's (UNESCO) eAtlas for Education 2030. Available at http://tellmaps.com/sdg4/#!/topic/TARGETS, this group of interactive maps allows students to see underlying data for SDG4 – Quality Education. Finally, World Bank's DataBank is an analysis and visualization tool that contains collections of time series data available at http://databank.worldbank.org/data/home.aspx

CONCLUSION

Data and statistical literacy can assist students and faculty to enhance their research and scholarship including new visualization techniques. The sources described in this article can be used as a springboard to find and extract data for analysis. Data needs to be analyzed to support a thesis or a research assignment, whether this is done in Excel, SPSS or Stata. In addition to finding the proper data, students can display the data in many new ways, using data visualization tools provided by the UN and NGO's. There are also many open source tools that can be used, such as with "R" programming (library.shu.edu/learnr). To stay current, use Twitter feeds provided by United Nations agencies and NGO's and the work of UN Libraries such as @SustDev, @SDG2030, @UNdata, @UNStats, @UNLibrary. These feeds will alert you to when updates are available.

CITATIONS

Carlson J, & Johnston L 2015, *Data Information Literacy: Librarians, Data, And The Education Of A New Generation Of Researchers*, n.p.: West Lafayette, Indiana: Purdue University Press, Seton Hall University, EBSCO*host*, viewed 18 August 2016.

Carlson, J and Stowell Bracke, M, 2015, Planting the seeds for data literacy: lessons learned from a student-centered education program. *International Journal of Digital Curation*, *10*(1), pp. 95-110.

Carter J, Noble S, Russell, A and Swanson E 2011, Developing statistical literacy using real-world data: investigating socioeconomic secondary data resources used in research and teaching. *International Journal of Research & Method in Education*, *34*(3), pp. 223-240.

DeLuca L, 2016. Email to Martin Edwards and Nabeela Alam, July 18.

DeLuca L, 2015. Making Peace with United Nations Data: Learn to navigate UN data information sources. Webinar. North Carolina Library Association Government Section. Available at: https://www.youtube.com/watch?v=kPzpUSOs_U4 [Accessed 5 Aug. 2016].

DeLuca L, Elichko S, Morse C, 2015, Data and Statistical Literacy E-Forum. Law and Political Science Section (LPSS) Program. Available at: http://www.ala.org/acrl/aboutacrl/directoryofleadership/sections/lpss/lpsswebsite/conferences/conferences [Accessed 5 Aug. 2016].

Maybee, C and Zilinski L, 2015, Data informed learning: A next phase data literacy framework for higher education. *Proceedings of the Association for Information Science and Technology*, *52*(1), pp.1-4.

New Media Consortium, 2016, *NMC Horizon Report: 2016 Higher Education Edition*. Available at: http://cdn.nmc.org/media/2016-nmc-horizon-report-he-EN.pdf [Accessed 5 Aug. 2016].

Purdue Online Writing Lab, Reference List: Electronic Sources (Web Publications): Data Sets, Graphic Data. Available at: https://owl.english.purdue.edu/owl/resource/560/10/ [Accessed 5 Aug. 2016].

Schield M, 2004, Information literacy, statistical literacy and data literacy. In *IASSIST Quarterly (IQ)*.

Tenopir C, Hughes D, Allard S, Frame M, Birch B, Baird L, Sandusky R, Langseth M and Lundeen A, 2015, Research Data Services in Academic Libraries: Data Intensive Roles for the Future?. *Journal of eScience Librarianship*, *4*(2), pp. 4.

Thornton, S, 2012, Trying to learn (politics) in a data-drenched society: can information literacy save us?. *european political science*, *11*(2), pp. 213-223.

APPENDIX: REFERENCE CHECKLIST – UN DATA INQUIRIES

Is the patron looking for:

- Agency
- Initiative – Sustainable Development Goals, Zero Hunger Challenge
- Indicator – which agency publishes this indicator?
- Topic – Population density, trade – allow you to review key concepts related to agency or initiative

Combine your search terms with some of these terms:

Data, dataset, repository, digital library, datasets, data visualization, digital library, infographic

Does the data exist for the time period the patron is looking for?

Does the data exist for the region/country/initiative the patron is looking for?

Can they shift their research question to data that is available?

What other experts can the patron consult to verify other data sources? (Liaison librarian, ask DAG, data administrator at specific agency, professional listservs such as LPSS or INT, IASST)

Does the patron need assistance with Excel, Stata or SPSS? Where on your campus can the student get assistance with this?