UNITED NATIONS ECONOMIC COMMISSION FOR EUROPE
CONFERENCE OF EUROPEAN STATISTICIANS

UNECE Work Session on Gender Statistics
(Geneva, Switzerland, 26-28 April 2010)

Working paper 7
14 April 2010

Session III B. of the provisional agenda

USING INNOVATIVE TOOLS AND MEASURES TO DISSEMINATE
GENDER STATISTICS

Note by OECD[1]

Invited paper

I.   INTRODUCTION

1.   Statistical offices can use innovations in web 2.0 technology and the new field of data science to join the growing debate on how to measure progress in societies. At the OECD Development Centre (OECD-DEV) gender statistics have been on the cutting edge of these developments. Web 2.0 technologies have revolutionized the manner in which statistical agencies produce, analyse, and share statistical data. Sources of freely accessible data sets are proliferating in the public sphere—creating the need for flexible analytical tools for parsing and visualising data. Collaborative platforms like Wikigender and Wikiprogress have broadened the community that creates and uses statistics. This document demonstrates why and how OECD-DEV is using innovative web 2.0 solutions to create, use, and disseminate its gender statistics to an increasingly wide audience.

2.   Statistical offices are well poised to take part in a growing international debate on measuring progress in societies by embracing new innovations related to data science and web 2.0 technologies. At the OECD Development Centre, gender statistics are on the cutting edge of these developments. This document reflects on how gender statistics at the OECD Development Centre are trying to help redefine how societies measure progress, how web 2.0 information and communication technology (ICT) tools are changing the way these statistics are used to create knowledge, and how this knowledge is in turn shared and disseminated to external audiences.

---

[1] Prepared by Christopher Garroway, OECD Development Centre

II.   WHY INNOVATIVE TOOLS AND MEASURES ARE IMPORTANT

3.   Gender statistics are increasingly important to the policy-making community particularly in light of a growing focus on non-income related measures of wellbeing used to measure progress in societies. Increased demands for an evidence-based approach to policy have also highlighted the need for statistical measures to address the root cause of social problems, rather than simply the outcomes that result from them. At the same time, the increased complexity of debates on public policy and the proliferation of quantitative evidence on all sides of an issue have left many non-specialist observers uncertain about what data they can trust, and have increased demands that data production be as transparent as possible.

4.   Web 2.0 ICT tools overcome many of the challenges associated with demands for better statistical measures and a more transparent and responsive statistically-informed decision making process. Wikipedia.org describes Web 2.0 technologies as *"web applications that facilitate interactive information sharing, interoperability, user-centred design, and collaboration on the world-wide web"*. Importantly, web 2.0 technologies exhibit both a "social" aspect and a "data-intensive" aspect that can help decision makers and citizens work together to gain better knowledge of society using statistical data.

A.   Evidence Based Policy Making and Measuring Progress in Societies

5.   The recent economic crisis of 2008-09 underscores efforts underway since the 1980s to move away from relying solely on national accounts to measure national progress. While viewing progress and development in terms of improvements in a nation's productive capacity made sense in the post-war economic climate; today, for many middle and high-income countries, it is increasingly important to view economic growth as merely a supporting pillar of human wellbeing rather than its sole objective (Giovanni, Hall, Morrone, and Ranuzi, 2009).

6.   In 2008, French President Nicolas Sarkozy set-up the Stiglitz-Sen-Fitoussi commission to study alternatives to existing national accounts-related statistical concepts, like GDP, for measuring society's well-being and its sustainability (Stiglitz, Sen, Fitoussi, 2009). The Commission report, which appeared in late 2009, encourages the development and use of new sets of indicators to inform evidence-based decision making across society as a whole.

7.   In many ways, efforts to improve statistical measures of progress are a natural outgrowth of the tendency over the past decade of many public organizations to increasingly focus on evidence-based approaches to policy. Evidence-based policy making emerged from a government modernization and reform movement within the UK government, which sought to address criticisms that government policies respond only to short-term concerns and do not seek to solve long-term, systemic problems. A 1999 UK Government white paper laying out the approach says the following:

> People are becoming more demanding… as citizens or as businesses affected by the policies and services which government provides. To meet these demands, *government must be willing*

> *constantly to re-evaluate what it is doing so as to produce policies* that really deal with problems; that are forward-looking and shaped by the evidence rather than a response to short-term pressures; *that tackle causes not symptoms;* that are measured by results rather than activity; *that are flexible and innovative rather than closed and bureaucratic;* and that promote compliance rather than avoidance or fraud. To meet people's rising expectations, *policy making must also be a process of continuous learning and improvement.*

<div align="center">"Modernising Government" (1999) with emphasis added</div>

B.   Using Web 2.0 Technologies to Transform Statistics into Knowledge or Decisions

8.   Over the past decade, the increased importance of evidence-based approaches to policy making and the expansion of the free and open internet using web 2.0 technologies have raised both the demand and supply of data produced and shared by statistical offices and agencies. This "data deluge" is remaking the statistical profession and is reshaping the skills and approaches used by professional statisticians. Importantly, this change has in many cases made the role of statisticians more central and visible to decision making processes and policy debates. As Hal Varian, author of one of the principal graduate microeconomics textbooks and chief economist at Google, puts it: *"I keep saying the sexy job in the next ten years will be statisticians. People think I am joking, but who would've guessed that computer engineers would've been the sexy job of the 1990s?"* (Varian, 2009) .While statisticians continue to rely on and perfect traditional techniques such as statistical analysis and data management, they increasingly employ skills such as storytelling & visualization to solve problems, often with the help of a growing toolkit of new programming languages and software frameworks (see the special report in the 25 Feb 2010 issue of the Economist).

9.   The ability to visualize and communicate data is becoming more important to statisticians, as they move beyond traditional roles as gatekeepers of data to newer roles as "data scientists" who identify, analyze, and contextualize important pieces of data. Even with good data and rigorous statistical techniques, if the results of an analysis are poorly visualized, they often won't convince. One of the key pioneers of data visualization, Ben Fry, formerly of the MIT Media Lab has proposed the following key steps for effective data visualization (see Fry, 2008):

1) acquire and obtain data
2) parse and structure data to create meaning
3) filter out all but the data of interest
4) mine the data to discern patterns or place it in context
5) represent the data with a basic visual model
6) refine the representation to make it more appealing and engaging
7) allow users to interact, manipulate and control what is visible.

10.   Steps 1-2 correspond roughly to the traditional role played by statistical agencies and offices. Steps 3-5 reflect analysis tasks that until recently often took place far removed from the statistical production cycle. Steps 6-7 clearly are communication-related tasks that also often took place in isolation from statistical production and analysis. Today, all these tasks are taking place in a much more tightly integrated process and timeframe, accelerated by technological tools.

11.   In addition to more tightly integrating the timeframe in which statistical production, analysis and dissemination processes take place, web 2.0 technologies include collaborative technologies such as wikis and other open-source models that have widened the community participating in data production, analysis, and dissemination. New web 2.0 "social" tools effectively lower the costs of co-ordinating group action. One defining characteristic that distinguishes web 2.0 technologies from earlier technological innovations is their potential to use the decentralized network structure of the internet to harness collective intelligence and effectively employ the wisdom of crowds to solve problems. Across more and more domains, peer-produced user-generated content is competing with traditional one-to-one communication and one-to-many broadcast channels (Shirky, 2008).

12.   Not only does the free and open internet provide a collection and distribution network for data, but it also provides an extensive infrastructure for peer-production of data and analysis along the open-source model. One of the key ways that statistical offices can work with the wider community as a whole is through peer-production using collaborative online tools, like a wiki—popularised by the well-known website Wikipedia.Org, but available for anyone to use through open source software, like MediaWiki. For more information, please visit www.wikimedia.org.

13.   Traditional models of statistical data production and analysis required content producers to first filter good results from the mediocre and then publish—under an open-source wiki model however, results can be published first and subject to filtering (often socially, through so-called crowd-sourcing) after the fact. Much in the same way that web 2.0 technologies have the potential to more tightly integrate and shorten the statistical production process, they also require statistical offices and agencies to realize that the production of statistical data and analysis can remain largely a work-in-progress. For example, on a wiki, anyone can update content at any time. This may imply a certain loss of control by content producers, but the economies of scale achieved by the collective knowledge creation can largely outweigh this (O'Reilly, 2005).

III.   HOW OECD USES INNOVATIVE MEASURES AND TOOLS TO DISSEMINATE GENDER STATISTICS

14.   Typically, gender statistics focus on outcome variables. The construction of the OECD Development Centre's Social Institutions and Gender Index tries to respond to demands for policy-relevant data and innovative measures of progress by exploring the root causes of gender discrimination in societies around the world. The dissemination of the SIGI through the use of data visualization tools is an excellent example of how web 2.0 technologies can facilitate and add value to the dissemination of statistical indicators. Wiki platforms, like WikiGender and WikiProgress are also likely to become excellent avenues for disseminating gender-related statistical data and analysis thanks to new innovations in on-line database technologies, like WikiProgress.Stat.

A.   The Social Institutions and Gender Index as an Innovative Measure of Progress

15.   The OECD Development Centre's Social Institutions and Gender Index, a new composite measure of social institutions that determine gender equality, attempts to address one of the key

challenges of evidence-based policy making by tackling the causes of gender disparity, rather than its effects. The SIGI moves beyond existing gender-related indices, which focus on outcomes such as female life expectancy or access to primary education, and introduces a new measure - social institutions - as the key determinant for women's involvement in economic and social life. The index complements and improves existing measures in several ways. While conventional indicators of gender equality capture the consequences of gender inequality, the SIGI focuses on the root causes behind these inequalities. SIGI draws on 12 institutional variables relating to household behaviour and social norms, including such practices as child marriage and genital mutilation but also factors like freedom of movement and restrictions on inheritance and access to property rights.

Figure 1: Build Your Own Gender Index

Source: http://my.genderindex.org

B.   Exploring Data with My.GenderIndex.Org

16.   While the SIGI index provides a convenient benchmark that aggregates discriminatory practices into a single composite measure, it can be difficult for non-specialists to fully understand how the SIGI index can be decomposed into its constituent sub-indices. SIGI findings are being disseminated on-line through a website using the "Build Your Own Gender Index" data visualization tool, which allows users to explore various alternative forms that the composite indicator can take depending on how users weight and combine the underlying data.

17.   The My.GenderIndex.org data visualization framework was designed to help make the statistics underlying the SIGI relevant and meaningful to non-expert users. The technology underlying My.GenderIndex.Org originated as part of the Cityrank.ch project at the Ecole Federale Polytechnique de Lausanne Media and Design lab (Flaxman, Huang, Stephenson, Comtesse, 2009). The main of both web applications idea is to provide flexible structure to a body of data that otherwise may be neither extensive nor well organized. Like cityrank.ch, my.genderindex.org provides a flexible framework for organizing, understanding, and presenting data. To make the data more relevant to users, users create and share personalized rankings based on the original OECD SIGI data sets.

18.   The "Build Your Own Gender Index" tool on My.GenderIndex.Org that allows users to explore the social institutions data in a transparent manner, dropping and adding variables, and changing the weights which individual social institutions carry in the composite measure. It is also useful tool for comparing how countries rank according to various individual social institutions or according to combinations thereof. The tool also allows users to explore the data at a regional level, as well. This type of data exploration can be very illustrative in its ability to help users understand how variation in the SIGI scores breaks down into its constituent social institution variables, and across geographical space. For more information, please visit http://my.genderindex.org

## C.   Visual Storytelling with OECD eXplorer

19.   Data visualizations can generally be categorized as either exploratory or communicative (Driscoll, 2009). Exploratory data visualization tools are generally intended to help the data analyst understand the raw data and help him or her develop a hypothesis about the data. In this case, the users of the visualization tool will be typically few in number and will very likely already be interested and or knowledgeable about the subject matter. To a certain degree, the My.GenderIndex.Org visualization framework fits into this category, since at the very least it is aimed at people interested in exploring the root causes of gender equality.

20.   In contrast, communicative data visualizations intend to communicate to a wider audience; the goal is to visually advocate for a hypothesis that the analyst has already developed on his or her own. A data visualization tool falling into this latter category, which has gained increasingly wider usage at the OECD is the eXplorer tool– an interactive graphical tool to analyze and explore OECD statistics, as whole. This cutting edge tool permits exploring, analyzing and communicating regional statistics, using a very interactive interface that combines maps with other visual presentations for selecting the groups of regions of interest to the user. Currently on-line, the tool presents the full set of OECD regional statistics updated to 2007 and 2008 with interactive maps and offers the possibility to explore trends over time (time animation). Users can view the structure of regional economies through scatter plots, parallel axes, and piecharts to compare regional and national data. Integration with a Google map layer was added recently, allowing for better identifying geographic location.

21.   One of eXplorer's most important and innovative functions is its ability to present stories about statistics embedded in an interactive visualization. This also includes the possibility to load and explore the user's own data together with preloaded OECD indicators, write your own stories and share them with others. The development of OECD eXplorer is the result of a partnership between OECD and the National Centre for Visual Analytics (NCVA, http://ncva.itn.liu.se/) at Linköping University, Sweden. Since October 2008, NCVA and OECD have worked intensively on the eXplorer tool improving it and enhancing it to meet demands of data users.

22.   Recent developments of an OECD eXplorer wizard has now made the development of domain specific eXplorer tools much easier. The OECD Development Centre is currently working on a Gender eXplorer tool to explore the statistics available through the Gender Institutions and Development Database (including the SIGI data.) The built-in story functionality

will enable the data visualization tool to provide tailored introductions to wide variety of interesting topics highlighted by the GID-DB. See Figure 2 for an example.

Figure 2: OECD eXplorer

Source: http://www.oecd.org/gov/regional/statisticsindicators/explorer
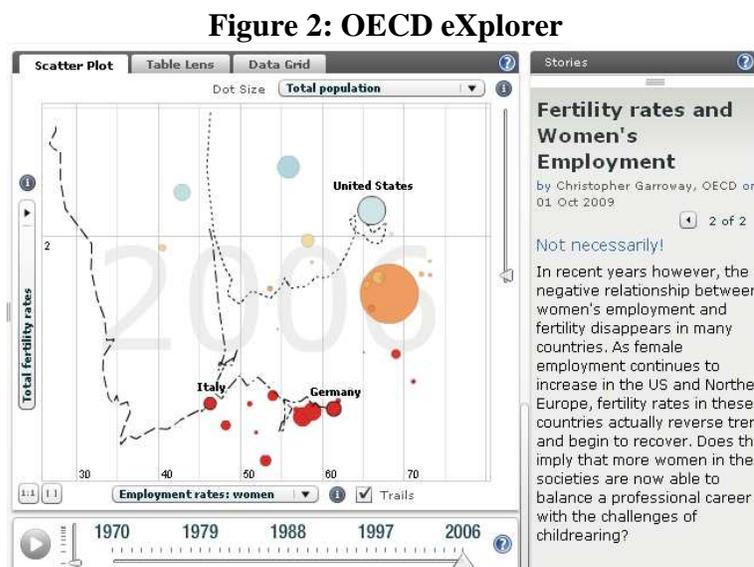
D.   Harnessing the Collective Knowledge of the Participatory Web with Wikis

23.   WikiGender was launched by the OECD Development Centre in 2008, and following its initial success other affiliated wikis like WikiProgress have emerged as initiatives supported by the OECD. WikiGender and WikiProgress together have burgeoning online communities composed of thousands of users and contributors.

24.   As WikiGender and WikiProgress expand, new functionality for loading and manipulating statistical data promises to make the peer-production model increasingly relevant not only to the qualitative content of the wikis, but to their quantitative content as well. External datasets can be loaded directly into the wikis through an interface similar to the OECD.stat online database, called wikiprogress.stat. Plug-ins have been developed to make this data embeddable in wiki entries and to also link this data to data visualizations using the eXplorer tool.

25.   Wikis are an excellent medium for more regular interaction between producers of data and the users of data. While costs of user-generated statistical content remain relatively high, compared to other media like journalism, and while accuracy in data should not be compromised for the sake of timeliness, the economies of scale and the broad appeal made possible by wiki platforms are hard to ignore. Although much data production clearly cannot be subject to the peer-production model, fruitful data analysis and storytelling can take place on platforms like WikiGender and WikiProgress.

Figure 3: WikiProgress.Stat

Source: http://wikiprogress.org/stats/

IV.CONCLUSION

26.   Innovative measures like the SIGI and innovative tools like http://my.genderindex.org ,
OECD eXplorer, http://wikigender.org, and http://wikiprogress.org have enormous potential to
enlarge the community using gender statistics and broaden their policy relevance and use to
measure progress. It is important to remember, however, that these innovations remain very new
and still can be improved through the help of external stakeholders and partners. As the OECD
Development Centre continues its participation in the progress debate and moves forward in its
adoption of web 2.0 technologies, it is increasingly important to compare and benchmark with
other organizations and offices in order to learn from the experiences of others and to stay abreast
of advancements made elsewhere.

27.   Web 2.0 technologies are bridging the gap between producers of gender statistics and
users. Data visualization and storytelling are widening the appeal and use of gender data in policy
and decision making worldwide. Collaborative platforms like wikis have broadened the
community using gender statistics and working to develop useful indicators of progress.
Statistical offices and agencies need to embrace these new tools and approaches in order to
advance the objectives of gender equality, and convince society as a whole that ending
discrimination against women will improve everyone's well-being.

References

The Economist (2010) "Data, Data Everywhere: A Special Report on Managing Information." February 25, 2010.

Driscoll, Michael. (2009) "The three sexy skills of data geeks" May 27, 2009
http://dataspora.com/blog/sexy-data-geeks/

Flaxman, S., J. Huang, J. Stephenson, and X. Comtesse. (2009). "CityRank: a dynamic tool for exploring and generating new indices of cities." Information Design Journal. 17:3, January 2010.

Fry, Ben. (2008) Visualizing Data. O'Reilly Media, Inc: New Sebastopol, CA.

Giovannini, Enrico, Jon Hall, Adolfo Morrone, Giulia Ranuzzi. (2009) "A Framework for Measuring the Progress of Societies" OECD Draft Working Paper, September 2009.

MediaWiki, www.mediawiki.org

O'Reilly, Tim. (2005) "What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software," http://oreilly.com/lpt/a/6228

Shirky, Clay. (2008) Here Comes Everybody: The Power of Organizing without Organizations. Penguin Press: New York, NY.

UK Government. (1999) "Modernising Government", March 1999.
http://archive.cabinetoffice.gov.uk/moderngov/download/modgov.pdf

Varian, Hal. (2009) "On How the Web Challenges Managers," McKinsley Quarterly, January 2009.
www.mckinseyquarterly.com/Hal_Varian_on_how_the_Web_challenges_managers_2286

Stiglitz, J.E., Sen, A., Fitoussi, J.-P. (2009) Report by the Commission on the Measurement of Economic Performance and Social Progress.
www.stiglitz-sen-fitoussi.fr/documents/rapport_anglais.pdf

Wikipedia http://en.wikipedia.org/wiki/Web_2.0