SEVENTH INTERNATIONAL CONFERENCE ON AGRICULTURAL STATISTICS

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS

Modernization of Agricultural Statistics in Support of the Sustainable Development Agenda
(Rome, Italy 26-28 October 2016)

Parallel Session – Cross Thematic set H:
Data dissemination & communication / Use of statistics for policy making & research
[48] The new professional profile of Agricultural Statisticians

Information Technology Skills in the Profile of Modern Agricultural Statisticians

Joseph L. Parsons
United States Department of Agriculture
National Agricultural Statistics Service
1400 Independence Ave, SW
Washington, DC, USA
joe.parsons@nass.usda.gov

DOI: 10.1481/icasVII.2016.h48c

Abstract

Information technology skills are needed at each stage of official agricultural statistics production; including frame development, data collection, survey cleaning, analysis, summarization and disclosure, and publication. This topic will address the skills required tocomplete each phase of the statistical data collectionprocess and how statisticians can effectively partner with information technology specialists within the Ministries of Agriculture and National Statistical Organizations.

This paper will provide background on the evolution of information technology (IT) at NASS as it relates to the production of official agricultural statistics. It will also discuss the skills statisticians need and ways in which those skills can be developed.

Keywords: Skills, Information Technology, Data Collection, United States Department of Agriculture, USDA, National Agricultural Statistics Service, NASS

1. Introduction

The National Agricultural Statistics Service (NASS) is the principle agricultural statistics data collection organization within the United States Department of Agriculture. The agency conducts hundreds of surveys every year and prepares reports covering virtually every aspect of agriculture in the United States. NASS's mission is to provide timely, accurate, and useful statistics in service to U.S. agriculture.

Beginning in 2009, NASS started a journey to pursue three transformations, primarily to provide savings in staff resource costs and improve the quality of our statistical products by enabling the centralization or regionalization of survey operations. The transformational initiatives involved: (1) centralizing and consolidating network services from 48 locations (Parsons and Gleaton, 2011); (2) standardizing survey metadata and integrating survey data into easily accessible databases across all surveys; and (3) consolidating and generalizing survey applications for the agency's diverse survey program (Nealon, 2013). During this same time period, NASS reorganized its headquarters operations, opened a national operating center and consolidated the vast majority of its field staff into twelve regional offices. About one-half of agency statisticians and technology professionals are located in the field with the remainder located at a national operations center or a headquarter location. These changes all aspects of agency operations including the technology skills required of much of our workforce.

Most NASS hires are agricultural statisticians, mathematical statisticians or are in the information technology job series. The agricultural/survey statisticians tend to have a background in agriculture and economics with the requirement of having completed several math and statistics classes. Many of theagricultural/survey statisticians we hire today have a masters degree and a few have doctoral degrees. Our mathematical statisticians have at least a bachelors degree, typically in mathematics or statistics.Most have a masters degree and we hire several with doctoral degrees. Those with specialized degrees are concentrated in our research and methodology teams. Our information technology specialists are often hired mid-career and have varied backgrounds depending upon their specialization (e.g. database administrator, software developer).

In this paper we will discuss the need for technical skills and the development of those skills for statisticians collecting and disseminating official statistics on items related to agriculture. We will share the experiences at the National Agricultural Statistics Service and relate how those experiences might inform others with similar opportunities and challenges. We will focus on agricultural/survey statisticians in the first years of their professional career, but will offer some limited observations on the development of other key groups of employees.

2. Developing Technical Skills of Statisticians

NASS and other similar statistical organizations are 'data factories'. The 'tools' in our factories are for the most part, information technology applications and staff need theskills to leverage, develop and enhance those tools. The layout of the factory floor of those generating official statistics will be pretty familiar across different organizations. The Generic Statistical Business Process Model (GSBPM) components describe and define the set of business processes needed to produce official statistics (Appendix A). The major business processes include: specifying needs, design, build, collect, process, analyze, disseminate, archive and evaluate.

Each of these major business functions have information technology tools that underpin the statistical processing activity that must be completed to produce relevant official statistics. For example, NASS uses Blaise software for data collection in its call centers and for interactive editing of some of the surveys conducted. Statisticians located in regional field offices work in Blaise frequently editing data and reviewing records. A smaller group of statisticians, working with survey methodologists, design Blaise data collection and editing instruments. These statistician/designers interact with those information technology professionals that support the application, the infrastructure, centralized database and metadata on which the application depends. Those statisticians providing this backend support must interact and collaborate with others that support the other enterprise tools so that the entire 'factory' runs effectively.

Over time NASS has created protocols to develop and train newer agricultural statisticians. For the cohort of newer agricultural statisticians involved primarily in data collection and estimation working in our regional offices, the training efforts are relatively standardized. Additionally, every employee in NASS is required to have an individual development plan (IDP). An IDP is a tool to assist employees in career and personal development. Its primary purpose is to help employees reach short and long-term career goals, as well as improve current job performance (US Office of Personnel Management). For newer agricultural statisticians several training events and career experiences are included in their plan along with specialized items that address their particular career interests or job responsibilities. NASS has also developed career guides to assist employees in managing their career development (NASS, 2016).

Each year our training and career development group conducts a training needs assessment. Typically, the assessment is sent to first and second level supervisors to prioritize training needs that are widely recognized. Managers are informed by the gaps they observe in the staff they supervise and in the IDP consultations they conduct with their staff. In addition, statisticians receive a version of the assessment so they can take a personal assessment of their skills in these essential areas.

The assessment helps to prioritize the areas on which to train and further focus how the training is specifically delivered. The following is a snippet of a recent assessment to help drive the annual formal training agenda. As shown, many of the areas are a blend of concepts and use of the actual tools.

Figure 1: Skills Assessment Grid

	What percent of your unit's ag and math stats:				
	Do not regularly use this skill in their current position	Are NOT AT ALL Proficient and need significant additional training or experience	Are SOMEWHAT Proficient but need some additional training or experience	Are COMPLETELY Proficient and do not need any additional training or experience	Total
cSMS Functions--Set DCMS codes and identify survey modes (Paper, CAPI, CAWI, CATI)					100%
cSMS Functions--Use Events tool to set data collection flags based on DCMS codes					100%
Blaise Processes--Identify records of concern					100%
Blaise Processes--Conduct forms editing - identify and interpret warnings, and resolve data inconsistency					100%

Most of our newer statisticians work in team settings and are paired with one or more experienced statisticians. A great deal of informal on-the-job training takes place in this setting. Our newer statisticians generally bring good generalized technical skills such that they are very comfortable with typical information technology office tools. However, our 'factory' has many specialized business processes and information technology tools to support those processes. Our experienced team leads are a key resource in developing the skills of our newer staff and helping them understand the survey concepts and the specific applications and tools used within the agency.

We blend this on-the-job training with more formal training sessions developed from input received on the training assessment and constrained by budgetary resources. Formal training sessions are usually a business week in length and are typically held at our training facility located within our national operations center. These sessions typically include thirty to fifty participants. Most of our training is focused on a particular survey or survey program or task.We also have a series of courses that focus on survey and estimation concepts. In every training case, staff bring their laptops, connect wirelessly to our network and virtual desktop environment to work with applications and tools that are integral to the training.

Our distributed operating environment is enhanced by being able to conduct hybrid learning environments. For example, we frequently are able to include additional instructors and technical expertsvia video teleconference (VTC). This allows us to tap a larger set of skilled trainers for interactive training sessions. In addition to the week-long sessions, we conduct frequent VTC sessions of one or two hours. Some focus on survey management issues, but others – through screen sharing – help to educate a distributed workforce on new or updated information technology tools.

NASS has made strong effective efforts to move toward enterprise level applications for case management, editing, summarization, analysis and other key functions (Drennen and Parsons, 2014; Parsons and Duxbury, 2013). These enterprise tools can accommodate highly varied survey efforts. They rely on centralized relational databases and curated metadata (Nealon, 2013). NASS also moves data from the transactional databases that support these applications to

analytical databases. We train newer statisticians to leverage standardized queries and predefined analytical tools to support the generation of accurate official statistics.

However, not all answers to relevant data issues will be found in these standardized queries and predefined analytical tools. We train more experienced agricultural statisticians to leverage business intelligence tools to explore and resolve data issues not easily addressed directly from within the tools. Our data services team that includes database administrators and metadata specialists often assist in hosting sessions on effective data exploration.

Some of the applications and technology that our agricultural statisticians use are not developed from within NASS. For example, statisticians in our regional offices frequently use applications from another USDA agency to mine farm program administrative data to verify operating status and estimate for nonresponse.

Most of our regional agricultural statisticians are generalists and work across the 'factory floor'. However, we have some niche tasks in which our regional agricultural statisticians develop expertise. For example, our field data collection staff use Apple iPads for data collection, data transmission, and survey training. A corps of our regional statisticians help keep this fleet of several thousand devices operating and their users trained. We hold periodic training to support this corps of employees as tools and procedures are updated.

Many of our headquarters agricultural statisticians are former regional statisticians. Frequently, the headquarter roles have a specific focus such as frame development, gathering survey requirements, or estimation and dissemination. Their general background and experience is invaluable to be able to connect their specialization to the broader work of the organization.

Conversely, many of our mathematical statisticians are hired with skills in sampling or statistical programming and placed in specialized roles. We use temporary assignments and other experiences to expose these more specialized statisticians to other roles and tools used within the organization. Most recently our methodology team hired nearly twenty new mathematical statisticians into a variety of jobs. We have taken advantage of having an entire cohort of new specialized statisticians to hold a series of training sessions on the tools and concepts unique to the organization.

While tools and technology are imperative to the success of the modern agricultural statistician, the most important attribute is critical thinking. In fact, we must ensure that training on the "how" does not crowd out the "why".

The profile of the modern agricultural statistician and the agricultural statistical agencycontinues to evolve. The tools we use and the focus on relational databases and the ability to manipulate data continues to grow. As we move forward, new skills will become more important. For example, we are using web scraping techniques to develop specialized lists of farm operations and explore the use of the technique to measure list sampling frame coverage. Specialized statisticians and geographers at NASS leverage big data to develop geospatial data products and inform other work of the organization. Almost certainly in the future, sensor, drone and other big data sources will inform our work and drive the technology skill set of our

statisticians.

3. Conclusion

Another description of the nexus of the computer scientist, quantitative social scientist and statistician is the ‘data scientist’. Undoubtedly, NASS and other organizations that produce official government statistics will increasingly need employees that are comfortable in the space that blends computing and computer science skills and statistics. As has always been the case, it is challenging to produce accurate and relevant official statistics for agriculture, but there has never been a more demanding, exciting, or interesting time to work in the field of agricultural statistics.

REFERENCES

Clark, Cynthia Z.F., Gia F. Donnalley, Roger Tourangeau (2004) The Joint Program in Survey Methodology and its Impact on the Federal Statistical Agency Workforce, *Proceedings of the American Statistical Association*, Toronto, Canada.

Clark, Cynthia Z.F., (2013)Mentoring and Development of Government Statisticians: Experiences as a Senior Government Official and Parent, *Proceedings of International Statistical Institute*, Hong Kong.

Coutu D L. (2002) The HBR Interview -- Edgar H Schein: The Anxiety of Learning,*Harvard Business Review*, Mar 2002, Volume: 80 Issue: 3 pp.100-106 (7 pages)

Drennen, Lorna J. and Joseph L. Parsons (2015) Information Technology Development and Implementation of a Case Management Services at a Federal Statistical Agency, *Workshop on the Modernization of Statistical Production*, Geneva, Switzerland, April 15-17, 2015

Gottschall, C. Personal Communication "Details on 2016 Skills and Training Needs Assessment"

Hillman, Paul (2010) Technology Report, Issue 27, accessed January 2, 2014. http://www.macul.org/downloads/2010_speaker_handouts/cdh75_fail_20100305_175526_42.pdf

Kegan, Robert and Lisa LaskowLahey (2001) The Real Reason People Won't Change, *Harvard Business Review*, Nov. 2001, Volume: 79 Issue: 10, Reprint R0110E, pp 1-9, accessed online January 2, 2014.

Kotter, John P. (2007) Leading Change Why Transformation Efforts Fail,*Harvard Business Review*, January 2007, Volume: 85 Issue: 1, Reprint R0701J, pp 1-10 accessed online January 2, 2014.

Kotter, John P. and Leonard A. Schlesinger (2008)Choosing Strategies for Change, *Harvard Business Review*, July-August 2008, Volume: 86 Issue: 7, Reprint R0807M, pp 1-9, accessed online January 2, 2014.

McFee, Andrew (2006)Mastering the Three Worlds of Information Technology, *Harvard Business Review*, Nov. 2006, Volume: 84 Issue: 10, Reprint R0611J, pp 1-9, accessed online January 2, 2014.

National Agricultural Statistics Service, downloaded from NASS intranet July 20, 2016 "NASS Career Path Guide, Agricultural Statistician1530 Series"

Nealon, Jack and Elvera Gleaton (2013) Consolidation and Standardization of Survey Operations at a Decentralized Federal Statistical Agency, *Journal of Official Statistics* Vol. 29, No. 1, 2013, pp. 5–28, DOI: 10.2478/jos-2013-0002.

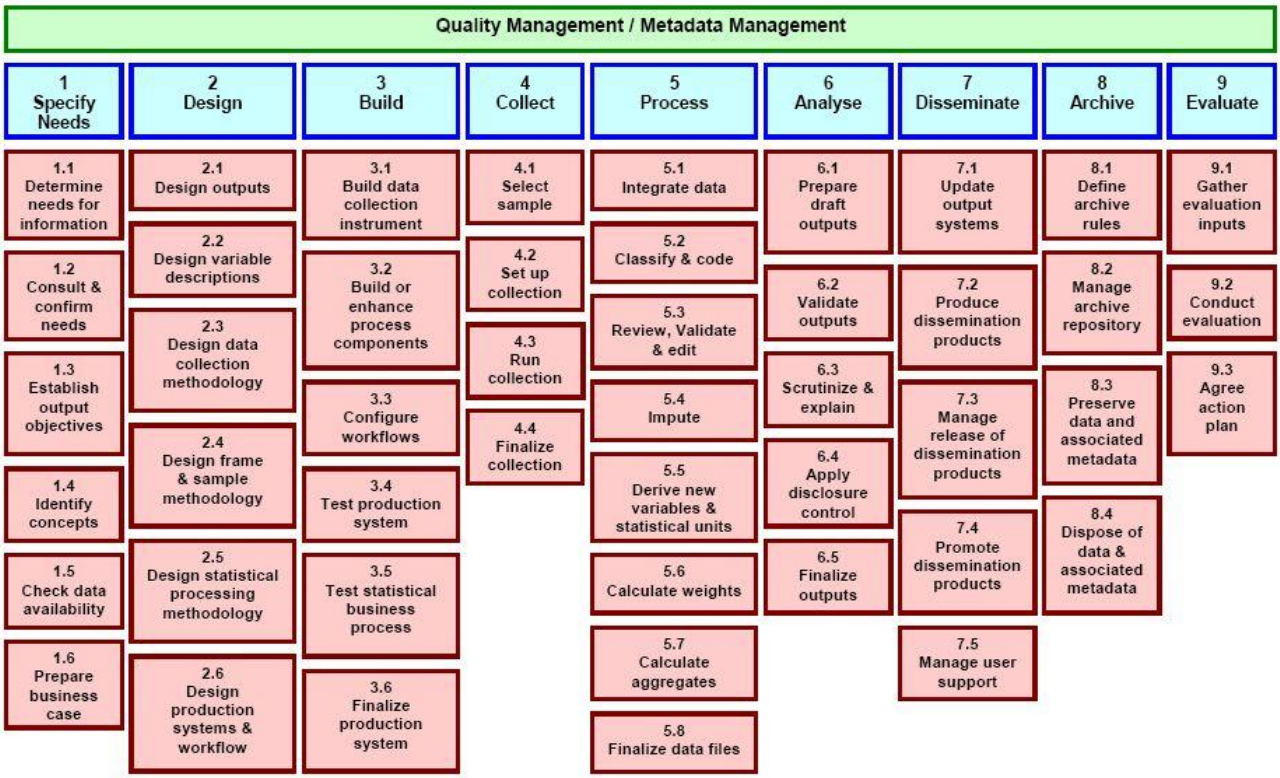
Parsons, Joseph L. and Brandon W. Duxbury (2014) Information Technology Centralization and Modernization Efforts and the Impact on Organizational Culture at a Federal Statistical Agency, *Meeting on the Management of Statistical Information Systems*, Dublin, Ireland April 14-16, 2014). Accessed February 10, 2015 http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.50/2014/Topic_1_USA_Parsons.pdf

Parsons, Joseph L. and Elvera T. Gleaton (2012) Virtualizing and Centralizing Network Infrastructure at a Decentralized Federal Statistical Agency, *Meeting on the Management of Statistical Information Systems*, Washington, DC, May 21-23, 2012, accessed January 2, 2014 http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.50/2012/06_USA.pdf.

Training and Development Policy Wikihttps://www.opm.gov/wiki/training/Individual-evelopment-Plans.ashxdownloaded July 15, 2016.

United Nations Economic Commission for Europe (UNECE), "**Generic Statistical Business Process Model**," UNECE on behalf of the international statistical community, Version 5, December 2013 http://www1.unece.org/stat/platform/display/GSBPM/GSBPM+v5.0

APPENDIX A: Generic Statistical Business Process Model (UNECE); The General Business Architecture for a Statistical Agency



http://www1.unece.org/stat/platform/display/GSBPM/GSBPM+v5.0

The GSBPM describes and defines the set of business processes needed to produce official statistics. It provides a standard framework and harmonized terminology to help statistical organizations to modernize their statistical production processes, as well as to share methods and components.