Supporting Faculty Reporting to the USDA at the University of Florida

Valrie Minson, Melody Royster

Tags: collaboration and partnerships, citation analysis, outreach, technology,
Author keywords: government reporting,

Abstract:
Each year the University of Florida (UF)'s Institute of Food and Agricultural Sciences (IFAS) reports research and extension efforts to the United States Department of Agriculture (USDA) and National Institute of Food and Agriculture (NIFA). These reports include a wide range of research accomplishments including Extension documents, refereed articles, patents, images, grants, and other non-scholarly publications. In 2008, IFAS transitioned from a home-grown software system to utilizing EndNote Web (ENW) software, available through the UF Libraries, for the management of the refereed publication reporting process. In 2009 IFAS found the level of support required to sustain the reporting process with ENW was too great and approached the Libraries with a request for additional support. Several years into this relationship, the Libraries have refined the process and improved the level of publication reporting. In addition, the Libraries now have a greater understanding of the publishing trends of their agricultural science clientele and are seen as valuable resources when publication questions arise. This paper will discuss the established agreement between the two organizations, an overview of the support process and workflow, the various challenges experienced, what the copious amounts of available data can tell us about agricultural publishing at the University of Florida, and what future possibilities this data may herald for future joint projects.

Acknowledgements: The authors thank Bart Birdsall, Katie Newman, and Michael Flowers for their assistance with data collection and their excellent work as reporting coordinators.

Background: faculty publication reporting: IFAS and the USDA
The University of Florida (UF)'s Institute of Food and Agricultural Sciences (IFAS), in order to meet federal funding requirements and to showcase the rich publication assets of a vibrant agricultural research community, required researching faculty to annually report their refereed articles. These publication lists, as well as additional research and extension accomplishments, were submitted annually to National Institute of Food and Agriculture (NIFA) under the AREERA Act – Agricultural Research, Extension, and Education Reform Act of 1998. For many years, faculty reported publications and other effort using a home-grown reporting system, but switched in 2008 to the use of the bibliographic management software EndNote Web, included in the Library’s subscription to the Thomson Reuters’ Web of Science database. Moving to EndNote Web removed the financial burden of continued software development while maintaining a comparable workflow that placed limited burden on faculty. IFAS, in the first year working with the new system, utilized an existing IFAS Information Technologist for instruction and one-on-one support to the more than 600 publishing faculty located in 28 units. After the initial year, IFAS administrators felt the required level of support was too great and approached the George A. Smathers’ Marston Science Library to assist with annual faculty publication reporting. In the fall of 2009, IFAS and the Libraries established a Memorandum of Understanding (Appendix) to establish the terms of the support and began several years of a successful and reciprocally-beneficial support relationship. The Libraries have provided the support from 2009 until this last year and have gathered approximately 1000 or more citations for each reporting year. The acquisition of a “complete” set of publication data opens the door to interesting and new data analyses that can reveal quite a bit about scholarly behavior. This paper will present an overview of the support processes and workflow, the various challenges experienced, what the copious amounts of available data can tell us about agricultural
publishing at the University of Florida, and what future possibilities this data may herald for future joint projects.

**Why the libraries?**

Through the years, IFAS and the Libraries have mutually benefited from developing strong and lasting partnerships which included library consultations on issues related to publication tracking, author disambiguation, data management, metric analyses, digitization and metadata, ethical conduct, and campus outreach. Our successful preservation of born-digital extension publications, as well as a strong library instructional program with an emphasis on bibliographic management software such as RefWorks and EndNote Web, identifies the libraries as both knowledgeable of the publication landscape and capable of providing quality instructional support for a difficult and frustrating process. Librarians’ familiarity with faculty publications, database searches, and the ENW software contributed significantly to their expanding role in the faculty reporting practices and were ideal partners for managing IFAS’ citation reporting process.

**Roles, timeline and data collection workflow**

The faculty reporting process includes a number of key roles/positions and workflows for the efficient management of the reporting process. The Reporting Coordinator is the primary support contact and assists with all aspects of the reporting process. Roles and responsibilities include:

- **IFAS Faculty:** create list of refereed publications; share list with Unit Coordinators
- **IFAS Unit Coordinators:** create the unit list of publications, de-duplicate list, and review data for accuracy; share list with Reporting Coordinator
- **Reporting Coordinator (Library):** create and update support materials; disseminate reminders and deadlines to Unit Coordinators and faculty; review unit lists for accuracy and compile the IFAS master list and statistics; document the process
- **Librarian:** supervise and train the Report Coordinator; review data quality; oversee process, as needed; review lists and statistics prior to submission to IFAS Administration
- **IFAS Administration:** communicate expectations, deadlines, and policy

The work, from first notification email sent to faculty to the final report sent to IFAS Administration, falls between November and April and is separated into five distinct workflows:

1) Support materials and timeline development,
2) Self-reporting,
3) Unit reporting,
4) Unit and master list processing,
5) Sanitization and analysis

**Workflow:**

1) **Support Materials and Timeline Development:**
In November, the Reporting Coordinator, the Librarian, and IFAS administration develop a timeline of the work, which includes identification of key times for communication. The Reporting Coordinator reviews and updates the online print guide, the video tutorials, and the support web pages. The Reporting Coordinator also creates spreadsheets in order to document the questions asked throughout the cycle and to maintain statistical counts for the final report.

2) **Self-reporting:**
In early December, faculty receive their first official notification of the upcoming deadline from IFAS administration. A follow-up email from the Reporting Coordinator with support information includes
additional details on the expectations, on important deadlines, and on requesting additional support. Faculty are given two months to complete their publication reporting. They are asked to:

- Populate a folder in EndNote Web with their peer-reviewed publications.
- Submit accepted, in-press, or submitted publications in the official publish year.
- Include the tag “refereed” in the Type of Work field.
- Include author(s), article title, source, volume, issue, year, and persistent URL when available.
- Share their folder with their unit coordinator by late January.

Faculty are encouraged to export publications using Web of Science or other databases relevant to their subject areas, although faculty often find it easier to hand-enter the publications. In 2009, IFAS administration required faculty to include retrospective publications from previous years, but in subsequent years this was optional or even discouraged. As a final step in self-reporting, when faculty have finalized their publication lists, they distribute the list to Unit Coordinators through the “share” function found in EndNote Web.

3) Unit reporting:
The first two weeks of February, Unit Coordinators compile the individual faculty folders into one unit-level master list. Unit Coordinators are asked to:

- Remove multiples of the same article (appearing when an article is co-authored by multiple faculty in the same unit)
- Review for accuracy

The Unit Coordinator then shares the unit master list with the Reporting Coordinator.

4) Unit and master list processing:
By the end of February, the Reporting Coordinator has received the unit lists and begun the heavy lifting of the project. The coordinator must:

- Remove duplicate publications from the 28 separate unit folders
- Review each unit list for accurate reporting of “refereed”

The Reporting Coordinator compiles an IFAS-All master list of publications, completes a final de-duplication process, and documents and submits statistics to IFAS administration.

5) Sanitization and analysis:
Separate from the above steps, the librarian completes a final, quick data sanitization step in order for the data to be showcased publicly by IFAS. In past years the data was displayed with little regard to how “clean” the data appeared. The following steps were included in the collection workflow and are now part of the added data analysis sanitization process:

- Export individual unit data from EndNote Web to Excel for initial sanitization (capitalization changes)
- Import Excel into Google Refine and sanitize the data (uniform source titles, additional de-duplication)

Challenges of data collection
The Libraries encountered more than a few challenges, including limitations with the tools, differences in user perceptions, and general breakdown in communication. We overcame many of the challenges through the development of altered processes, conversion to a new technology, or adjustment of our communication plan. Below are the challenges faced and the solutions employed for acquiring more accurate reporting data.
Tool limitations:
Bibliographic management software applications are excellent for bibliography creation but are not always intuitive for "occasional" users. Both faculty and unit coordinators utilize this software only for annual publication reporting and found exporting/importing and database selection difficult to navigate. Rather than learn how to export or import citations, or take time to locate the best database, many faculty chose to manually input publications. To overcome the difficulties of using bibliographic management software, the coordinator offered additional trainings, presented at unit meetings, created short video tutorials, and encouraged the use of Google Scholar as a general export option. The Reporting Coordinator also provided deadline incentives by providing export support to the few faculty and unit coordinators working in advance of the scheduled deadline.

Accurate folder sharing between researcher, unit coordinator, and coordinator accounts is important to the overall process as we compile separate publication lists; unfortunately, the software's folder sharing features, while excellent, were not designed for large-scale sharing. To reduce sharing issues that arise each time a new unit coordinator is assigned the role, and for easier troubleshooting account issues, we created and assigned EndNote Web unit accounts to each unit coordinator; despite the account creation, as well as increased communications, researchers still occasionally share with an incorrect email account.

Another limitation of bibliographic management software is visible when attempting to manage thousands of citations, particularly when the citations are shared between multiple accounts. For example, the workflow for the Reporting Coordinator included de-duplication of the unit and IFAS master lists. The de-duplication step was accomplished from within the software but without utilizing the “duplication checking” functionality, as the software does not easily allow de-duplication within a particular folder.

Lastly, glitches in the software are a frequent occurrence, particularly between the shared accounts. In one instance a faculty folder was empty but the unfiled folder listed 1174 citations. In another instance, citations visible from one account were not visible to a second shared account. These types of glitches were resolved quickly with a phone call to technical support but posed a threat to the accuracy of the data, particularly if the unit coordinator failed to identify that data was missing. These issues were incredibly frustrating to faculty and unit coordinators who spent valuable time troubleshooting. If we continue to utilize the software for faculty reporting, these types of glitches are expected to continue.

User perceptions:
On occasion faculty perceived the process as evaluative. There may be some anecdotal evaluation at the administrative level (i.e., unit X has fewer publications this year); however, evaluation of the individual or unit was not the purpose. Despite this, faculty bolster their lists by reporting beyond what is required (including “in process”, “in press”, or “submitted” publications, as well as posters, presentations, patents, etc.). The inclusion of items outside the refereed requirements means filtering for refereed publications is a crucial step. For easy filtering, faculty are required to include a “refereed” tag in citations, but frequently make reporting mistakes which include: failing to include the “refereed” tag, tagging non-refereed publications as refereed, or incorrectly tagging with data that adversely affect the search filter results (i.e., “referred” or “non-refereed”). To increase refereed reporting accuracy, the Reporting Coordinator spot-reviews publications by searching Ulrich's Web for the refereed status of a journal. To further compound the issue, many journals are listed in Ulrichs as non-refereed when, under further examination, the titles were found to be refereed. To avoid disagreements, the Reporting Coordinator will add a refereed tag to a citation, but will only remove a tag if the title is clearly an Extension document. Extension documents do undergo a type of review process, but this review process does not fit with the definition of refereed for NIFA reporting. In the future, to avoid
the time-consuming process of reviewing titles in Ulrichs, we will request that faculty submit only peer-reviewed publications.

Communication breakdowns:
Communication is a great challenge on any multi-step project involving many people and organizational units. In our first year the library sent out emails to the unit coordinators, asking they disseminate the announcements to the faculty. Unit coordinators were not always diligent in forwarding the announcements. We have since outlined a thorough schedule of key times to send out communications and reminders. These communications are sent to all people involved in the reporting process, with occasional communications coming from IFAS Administration as official reminders, and offer support incentives to units finalizing their lists early. Over time we have also shortened the time allotted to faculty to work on their lists, as support queries indicated that faculty failed to utilize the extra time.

One additional and important communication challenge faced by the libraries is our insertion in an unpopular process. It is important that the libraries continue to be thought of as a neutral support entity separate from evaluative processes. Library communications should emphasize the libraries as providing support for a difficult process, rather than as initiators of the process. When faculty email the libraries regarding their frustration with the process and interruption to their work, it is important to continue to provide gentle communications regarding our support role and refer all policy decisions to IFAS Administrators.

Data Analysis
This ongoing project provided the Libraries with several years (2009-2012) of comprehensive citation lists containing fairly accurate data. Prior to the collection of this data we had a general understanding of the research habits of faculty, but much of our knowledge was anecdotal. This data provided an opportunity to understand researchers from their publishing habits and to cater services and collections to specific needs. For example, the agricultural sciences covers many different disciplines, making it difficult to choose just one database for exporting a particular faculty member’s publications, and yet, there is a generally held belief that Web of Science has relatively complete agricultural science journal coverage. The data allows us to answer two basic questions (in which journals are faculty publishing? and which databases index the greatest proportion of those journals?) that can provide a better understanding of the databases’ value. Below are the questions we sought to answer and the methods used, to answer these questions. Included are the amount of processing time (N=none; L=limited; M=medium; H=high; V=very high) required and (***) methodology or tools utilized. None (N) indicates there is no processing beyond the previously identified data sanitization workflow. Many of these questions can be applied at both the IFAS and unit level and compared across multiple years.

- How many publications were authored? (N)
  Curation: tally publications
  Provided general benchmarking and accreditation reporting support regarding quantity of publications

- What is the faculty to publication ratio? (N)
  Curation: Current faculty counts are required
  Provided information on unit productivity. This answer helped us understand which units had the greatest number of publications per faculty member.

- What percentage of publications that are single-, double-, or triple-authored? (M)
Curation: Hand-count authorship, or utilize statistical program such as SPSS to assist
Provided a preliminary understanding of co-authorship trends for each unit. This answer helped us
assess which units published large-authorship papers, similar to the Physics Department. Single,
double, and triple authorship were organized by unit count. For local analysis, librarian reviewed
which units were multi-authored (local interest only).

- In which journals are the researchers publishing? (L)
  Curation: Extract journal counts using SPSS or Google Refine
  It is interesting to note that of 500 journal titles in 2012, 250 were only published in once by IFAS
  faculty. These counts help us understand the value of a journal to the institution and can assist with
  additional questions about journals and databases.

- What types of journals (open access or subscription)? (H)
  Curation: Search Ulrichs by journal title and document OA status
  Useful for understanding whether OA journals are used for UF research dissemination, as well as
  whether annual OA publishing fund initiatives are effectively utilized.

- Which databases index the journals? (H)
  Curation: Search Ulrichs by journal title and document indexing databases
  Useful for understanding which databases are best for exporting faculty publications.

- Does VIVO@UF or Google Scholar list the citation? (M)
  Curation: Search VIVO and Google Scholar by article title and document availability
  This search would tell us which source might be the strongest source for dynamic harvesting of
  citations. Additionally, searching VIVO reveals the effectiveness of the publication harvesting effort
  from Web of Science and presents opportunities for researching author disambiguation. Lastly, the
data can be used to talk to faculty about increasing access to publications through consistent author
  affiliations.

- What percentage of publications are co-authored with other IFAS researchers? (H)
  Curation: Acquire an IFAS employee listing and note unit affiliation (code by home or tenure unit).

Results
We were able to examine and compare some data between 2011 and 2012, as the data from these years
benefited from improved sanitization processes, but the most thorough and complete analysis was applied to
2012 data. Primarily the greatest challenges of the analyses were navigating and ordering the sanitization
steps, the hand-processing required for greatest accuracy, and planning collaboration analyses when many of
the articles were co-authored by more than 12 authors. Secondary challenges included navigating unfamiliar
tools (Google Refine, SPSS, advanced Excel) or working within the limitations of the available tools (such as
Ulrichs).

The Reporting Coordinator and Librarian spent approximately 240 hours processing the 2012 citation data,
resulting in the following statistics:

- From 2011 to 2012 there was a 9% increase in the number of publications reported.
- In 2011, 10% of publications were co-authored with other IFAS units. In 2012, this percentage
  increased to 17%.
The units with the greatest number of publications per researcher were Southwest Florida Research and Education Centers, Wildlife, Soil and Water Science, Citrus Research & Education Center, Agronomy, West Florida Research & Education Center, Indian River, and Plant Pathology.

In 2012, researchers published articles in 562 unique journals. 312 of these journals published only one article with an IFAS researcher listed as an author. The remaining 248 journals published two or more unique articles. The 10 journals with the highest frequency of articles, included: HortScience (48), HortTechnology (43), Plant Disease (28), PLoS One (28), Florida Entomologist (27), Arthropod Management Tests (23), Journal of Animal Science (18), Crop Science (16), and Phytopathology (16). Of the 248 titles, 18 are open access.

According to Ulrichs, the 248 journals with 2 or more articles were indexed with the following counts:

<table>
<thead>
<tr>
<th>Number of journals indexed</th>
<th>Database</th>
</tr>
</thead>
<tbody>
<tr>
<td>186</td>
<td>Agricola</td>
</tr>
<tr>
<td>227</td>
<td>CAB Abstracts</td>
</tr>
<tr>
<td>219</td>
<td>Scopus</td>
</tr>
<tr>
<td>213</td>
<td>Web of Science</td>
</tr>
</tbody>
</table>

The Libraries are still processing and analyzing whether citations can be discovered in Google Scholar and VIVO.

**Future directions**

Our preliminary findings are important for supporting us in assessment of future directions. This ongoing project is funded by IFAS for the primary purpose of faculty publication reporting; thus the Library’s primary goal has been to increase the reporting accuracy while reducing the difficulties encountered by faculty when self-reporting. Over the years we streamlined the process significantly and reduced the number of complaints received. EndNote Web was chosen by IFAS under a different administrative team and is not a technology we feel strongly wedded to for this work. Eventually we hope shifting from EndNote Web to a different system will reduce frustration levels. We are considering a number of alternate solutions for data collection, as the majority of faculty do not use the export publication function provide by EndNote Web. Utilizing benchmarking resources such as Academic Analytics or harvesting publications from VIVO and supplementing with a simple form may be other options to consider. We will have to be cautious if we shift to a new solution, as this workflow has the potential to be used as a solution for the recent Office of Science and Technology Policy (OSTP) mandate requiring the long-term preservation and sharing of publications funded by federal agencies.

We are also exploring ways in which the data analysis may assist with Southern Association of Colleges and Schools (SACS) accreditation review or general benchmarking support. IFAS is interested in a number of metrics for SACS reporting, as well as metrics to identify collaboration trends. The basic publication reporting provided by this work is significant; for example, one of the IFAS SACS goals is to increase scholarly productivity (as measured by refereed journal publications per faculty FTE) by 5% in 5 years. The current process will allow for tracking this type of productivity and it may be, with further exploration that our reporting can continue to expand support for SACS and benchmarking efforts. Additionally, the final NIFA report includes assignation of publications to NIFA Knowledge and Program Areas. We are also exploring methods to painlessly assign these categories to the relevant citations, for assisting with additional reporting requirements not met by our current process.
We hope to both continue the publication support and to continue to expand the analyses to meet the changing needs of IFAS administration. One area of interest is the exploration of non-IFAS and non-UF collaborations. We are also considering continued analyses on the scholarly impact of our researchers. These types of analysis are generally approached from a discipline and are time-consuming and difficult to accomplish, albeit valuable for understanding team science as applied to agriculture. Initiatives such as ORCID, as they further develop, will be helpful in understanding a researcher’s collaboration or scholarly impact.

This data will also feed the improvement of library collections and support services. The identification of key journals for IFAS provides insight into the significance of particular journals, many of which we subscribe to only in print. Acquiring these titles in electronic, despite high financial burden, is something to consider given their importance. This same approach can also be applied to databases. The Libraries have long subscribed to Web of Science, rather than Scopus, and there has been some question as to whether Scopus is the more appropriate database. The analysis has identified the primary journals in which IFAS faculty and IFAS units publish have better indexing in CAB Abstracts and Web of Science, although the question still remains as to the thoroughness of the indexing within these two databases. Library support services may also be improved. Currently the libraries are developing data support framework. This project has provided librarians with greater familiarity with spreadsheet technologies (Excel), data transformation programs (Google Refine), statistical programs (SPSS), and data visualization software (Sci2), leading to additional workshop and instructional support. Lastly, we hope to develop additional workflows that increase the number of faculty publications submitted into the libraries' institutional repository.

**Conclusion:**
This project was highly successful in many different ways. It helped the libraries build a trusting and unique relationship with the agricultural science community and provided the libraries with data that assisted with collection and service development. In the last few years, data has become an important topic within the academic community and this project provided hands-on data and technology experience. These relationships and skills are necessary for the library to continue to meet ongoing data needs, and this project, firmly integrated in the agricultural science community and with clearly-defined outcomes, has given the libraries greater prominence within the faculty and administrative community.
Appendix:

**The Memorandum of Understanding (MoU)**

The agreement was perpetual until either party cancels, in writing, with a minimum of 45 days' notice of cancellation.

Funds would be used to employ a position (Reporting Coordinator) to assist IFAS with faculty publication tracking. The general responsibilities include:

- Assist all IFAS faculty with publication reporting related questions
- Assist all IFAS faculty with EndNote-related technical questions
- Assist EndNote Unit Coordinators with creating unit publication lists
- Ensure data validity by checking all publication data for quality, completeness, and applying reporting standards, and correcting records which are not sufficiently complete or accurate
- Provide customized lists and/or reports for Research Deans on request
- Serve as IFAS-wide phone/email contact for EndNote

This position is supervised by a Library faculty member and may work in either the Marston Science Library building or remotely, as deemed appropriate. IFAS funds the purchase of a computer every 3.5 years to be used by the position and the computer is supported by the library IT department. The UF Libraries provide a suitable workspace, a telephone and connectivity. The employee delivers monthly progress reports to the Library supervisor and shares the report with IFAS.