

Faculty Research Practices in Civil and Environmental Engineering: Insights from a Qualitative Study Designed to Inform Research Support Services

Study time frame:

Study began in Fall of 2017, interviews conducted in Spring of 2018, analysis in Summer of 2018 and report written in Fall of 2018.

Authors:

Whitney Hayes, Virginia Pannabecker, Yi Shen, Erin M. Smith, Larry Thompson

Roles:

Authors are listed alphabetically by last name. Pannabecker, Shen, Smith, and Thompson contributed equally to the full study. Hayes joined later in the study and contributed equally to data analysis, synthesis, and report writing.

Abstract

This qualitative study analyzes and reports on in-depth interviews with eight faculty from the Charles E. Via, Jr. Department of Civil & Environmental Engineering at Virginia Tech. This project is part of a larger series of studies from Ithaka S+R on the research support needs of scholars by discipline. Ithaka S+R provided guidance on research methodology and data analysis. The goal of the study was to better understand the overall and day to day research practices of civil and environmental engineering (CEE) faculty to inform priorities and needs for research support services. The interviewees represented different career levels (Assistant, Associate, and full Professor) with at least one being from each of the Virginia Tech CEE department's five fields of focus. Findings address five thematic areas from the interview transcripts: research, collaboration, data practices, information discovery and management, and research communication and dissemination. Implications identify four major support needs and service opportunities: supporting faculty goals to share more of their work in open access venues, supporting use of Google Scholar as a major research tool, data education and services for all - from students to faculty, and the continuing importance of student training in information source evaluation and critical appraisal.

Introduction / Aims

This study is an in-depth, small scale, qualitative analysis of the research practices of faculty in Civil and Environmental Engineering (CEE) at Virginia Tech to better understand the resources and services these faculty members need to be successful in their research. This information will be used to articulate the research activities and needs of CEE scholars including identifying improvements to pre-existing research support services provided by University Libraries at Virginia Tech and opportunities for developing new services for CEE or more widely.

This project is part of a larger series of studies from Ithaka S+R on the research support needs of scholars by discipline. The local study reported here is connected to a suite of parallel studies being developed locally at other US-based higher education institutions with CEE departments. Ithaka S+R, a not-for-profit research and consulting service that helps academic, cultural, and publishing communities, provided guidance on research methodology and data analysis. The research project as outlined here was implemented exclusively by the Virginia Tech authors on this report. The anonymized aggregated data and analysis will also be used towards a comprehensive report covering findings from all participating institutions that will be written and made publicly available by Ithaka S+R.

Local Context

The Charles E. Via, Jr. Department of Civil & Environmental Engineering at Virginia Tech offers a comprehensive program of research and instruction. The department confers graduate degrees in five distinct fields: Construction Engineering and Management; Environmental and Water Resources Engineering; Geotechnical Engineering; Structural Engineering and Materials; and, Transportation Infrastructure and Systems Engineering.

Within the department there are currently 79 faculty members, 545 undergraduate students, and 328 graduate students. U.S. News and World Report (2017) ranks the graduate program tied at #7 with MIT, and the undergraduate program tied at #7 with Purdue.¹

Participants

This study reports on interviews with eight Virginia Tech CEE faculty members at different career levels (Assistant, Associate, and full Professor), with at least one being from each of the department's five fields of focus.

¹ *U.S. News & World Report*. "Best Civil Engineering Programs."
<https://www.usnews.com/best-graduate-schools/top-engineering-schools/civil-engineering-rankings>.
Accessed November 1, 2018.

Methods

The authors planned the Virginia Tech study based on the Ithaka S+R required methodology and protocol focused on a qualitative case study of CEE faculty research experiences and needs at Virginia Tech. The study protocol was approved by the Virginia Tech Institutional Review Board in 2018.

Interviewees were recruited from the CEE faculty of Virginia Tech. Eight researchers responded with an interest in being interviewed, and upon confirming and signing the study consent form (Appendix I) each were interviewed by the Engineering Librarian Thompson, an author on this study, using the IRB approved Interview Guide questions (Appendix II).

The interviews were recorded on two devices and recording files were transferred to a secure location as were copies of the signed participant consent forms. Transcription of the interviews was completed through Scribie, a commercial transcription service.² Personally identifying information (PII) was removed from the transcripts by the study authors. The de-identified transcripts were shared with Ithaka S+R for use in the overall compiled report of aggregated results from all participating institutions.

The study authors read and coded each transcript for key issues, comments, and themes, focusing on discussion of research and service needs for this report.

Findings / Results

Research

Scope of CEE Research

The current research areas for the interviewed faculty are broad, including transportation, structures, geotechnical, environmental / water resources and construction. Within these broader areas researchers are specifically focused on: coastal engineering, including coastal erosion and instability, coastline failures during extreme events, coastal hazards, and nature based infrastructure for hazard mitigation; evaluating, improving and extending the life of aging infrastructure, primarily steel bridges; earthquake engineering for steel buildings; steel beam-to-column connections with an emphasis upon earthquakes; resilient water infrastructure systems; transportation safety; lake and reservoir management; indoor environments; and, systems approaches to integration of knowledge domains.

² Scribie transcription service website. <https://scribie.com/>. Accessed November 1, 2018.

Stakeholders

There are three primary stakeholders involved with research in the field of CEE, with the first being funding agencies, which finance the facilities necessary for more expensive research projects. This funding is received primarily from government entities, (e.g. National Science Foundation, Federal Highway Administration, Environmental Protection Agency) and trade/manufacturing associations (e.g. American Institute of Steel Construction, Association of American Railroads). The second group of stakeholders are the researchers themselves. These are primarily the CEE faculty, but they also collaborate with faculty outside the CEE field, as well as comprise teams which include graduate students, technicians, government officials, industry experts, and others.

The third broad group of stakeholders impacted by civil and environmental engineering research includes engineers and practitioners, technical organizations, government agencies, and members of the public. Civil engineers and practitioners working with the built environment are impacted by and sometimes engage with researchers to contribute to new models, standards, instrumentation, and techniques for CEE research. Technical organizations rely on current research to issue standards that provide details for building design requirements. Government agencies may use the research to implement policies for such things as coastal development or erosion control. And, ultimately, because civil engineering works with the built environment, and because everyone is affected by that environment, every individual is a stakeholder in civil engineering research.

Methodologies Used

The strategies for conducting research vary greatly depending upon the research environment and approach. While civil engineering, in all its areas, deals with the constructed environment, the research environment runs the gamut from office to field to lab. Therefore, the approach to research is as varied as the researchers. One of the interviewees referred to themselves as a data gatherer, another as an experimentalist, and still another as a field worker. All of the researchers view their work as being clearly related to other disciplines, not only within civil engineering, but also to other science and engineering disciplines, as well as to the communities their research impacts.

Regarding specific research methods, modeling is one that is common to several researchers. One interviewee describes their process as, "I like to have models and measurements. So, I collect data, develop models, put the two together, and that's my general approach." To this researcher, data collection involves generating their own data, as well as using data from others. Computational approaches mostly relate to data analysis, modeling, and simulation. Qualitative approaches are also noted as important by interviewees who use surveys, interviews, and other qualitative approaches, especially to understand stakeholder interests and needs. Lab experimentation is another common methodology within CEE research, the scope of which sometimes expands beyond a mere room or floor of a building. For instance, one

interviewee described their lab as, “I do lot of large scale research. We construct full-sized components of bridges and buildings, and we have a hydraulics system, hydraulic jacks and actuators that can apply hundreds of thousands of pounds of force to simulate real world situations.” This researcher, along with others who work within structures, the built environment, and connections with the natural environment, works in the field and combines lab research with testing approaches in real world conditions: “We are collecting soil samples, but we are even more interested in measuring [soil properties] in situ.” In addition to laboratory research, multiple interviewees cited literature reviews as a necessary component of their methodology, particularly in identifying methods, tests, and other approaches used by others that they can adapt to their work.

Research Documentation

There were a variety of preferences for documenting the research process, as well as the recognition that there is little consistency. One interviewee noted, “I would say that one of the big challenges is the differences in software. Some people use Word, some people use Google Docs.” Dropbox was preferred by one researcher over Google because, “I’ve found Dropbox is much better than Google for synching,” while a second researcher simply said, “I’m not a huge fan of Google Drive.” Conversely, another interviewee noted that “. . . we really started appreciating . . . the Google shared documents and Sheets. . . “ while at the same time noting that “I don’t know if I’m old school with that or not, but I still do also create another copy onto another drive.”

Interdisciplinarity

All of the researchers mentioned the interdisciplinary nature of their research. One researcher described how they incorporate knowledge and methods from disciplines outside engineering, such as reaching out to information sources or colleagues to build base knowledge on a topic related to a particular research project. Another interviewee mentioned the need to examine the methodology of other disciplines, such as medicine, in order to glean insights into qualitative analysis. Another researcher mentioned the difficulty that students have in communicating with researchers from other fields or seeing a problem from various perspectives because the student is intensively following a straight line path in one focused field. Still another researcher spoke of knowledge fragmentation and stated that “we need to bring knowledge back together again and start connecting it again.”

Collaboration

Often CEE research is application oriented. For example, one interviewee indicated, “I think that's why it's so attractive to collaborate with my group because at the end, numerical, analytical, as well as the lab people, they all work towards the application in the real world. So the data that we collect are usually really of interest for them for calibration and validation. And I need their information because in the field I'm affected by so many different things. I cannot

isolate one process. So they can isolate a process, inform me how that would affect [a target research material] and then we can compare that.”

The interviewees generally thought that collaboration is becoming easier with the advent of cloud computing and cloud storage to share and backup work. One researcher also indicated that the availability of a centralized and field-specific data repository is “a huge plus to our field.” However, communicating and sharing data and work across very disparate disciplines remains a major challenge as detailed further under ‘challenges for collaborations.’

Current Practice in Collaborations

There are many collaborations within and across fields and among academia, industry, government, business, and communities. These collaborations are often distributed and integrative. A few examples of collaboration domains include geosciences, meteorology, oceanography, environmental sciences, natural resources, ecological fields, business, social sciences, transportation, and biological sciences. Some examples of entities that the researchers collaborated with or received funding from include the Bureau of Transportation Statistics (BTS), the Department of Transportation, DMV, the American Association of Railroads, the American Institute of Steel Construction, the National Steel Bridge Alliance, and the Federal Highway Administration.

Interfacing multiple domains, such integrative work is often tailored to practical challenges and societal impact to gain “a more holistic view” of research problems. For example, one interviewee mentioned that “some people on earthquake engineering get involved with the social sciences folks and some of the decisions...like how much damage or what kind of performance ... are we okay with as a society?].”

In cross-sector collaborations, it is important to emphasize that feedback from cities or communities and “practice reviews,” or information gathering about current practices via meetings with practicing professionals, are particularly valued by a few of the interviewees as practical inputs in their applied research.

Other collaboration and supporting partners involve technologists, librarians, instrument developers and additional support personnel. One interesting example given by a researcher is about working with a new type of librarian that expands traditional libraries’ capabilities into working with “big data” and “unstructured text search.” In this example, a librarian in the National Transportation Library under BTS was “able to search these large streams of unstructured text to generate knowledge out of it.” In the interviewee’s opinion “... librarians are not what you think of... what you grew up with in school; their roles have changed dramatically.”

Challenges for Collaborations

Another theme is about networking from the experimentation side and the data collection side. As for experimentation, one interviewee described a National Science Foundation (NSF) project

focusing on building the Network for Earthquake Engineering Simulation (NEES), which aims to create a network of laboratories that anyone with a winning NSF proposal can access and use. On the data collection side, the researcher further described their efforts in establishing a centralized data repository for their test data as a result of NEES. Here, as the interviewee described, “the idea was that you could integrate this better, and so if somebody was trying to do more of a simulation, they could take data from the different sides.” But they encountered major challenges during this process. One overall challenge that revolves around disciplinary barriers, terminology differences, and measurement inconsistency, was described by the interviewee:

“It turned out to be an absolute nightmare because we had people coming from the tsunami side, from the geotechnical side, from the seismology side, from the structural side, and so on. And we all measured different things in our experiments. And if you try to build a database to be able to take all of that and in some fashion integrate it[,] [p]utting the data in is not that difficult, making it usable turned out to be incredibly difficult.”

Related to the above challenges, there are also collaboration barriers as a result of people using different information management tools and platforms. Data transfer and “currency of software” are also challenges. Language is another challenge, particularly with regard to international collaborations.

Data Collaboration Challenges: Security, Confidentiality, Liability, Reliability, Standardization

One interviewee mentioned that data source security, confidentiality, liability, reliability, and, standardization issues are all challenges particularly in collaboration with public sectors. In this case, the researcher works in water infrastructure systems and is developing a database. As they indicated, “cities are our real labs ... most of my data also comes under homeland security because these are ... all buried infrastructure.” Another issue is liability with regard to utility systems, “they don't want their data to go in a public domain because that becomes a liability for them. Then people will start filing lawsuits and this and that, liabilities, so they will not share the data with me if they don't feel confident that I will only use [it] for research... And that's why we have to sign a lot of confidentiality [agreements].” Furthermore, standardization and frequency of data collection are additional challenges that come with the data from city networks. In the interviewee's words, the “devils are always in the detail,” data is “time dependent” and it is important to understand the “physics of the data” to ensure the reliability of interpreting it. A major problem is that most of the data source areas have only collected data at the start and then at an endpoint when an issue has occurred. Additionally, “math will not do magic” if there is not enough data to construct reliable models.

Future Direction and Needs - Towards Knowledge Integration

As integrative research becomes increasingly prevalent, there is the need to address information and communication challenges in collaborative work. This leads to required changes in current educational practice to promote multidisciplinary understanding, data sharing practices, and cultural exchange between disciplines. In one interviewee's opinion, such changes will help create an "open trust space" and facilitate the "communication of ideas" across disciplinary boundaries.

Notably, one interviewee described "a system of systems" approach to sustainability and resilience work. By "coupling all these different systems together" and "integrat[ing] many different knowledge domains" such as environmental, economic, and social systems, this particular researcher is participating in the efforts to solve sustainability challenges and improve "societal well-being." As they further described, "we're looking at agriculture and energy and water and the intersection of all of those using a kind of a system of systems approach for ... [understanding] these infrastructure systems on coastal flooding." But the major challenges remain with "knowledge management systems," as the interviewee explained, "because we'll be trying to do things where there won't be information, or it's a mess, or it's all in the wrong format." So as researchers are trying to solve complex problems using a systems approach, the daunting task for library and information professionals is to develop information and knowledge management systems where people can pull all sorts of information into an integrative environment and use it in different ways.

Data Practices

Data are an important aspect of the research lifecycle, and the CEE interviewees produce and collect wide varieties of data. Most mentioned that they collect data from the field, data produced in highly controlled lab environments, data produced and obtained from computer simulations and models, and photography and video imaging data. Although data size ranged from small data sheets to large terabytes and petabytes of simulation data, researchers did not express worries about the storage size necessary to store and backup their data.

Data Management Practices

Like the varied lab structures that exist throughout these interviewees' labs, data management practices varied. With the wide variety of data collected, stored and analyzed, many researchers used print lab notebooks, local hard drives, external hard drives, online cloud storage systems like Google Drive and Dropbox to collect and manage their data. For researchers working in the field under less than ideal environmental conditions, hardcopy notes were easier to use than digital devices for data collection.

Regarding the use of software to manage data, there was much variety and conflicting preferences among the interviewees. The two software packages mentioned most frequently for data analysis were Excel and MATLAB. However, the cost of a standalone MATLAB license was mentioned as a deterrent, and therefore Python (a programming language) and its packages were attractive to some researchers for analysis. One researcher, when asked how they would store data said, "I'd try to put it into Excel because I think that's one of the more easily portable things". However, another researcher discouraged students from using Excel because "it's clunky," and used MATLAB almost exclusively. Google Sheets was mentioned positively by one researcher because of ease of sharing data, but many interviewees did not like the Google products. Others described the need to create their own applications or to use extremely specialized applications or simulation systems, such as one researcher who described a system they use in this way, "there's really only one company in the world now that manufactures that level of sophistication."

Moreover, researchers mentioned how time consuming it was to teach graduate students how to manage and organize data. One interviewee suggested that graduate students be required to take a short course at the beginning of the semester where data organization and metadata creation were taught so students can understand the entire process of experimentation.

Sharing and Reusing Data

"It's expensive to collect good data and so, quite often I'm writing a paper now that's using data from five or even 10 years ago."

As the field moves toward more computational modeling and less replication of previous lab experimentation, many researchers are looking to reuse raw data collected from previous experiments to build their computer simulations and models. Data used in this manner can be found in paper reports, where researchers can digitize it and extract the digital data from the paper report. One interviewee noted, "If data were made more available, you could almost create a whole research area where people could go and mine other people's data and reanalyze it and create more knowledge."

However, much of the raw data that can be reused "resides in the archives of the professors," said one interviewee. "And, when they retired, all of that was tossed out."

To prevent this valuable data loss, some researchers interviewed are proactively taking steps to ensure they can share their data even before collection and analysis. One researcher mentioned that they have been preparing their data-analysis coding with the thought of sharing them via open access. Also, several researchers interviewed have found that the CEE community is much more open to sharing large computer simulations, which tend to use primary data to re-create and re-analyze problems, rather than sharing their raw data.

Other primary modes of sharing data included: data repositories (managed by academic institutions, federal agencies and industry groups), data publications, cloud storage and personal websites. Some mentioned the mandatory data sharing requirements of the National Science Foundation as an impetus to share data, and some mentioned the stringent metadata requirements and other documentation for sharing through repositories.

Barriers to Sharing Data

Although the researchers were pleased to talk about the benefits of sharing and reusing data, many of them mentioned the following challenges and barriers that exist within the CEE field that prohibit active data sharing: (1) exposing researcher's warts; (2) the lack of data standards; and, (3) time.

Exposing researcher's warts

"The biggest drawback is that I have to expose every little wart to the inspection of my colleagues. I did something wrong during the test, something went awry, it has to be there. And, in the academic community, many people see that as way too much exposure...You can take anything and trash it."

After digitizing data from a third-party report, one interviewee commented, "We had a big discussion about whether or not to make the digitized data available, and ultimately, I didn't want to do it because I was worried that some of the researchers that conducted the original data or original tests would find it and say, 'This isn't my data...somebody took a ... plot, tried to digitize it and it's not correct.' And, so I didn't feel comfortable with that... If we had the actual original... data, that would be nice. That would be the thing that we could feel confident in sharing."

The lack of data standards

"That's the problem. It's not collecting or even maintaining the raw data but five years from now understanding what were the circumstances it was collected under."

Concerns with sufficient data description and standardization arose in many of the interviews. The lack of data standardization in many fields has created problems when trying to re-use and share data. Not knowing how a test was set up or constructed or what types of data are collected can cause quality challenges. One researcher stated that if they blindly follow the data that have been published to re-create a model, the results will be inconsistent since many of the measures are interpreted differently. "And, it's not because my model is bad, it's because the data is not the same."

Although some types of data have been standardized (like weather data), data collected from interdisciplinary teams lack standardization, which can cause confusion when sharing, as described in the quote from an interviewee in the Collaboration section about the NEES data repository.

Time Required

Preparing data to be shared is time-consuming, especially for repeated experiments that may use different instrumentation or collection methods. One researcher mentioned that to share their collected data, they had to document their curation protocols for how they processed their raw data. Although protocols for repeated experiments can be automated in a sense, when an instrument changes or the instrument is used to capture something else, the protocols have to be updated each time.

Data Privacy and Security Concerns

Some researchers mentioned that some of their data have restrictions and other security and privacy concerns. For instance, crash safety analysis relies upon police reports and Department of Motor Vehicle data, both of which have personally identifiable information (PII) in them. The PII is not given to the researchers, which means that rather than having a complete data dump to the researchers only selected fields are retrieved. Other research involves satellite images which can only be viewed on non-networked computers and by those who have been given clearance to do so. The data for water infrastructure systems is collected from numerous water agencies and is held in confidence by the researchers working on that study. Due to these privacy and security concerns, many researchers have to sign data usage agreements before using the data.

Future Directions and Needs

Since data reuse is growing in the CEE field, researchers seem interested in assistance with increasing responsibilities for creating metadata, designing data management plans and using archival and preservation systems. There is a strong need to provide easy-to-use, less time-consuming platforms that provide adequate data curation. Furthermore, the field needs a strong incentive to adhere to a particular set of data standards for curation and sharing. Information specialists and data librarians can provide direction to possible standards to use for curation; however, direction from industry or federal agencies can provide the incentive and push for researchers to adopt them.

In addition to industry and federal agencies providing a supportive voice in the adoption of data standards, the growth of text and data mining software may help create less burden for researchers in the creation of their metadata in the near future.

In the realm of data literacy, libraries can continue to provide strong data literacy education and consultative services to better fill in the knowledge gaps surrounding data management. Several researchers mentioned the need for data management training for graduate students (as well as researchers), and multiple researchers mentioned sending their students to the

library for assistance in methods and tools for data management and data analysis. Libraries should be more proactive in this space by pushing for these type of services to be integrated in the graduate education curriculum and/or the professional development curriculum for researchers.

Information Discovery and Management

CEE researchers access and use a variety of information sources to develop their research agenda, explore methodologies from within and outside their discipline, and to build personal and shared libraries.

Information Sources

Researchers' descriptions of their information-seeking processes highlighted the importance of interdisciplinary and multi-format source exploration. To stay up to date in their field and research topics, and to build their collaboration networks, researchers' main source types are: conference materials and publications, industry and trade news and magazines, theses and dissertations, and person-to-person communications. To gather research on a topic in preparation for a new or changing study, scholarly journal articles, reports, published conference proceedings, and theses and dissertations are the main source types used. Interviewees highlighted textbooks as a useful source for foundational and interdisciplinary research, such as: for a quick overview of an area prior to expanding one's research into a new field; as a source of quick general statements needed to set up a background section for a report or project; or to reference a formula or other specific, foundational information.

Reuse of Research for Methodology Development

Methodology research in particular focused on interdisciplinary information seeking. CEE researchers described interdisciplinary research as a key strategy to identify new ideas to apply to a particular problem or issue in one's own research. Multiple interviewees discussed this at length and highlighted examples of searching the literature in fields that one might consider distant from CEE. For example, one researcher, looking for new ideas for computational models to address a complex modeling need, found methods in medicine that could be applied to water research, "During [earlier work], I developed an algorithm to automatically [detect] ... cracks. ... You can detect [a] lot of ... things if you can define the features. How do you define a crack? You can't say a straight line ... so it is very difficult, right? So, I learned a lot from published research in the medical field [to address this problem, particularly,] how they detect arteries." Research and exploration of datasets from their own and other fields informs CEE research methodology development.

Discovering and Accessing Information

CEE researcher information discovery strategies are almost entirely conducted online with occasional need for physical information sources. While all interviewees mentioned the value of peer-reviewed materials, most do not rely exclusively on library databases to conduct their literature searches, and most favor Google and/or Google Scholar as primary search discovery tools. As noted in this report's section on Collaboration, within the Current Practices in Collaboration, text mining, described by an interviewee as a 'capability to search large streams of unstructured text and generate knowledge out of it,' is an emerging technique of interest for information searching and refining - to search and pull project-specific information or to identify themes from large pools of textual sources.

Google Scholar

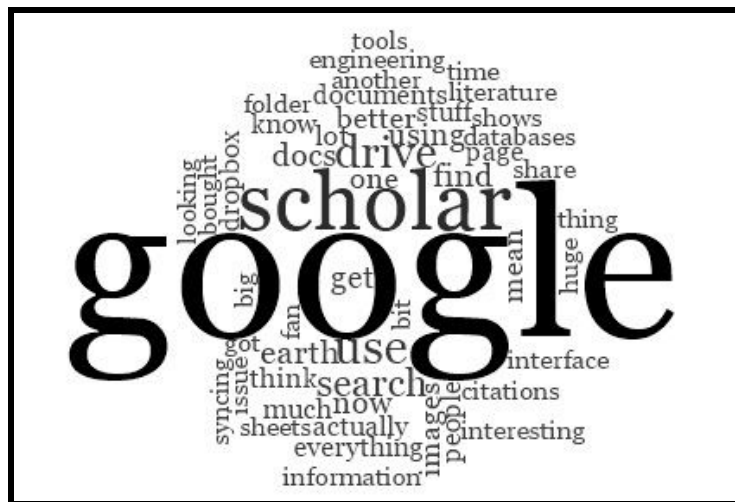


Figure 1. Word frequency count based on a search for "Google" within CEE researcher interview transcripts (parameter includes approximately 5 words on either side of search term).

The majority of researchers interviewed indicated that Google Scholar (GS) is the key search engine of choice to identify publications in their field and on a current topic. As one researcher commented, "I don't think that's the way it should be, but there's no time. And I think when we're trying to do things, we are just rush, rush, rush, let's just see what we get. ... Sometimes I'll go there, then I'll go to the Engineering Village and Web of Science afterwards. After I get my first cut through." Comments also revealed that interviewees identify GS with researcher-specific or academic scholarship partly due to their equally extensive use of GS citation profiles. The same researcher who commented above follows with, "I think Google Scholar Citations is the one that [is] probably the most widely used just because it's the most accessible. Most [of] the people who are publishing and doing well in publishing have public Google Scholar sites in my profession." They also go on to note the use of GS profiles to pull citation counts for promotion and tenure dossiers, and to create alerts to find out when others cite one's publications.

Following a GS and possibly also a Google search, prior to or instead of using a disciplinary database, another interviewee described their next step as reviewing and following up on references from papers found via GS results. Out of the researchers interviewed, 7 use GS as a primary research search tool, with only 1 researcher describing a primary focus on Engineering Village or another database (Table 1).

Table 1. Use of Google Scholar and research literature databases by CEE researchers for literature searches.

| N Researchers Interviewed | Google Scholar Use Level | Other Search Tools Used |
|----------------------------------|---------------------------------|---|
| 2 | Exclusively use GS | None |
| 1 | GS is primary search tool | Possible follow up in Engineering Village or other databases |
| 1 | GS is primary search tool | In tandem with Engineering Village or other databases |
| 1 | GS is primary search tool | With Google and ResearchGate |
| 1 | GS is primary search tool | With Mendeley |
| 1 | GS is primary search tool | With Wikipedia (focus on referenced sources) |
| 1 | GS is secondary search tool | Primary search tool is Engineering Village or other databases |

Other search tools

Other search tools described, such as subscription research databases provided by the Library, included:

- Engineering Village
- Web of Science
- The CRC Handbook
- The library search box on the library website (an aggregated search of most online and physical resources at the library)
- Specific subject-area organizations, repositories, or sources, such as the National Transportation Library

CEE faculty interviewed also tend to explore information sources widely via Google or via other tools or sources: to find datasets or reports by others in the field, to review sources cited in Wikipedia articles, and to contact researchers or field practitioners directly for email discussions, phone or online meetings, or to arrange hosted visits to discuss topics of mutual interest.

Organizing and Managing Information

Information management relates to many aspects of CEE research and researcher practices. With regard to managing their personal or group libraries, interviewees primarily use hierarchical folder systems, either on a personal or external hard drive, or via an online storage platform, such as Google Drive or Dropbox. Most researchers described a personal organization system that groups information by research study or area and by specific topics within these. Some also used, or primarily used, a citation manager, with EndNote being the most common application mentioned. An important consideration for all interviewees was a method to share these libraries with their student researchers and collaborating colleagues. Sharing information is accomplished through making a copy of personal libraries, through sharing options in a platform already in use (Google Drive, Dropbox, citation managers that allow shared or online libraries), or via customized methods such as a lab website internal page.

Information Literacy

Information searching and management is often handled by student research assistants. As mentioned throughout this report, multiple interviewees brought up the importance of student training for searching, managing, and evaluating information sources.

CEE Information Guidance on Information Sources and Searching

In information searching practices, researchers encourage students to search widely and persistently, and interestingly, they direct students to library resources and subscription databases much more than they report relying on such tools themselves. In some transcript excerpts, it seems that perhaps researchers are still using library search tools more than described when they shared their searching preferences, especially for known item searches or for difficult searching needs when a review using many search tools and sources is needed. One researcher notes, "...even in the class I do the same thing, I tell them do [an] extensive literature review," and "I encourage my students to look [in] other fields. If something is working really [well] like, qualitative analysis in some other field, I would like to [use it]." Another describes a thorough and collaborative research process to inform a new method or protocol: "We go out there and find ... whatever we can. Typically ... I guide a student through that. This is one area where the library becomes almost indispensable because they have to look at all of that stuff. ... By and large, we are able to locate almost anything that we want. In some cases, we go much deeper which means that we take the whatever is in the literature, and then we go looking for the PhD theses." This researcher adds that they use their contacts network or the library to get difficult to attain materials. Another interviewee says, "Of course, I think Engineering Village is just... It's the place to go, and that's where I tell the graduate students to

go. I don't know if any of them really use it that much, but for graduate students who have complex searches, ... Engineering Village is, ... the best place to go.”

Information Selection and Synthesis - Critical Appraisal and Technical Writing Skills

Training students in critical appraisal of source quality and writing skills is another common thread. As one researcher describes, “That's actually one of the skill sets that I think our students are somewhat lacking ... I think the library actually has some initiatives that are helping to address this[:] ... how do we teach students or how do we teach the new generation coming in to evaluate what they're seeing in a way that it becomes useful. ... I mean ... being able to contextualize it to understand whether it's valid or what... to understand how much validity you should put into it.” This same researcher describes their teaching of this topic as, “when I'm teaching students and actually when I'm working with my grad students, I will tell them that that information may or may not be what you need, but you need to check it against multiple other information [sources].”

Technical writing skills are another area where researchers would love to see more support for student skills development. Researchers can identify difficulties that students have with writing, but they do not always have the time or educational background to support them in improving. “... one of the challenges I have is proficiency in technical writing of students. ... I rely on the students to create a draft but technical writing is often not their strong suit in engineering. And so, I have relied in the past, especially for international students, on taking that to the Writing Center. But there's a gap. The Writing Center's good at writing but not necessarily... Usually [they don't] have technical expertise in engineering ... [or expertise in technical writing as a specialized form of writing.]”

Information Use Barriers

Access, or lack thereof, to search tools and sources was mentioned by multiple interviewees as a factor that influences their use of certain information sources. Though some interviewees described familiarity with library services, such as interlibrary loan, that support acquiring sources that are not readily available, other researchers highlighted barriers to access. One researcher described issues with use of the university's virtual private network (VPN) to access subscription resources without going to the library's website, and with access to subscription resources for those not affiliated with a research institution, “we do a fair amount of literature review. And any more, particularly because I was in [part of state agency], I relied almost exclusively on what was accessible through the internet, through digital, and... Even getting to [the institution's] library could be an interesting challenge sometimes.”

Research Communication and Dissemination

Publishing within the Academy

When it comes to publication practices in the CEE community, the majority of researchers publish in journals or published conference proceedings. Several interviewees cited pressure from the academy to publish frequently, particularly for younger faculty members pursuing tenure. In these instances, as well as in cases where graduate students are on the job market, the reputation and impact factor of the journal matters to the researcher. There is also concern that a rush to publish can lend itself to a lack of “quality control.” One interviewee commented: “This is my own two cents about the whole academic system, but I think, unfortunately, we've created this currency of journal articles that you have to pay your dues by creating so many journal articles so you can get tenure. So I think oftentimes the quality, it's superseded by the quantity.”

Another concern regarding journals is the “lag time” in publishing. One researcher stated that regarding research dissemination, “journal articles...are the foundation, with the exception of the shift that's taking place because of real-time delivery, if you will. Journal articles take a year, two years.” A second interviewee cited the same issue, particularly for more reputable journals. In addition to journals, engineering researchers also publish reports that feature their research-- predominantly federal reports.

One interviewee discussed publishing via “data papers,” a new type of publication seen as valuable in their area of research, describing it as, “[There] are now data papers, where we consider it a paper, it's as valuable as a paper, where I sit down and write everything about my data set and tell you what's there and how to find it.” The researcher further describes this newer article type as comprised of a description of a particular dataset, a brief description of the study it relates to, the methodology used to create the dataset, as well as where the data can be found and any permissions or certifications required to access it. Here, according to this researcher, “the idea is the data lives on.”

Conferences are used as another means of disseminating research. One interviewee pointed out that conference attendance allowed access to more contemporary research-- whereas older journal articles, for instance, often no longer maintain the accompanying data.

Open Access

One point that arose several times is the concept of open access publications, which seemed to garner mixed reactions amongst engineering researchers. Several interviewees stated they were open to publishing work through this avenue but were concerned with the associated costs. “I think that if it was cheaper I would probably always be publishing Open Access.” Another asserted that they would never pay fees to publish open access in a journal. With cost

being the primary challenge, it seems unclear how the academic community will continue to engage with open access, especially as it gains traction within various online spaces. In fact, one interviewee cited this as a possible area of focus for library resources: “I think this open access thing, anything that the library can do to help that be seamless.”

Another challenge that CEE researchers face is the question of copyrights related to their published research. One interviewee in particular who maintains their own website stated: “I feel like it might be more beneficial for me to put [my articles] on there so that they get more exposure and potentially more citations to play the whole tenure game. But I am always unsure of what is legal, based off the copyright.”

Institutional repositories or disciplinary repositories are two areas where interviewees are working to share their research, and would like to do so more. Barriers to doing so mostly include the time required to identify copyright compliance, identify the repository or get access to submit to repositories, find the right documents to submit, and go through the submission process. In response to being asked about whether they submit their articles to the institutional repository, one researcher said, “So that’s a difficult one. So it depends a little bit on the journal, and how open we can handle that. Like putting it on the repository, I have to say that my students have been better than me with that. Because I’m pushing for it, but then I have to admit that myself, I’m sometimes not that good with it.”

Other researchers also mentioned needing help to disseminate research appropriately, citing possible online tools such as Overleaf, Mendeley, and ResearchGate, but needing assistance to fully utilize them.

Future Needs and Directions

An additional area of focus for possible library support lies within the realm of social media. On the whole, there is debate amongst scholars about whether social media is a productive or valuable tool. One interviewee mentioned social media as a potential site for data acquisition, but the rest were varied in their utilization of these sites. Another scholar requested assistance in gaining visibility through social media.

A consistent struggle across all areas of both publication and the circulation of scholarship is the issue of time constraints. In fact, one interviewee professed that publishing in certain fields of engineering becomes unappealing due to the amount of time involved, coupled with the lack of recognition. Another scholar cited similar concerns-- that using valuable amounts of time to publish detracts from time spent on research.

This finding perhaps leads to a valuable insight: career status can inform publication practice. For certain areas of the discipline and for certain types of faculty, publishing on the whole is a crucial component of one’s career. But for others, particularly tenured faculty, this pressure is significantly less, which leads them to publish only tangentially and mostly alongside their

graduate students. For instance, one interviewee stated: “Very few papers do I have that I’m first author. I’m just not into that mode.”

So while it’s true that all scholars interviewed do, to some degree, disseminate their research in some fashion, their frequency and motivations behind doing so differ. And finally, most scholars seem able to benefit from additional support with regards to the publication process.

Discussion / Implications

Virginia Tech, University Libraries Opportunities (new services, communication opportunities, and continuing priorities)

An overall reflection of the research findings has shown a recurring theme and common issue of fragmentation across the current academic practices, encompassing research, data, knowledge, teaching, and learning. And, some aspects of these difficulties identified during the interviews, such as research collaboration struggles between academic and public sectors that can be impeded by data liability and reliability issues, may be outside the library’s scope. However, this study’s analysis of the interviews identified four specific areas which the library should act on:

Supporting Faculty Open Access Goals

The first is making faculty publications openly available. While it is clear that CEE faculty are usually willing to make their publications openly accessible, it is also clear that many are unsure how copyright restrictions apply. One interviewee stated, “that would be a nice thing also for you guys to inform us as to exactly what I can do and not do,” while another said, “But I am always unsure of what is legal, based off the copyright.” The library offers a workshop several times a year, which specifically addresses the question “Where Can I Post My Publications?” but even with this outreach, researchers are still uncertain about their options. The perceived complexity of making items accessible is another hindrance. As noted, one interviewee fully recognized the advantage of making items accessible, and said, “my students have been better than me with that. Because I’m pushing for it, but then I have to admit that myself, I’m sometimes not that good with it takes time.” Therefore, for faculty to provide open access to their articles, two obstacles must be overcome: first, faculty must be made aware of the legality of open access, and then there must be a method to provide quick and seamless open access. Even if faculty can be shown the legality of providing access to their articles, if the method to do so is cumbersome or time consuming, it will not happen.

One possible solution for this that the library is just beginning to implement, is a recent connection of a faculty activity reporting software and the institutional repository. This connection makes submission of works a simple process where journal copyright policy guidance is provided within the activity reporting tool, and the metadata about the journal article (or other work) is pulled from the activity reporting system as well. The faculty member’s task is limited to locating the appropriate file/s to submit, and uploading them through this system.

Certainly, better outreach and marketing of this simplified method for sharing research is needed.

Engaging with Google Scholar as a Major Research Tool

The second area which the library needs to consider is the use of Google Scholar. A significant majority of the interviewees used GS as a primary, or even exclusive search tool. This impacts the library in two areas, the first of which is instruction and user support. While some librarians may already be incorporating GS into instruction and research support, the library as a whole cannot ignore its use by faculty and students, and must become thoroughly familiar with the GS search capabilities and how to integrate it into the search strategies of our users. Such instruction can address effective methods to use GS including connecting to full text via library subscription sources from GS (through adding library links in GS settings), optimizing one's search strategies for GS (advanced searching can be challenging), and setting alerts for new results from a search. Tools to retrieve data from GS, such as Harzing's Publish or Perish software, or citation software plugins may also be of interest for power users, to avoid the need to download citation information one item at a time. Training and how to instruction and support can also address benefits and challenges of using GS. Some of its most useful aspects include: using it as an initial search for an introductory and interdisciplinary view of research available on a new topic to identify highly relevant key articles, fields, and journals to inform further in-depth research via disciplinary databases or journal hand searching; or as a concluding search to complement traditional database searches to ensure one has identified results across disciplines, geographic regions, and languages as GS supports identifying scholarly research across more languages and geographical publishing regions than most traditional databases. Pitfalls or considerations for using GS to include in training, could include its unreliability for retrieving the same search results due to its customized retrieval for each user and its ever-changing source base. Also, while GS has been shown to be comprehensive for known-item searching³, it can be less reliable and less effective in providing comprehensive results for a direct topic search compared with some disciplinary databases - primarily due to its cutoff point of allowing only 1,000 results to be reviewed from a search, as well as the added time and difficulty involved in reviewing or exporting citations from Google Scholar due to its inadequate user interface and export options.⁴ A brief look at a review of 91 studies⁵ evaluating

³ Carmen Cole, Angela R. Davis, Vanessa Eyer, John J. Meier. "Google Scholar's Coverage of the Engineering Literature 10 years Later," *The Journal of Academic Librarianship*, Volume 44, Issue 3, 2018. Pages 419-425. DOI: <https://doi.org/10.1016/j.acalib.2018.02.013>. Alberto Martín-Martín, Enrique Orduna-Malea, Mike Thelwall, Emilio Delgado López-Cózar, "Google Scholar, Web of Science, and Scopus: A systematic comparison of citations in 252 subject categories," *Journal of Informetrics*, Volume 12, Issue 4, 2018. Pages 1160-1177. DOI: <https://doi.org/10.1016/j.joi.2018.09.002>.

⁴ Bramer, Wichor M., Dean Giustini, and Bianca MR Kramer. "Comparing the coverage, recall, and precision of searches for 120 systematic reviews in Embase, MEDLINE, and Google Scholar: a prospective study." *Systematic reviews*, Volume 5, Issue 1, 2016. DOI: <https://doi.org/10.1186/s13643-016-0215-7>

⁵ Gali Halevi, Henk Moed, Judit Bar-Ilan. "Suitability of Google Scholar as a source of scientific information and as a source of data for scientific evaluation—Review of the Literature," *Journal of Informetrics*, Volume 11, Issue 3, 2017. Pages 823-834. DOI: <https://doi.org/10.1016/j.joi.2017.06.005>.

GS in comparison with disciplinary databases supports these instructional points, however a full review of GS studies may provide different or additional implications for best practices in using GS.

GS may also have a second impact on the library at some point, in terms of collection development. If GS is the go to source for many of our users, does that impact how much we spend on subscription based indexes? Would money spent on some current subscription databases be better spent on other resources? To answer these questions, the library would need to further investigate and compare user searching preferences and behavior, full text usage and provision from such library databases, and overall questions of the long term sustainability of prioritizing GS as a true substitute for some disciplinary databases.

Data Education and Services for All - Students to Faculty

Data services addressing data literacy, data curation, and data management training are a third area for the library to build support for researchers. Researchers themselves identified several areas where they would benefit from assistance to share data. These included identifying existing professional community standards for data description and documentation to increase data reusability after sharing, support for best practices in data digitization projects, and support for the time and effort needed to document and format data for deposit in a repository.

As one researcher described, “So if I could just give you my data and you organize it in a repository, that would be already something. I can take care of the publications. I can take care of the collection of the data. But on that end, and I know that, for example, {my graduate student} has... I don't know if he worked with you directly or one of your colleagues, I know that you guys helped him a lot making that possible, and I truly appreciate that because I have to admit, I didn't have to spend any time on that. And that was great. And so that is helping a lot.”

As noted in this excerpt, and in the data practices section of this report, graduate student training is another key area for support. Libraries, as a hub of data management expertise, can provide curricula that emphasizes experiential data education that can be integrated into graduate education and professional development programs so graduate students have the skills to organize data for repository use. Moreover, libraries will also need to think beyond data repositories to other open research collaboration platforms like the Open Science Framework for ways to promote and encourage data sharing.

Critical Appraisal and Evaluation of Information for Research - Continuing Importance

While instruction and support for source selection, evaluation, and critical appraisal as part of the discovery and research process, is a common part of academic librarians' work, comments from interviewees in this study highlighted its continued importance. Working with faculty and students to ensure that students at all levels are practiced in methods to investigate and

understand the sources of their information, and in methods to verify the accuracy and validity of information they find to inform their research is key.

Appendix 1

Participant Consent Form

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY

Informed Consent for Participants

in Research Projects Involving Human Subjects

Title of Project: Research Support Services Study - VT and Ithaca S+R - Civil and Environmental Engineering Focus

Investigator(s): Ginny Pannabecker, vpannabecker@vt.edu / 540-231-7980
Larry Thompson, larryt@vt.edu / 540-231-8693
Erin M. Smith, erinSmith@vt.edu / 540-231-5663
Yi Shen, yishen18@vt.edu / 540-231-5329

I. Purpose of this Research Project

This research study seeks to examine the research practices of academics in Civil and Environmental Engineering (CEE) in order to understand the resources and services these faculty members need to be successful in their teaching and research. This Virginia Tech study is part of a larger multi-institutional study seeking to understand CEE faculty research practices and support services needs. The results of the research, including photographs, will be disseminated through conference presentations, scholarly articles, and/or as part of publicly available reports published online through VTechworks, VT's institutional repository, via a public Open Science Framework project site, and the Ithaca S+R website. The Ithaca S+R report that reports aggregated data from institutional studies will be issued using a creative commons license which will also enable it to be deposited in VTechWorks. Reports and other publications will be shared with participants.

II. Procedures

Your participation in the study involves a 60-minute audio-recorded interview about your research practices and support needs as a Civil and Environmental Engineering scholar. We also may take photographs to document your work space if you consent to photographs; however, you will not appear in the photographs. Although work spaces will not contain images of you, personal work spaces may be identifiable to persons who have worked in those spaces before. Your participation is completely voluntary. You are free to withdraw consent and discontinue participation in the interview at any time for any reason.

III. Risks

There are no known risks associated with your participating in this study.

IV. Benefits

You may experience benefits in the form of increased insight and awareness into your own research practices and needs. No promise or guarantee of benefits has been made to encourage you to participate.

V. Extent of Anonymity and Confidentiality

If you choose to participate, your name will not be linked to your interview responses or work space photographs at any time. We do not include your name on any of the interview data and there is no link between this consent form and your responses.

The Virginia Tech (VT) Institutional Review Board (IRB) may view the study's data for auditing purposes. The IRB is responsible for the oversight of the protection of human subjects involved in research.

VI. Compensation

There is no compensation provided for participating in this study.

VII. Freedom to Withdraw

It is important for you to know that you are free to withdraw from this study at any time without penalty. You are free to not answer any questions that you choose or respond to what is being asked of you without penalty.

Please note that there may be circumstances under which the investigator may determine that a subject should not continue as a subject.

Should you withdraw or otherwise discontinue participation, you will be compensated for the portion of the project completed in accordance with the Compensation section of this document.

VIII. Questions or Concerns

Should you have any questions about this study, you may contact one of the research investigators whose contact information is included at the beginning of this document.

Should you have any questions or concerns about the study's conduct or your rights as a research subject, or need to report a research-related injury or event, you may contact the VT IRB at irb@vt.edu or (540) 231-3732.

IX. Subject's Consent

I have read the Consent Form and conditions of this project. I have had all my questions answered. I hereby acknowledge the above and give my voluntary consent to participate in the study as described above including:

_____ being interviewed and being audio-recorded during the interview

_____ having my work-space documented through photographs

_____ Date _____

Subject signature

Subject printed name

Virginia Tech Institutional Review Board Project No. 17-1035
Approved January 21, 2018 to January 20, 2019

Appendix 2

Interview Guide

Semi-Structured Interview Guide

Research Support Services Study - VT and Ithaca S+R - Civil and Environmental Engineering Focus

Research focus and methods

- Describe your current research focus and projects.
- How is your research situated within the field of Civil and/or Environmental Engineering?
 - Does your work engage with any other fields or disciplines?
- What research methods do you typically use to conduct your research?
 - How do your methods relate to work done by others in Civil and/or Environmental Engineering [and, if, relevant in the other fields you engage with]?

Working with others

- Do you regularly work with, consult or collaborate with any others as part of your research process?
 - If so, who have you worked with and how?
 - Lab or on-campus research group
 - Other scholars or researchers [e.g. faculty at the university or other universities, student assistants, independent researchers]
 - Research support professionals: e.g. librarians, technologists
 - Other individuals or communities beyond the academy
 - Others not captured here?
- Have you encountered any challenges in the process of working with others? [focus on information-related challenges, e.g. finding information, data management, process of writing up results]
- Are there any resources, services or other supports that would help you more effectively develop and maintain these relationships?

Working with Data

- Does your research typically produce data? If so,
 - What kinds of data does your research typically produce? [prompt: describe the processes in which the data is produced over the course of your research]
 - How do you analyze the data? [e.g. using a pre-existing software package, designing own software, create models]
 - How do you manage and store data for your current use?
 - Do you use any other tools to record your research data? [E.g. electronic lab notebooks]. If so, describe.
 - What are your plans for managing the data and associated information beyond your current use? [e.g. protocols for sharing, destruction schedule, plans for depositing in a closed or open repository]

- Have you encountered any challenges in the process of working with the data your research produces? If so, describe.
- Are there any resources, services or other supports that would help you more effectively work with the data your research produces?
- **Does your research involve working with data produced by others? If so,**
 - What kinds of data produced by others do you typically work with?
 - How do you find that data?
 - How do you incorporate the data into your final research outputs? [e.g. included in the appendices, visually expressed as a table or figure]
 - How do you manage and store this data for your current use?
 - What are your plans for managing the data beyond your current use?
 - Have you encountered any challenges working with this kind of information?
 - Are there any resources, services or other supports that would help you more effectively work with data produced by others?

Working with Published Information

- What kinds of published information do you rely on to do your research? [e.g. pre-prints, peer-reviewed articles, textbooks]
 - How do you locate this information? [Prompt for where and how they search for information and whether they receive any help from others in the process]
 - How do you manage and store this information for your ongoing use?
 - What are your plans for managing this information in the long-term?
 - Have you experienced any challenges working with this kind of information?
 - Are there any resources, services or other supports that would help you more effectively work with this kind of information?

Publishing Practices

- Where do you typically publish your scholarly research?
 - What are your key considerations in determining where to publish?
 - Have you ever made your scholarly publications available through open access? [e.g. pre-print archive; institutional repository, open access journal or journal option]. If yes, describe which venues.
 - Describe your considerations when determining whether or not to do so.
- Do you disseminate your research beyond scholarly publications? [If so, probe for where they publish and why they publish in these venues]
- Do you use social networking or other digital media platforms to communicate about your work [e.g. ResearchGate, Twitter, YouTube]?
 - If yes, describe which venues and your experiences using them.
 - If no, explain your level of familiarity and reasons for not choosing to engage with these kinds of platforms.
- How do your publishing practices relate to those typical in your discipline?
- Have you encountered any challenges in the process of publishing your work?
- Are there any resources, services or other supports that would help you in the process of publishing?
- Prompt for further information if time and if not covered in answers to the above:

- Is it important to you to know how your colleagues and others (such as the public, policy creators, etc.) use your work? How do you track this kind of engagement? (i.e. citation metrics, altmetrics: tweets, downloads, comments, reviews)

State of the Field and Wrapping Up

- How do you connect with your colleagues and/or keep up with trends in your field more broadly? [e.g. conferences, social networking]
- What future challenges and opportunities do you see for the broader field?
- Is there anything else about your experiences or needs as a scholar that you think it is important for me to know that was not covered in the previous questions?
 - “If I gave you a magic wand that could help you with your research and publication process – what would you ask it to do?”