Self Evaluation and Personal Statement

Summary. The following briefly summarizes my main contributions and achievements over the last five years at Gonzaga University. Each of the following are explained in more detail throughout my self evaluation and personal statement.

1. Published over 30 peer-reviewed papers. Of these, five papers had GU computer-science students as co-authors (11 students in total).

2. Co-editor of two conference proceedings (books) published by Springer (for the Conference on Scientific and Statistical Database Management in 2011 and 2012).

3. Received over $320,000 in external research funding, including three National Science Foundation awards (as Co-PI) and one award from the Institute for Systems Medicine.

4. Funded 16 undergraduate students via external grants to work during the summer on computer-science research and development projects (40 hours per week for 10–12 weeks). Many of these students were also hired to work part-time with me on research during the academic year.

5. Taught eleven different computer-science courses and 27 total sections, averaging over 20 students per section. All but one of these courses (CPSC 421, Database Management Systems) were new courses for me. Two classes (Software Engineering and Data Mining) were new courses at GU.

6. Taught four sections of UNIV 099 (Strategies for Success)—a one-credit course designed to help students on academic probation. The sections were geared toward Engineering majors.

7. Received funding from the GU KEEN grant ($10,000) to purchase 18 personal robots and to develop corresponding teaching modules for Computer Science I (CPSC 121). I used these robots and modules in the Fall of 2012. This work led to a recent paper (with Dr. Yerion) that will be presented at a computer-science education conference this Fall.

8. Served as Editorial Board member (2006–2010) and Co-Editor in Chief for the Journal of Data Semantics (2011–2013), and served on over 25 conference and workshop program committees. Also served twice as student poster chair for CCSC-NW and proceedings chair for SSDBM.

9. Served on the Academic Council Policy and Planning Committee, the University Mission Advisory Board, the Sponsored Research Office's Grant Advisory Board, the CS Faculty Search Committee, the Herak Lab Renovation Committee, the Herak Center Space Committee (Chair), the CEDE Faculty Committee, and was elected to the University-wide Patent Committee and Commencement Awards Committee. Also serve as a Faculty Senate representative for Engineering.

10. Served as faculty advisor (2010–present) for the Gonzaga UPE honor society and Computer Science club. Also currently academic advisor for over 40 computer-science majors.

11. Received a Gonzaga University Exemplary Faculty Award in Spring 2012.

1 Teaching

I believe that my primary role as a computer science educator is to help students develop the skills they require to achieve their academic and professional career goals. In addition to proficiency in computer science understanding, these skills also consist of a student's ability to effectively communicate, to work productively within a team, to make informed decisions about computing technology, and to participate in problem-solving activities confidently and independently. Through my own experiences as a teacher and advisor, I have found that working with students to achieve these goals can be very challenging, but also extremely rewarding. To be a successful educator I believe it is important to be well organized and enthusiastic, to identify and clearly articulate learning objectives, to be approachable, and to devote time and energy in guiding students through course material and the learning process. Over the last several years, I have made a number of significant changes to the way I deliver my courses. These changes include moving away from PowerPoint-driven lectures to board lectures and by incorporating a number of the techniques that come out of the "active" learning methods of teaching. I strongly believe that as an educator, I must continually work on improving and adapting my teaching to better engage students and to help students learn the latest approaches in computer science—a field that changes rapidly.

Lectures. In previous years, I tried to structure lectures as a mix of PowerPoint presentation, board work, and in-class exercises. The board work often included concrete examples, walking students through homework and quiz answers, and key definitions of terms. Over the past three years, based on comments by computer-science faculty as well as some students, I have moved away from delivering content through PowerPoint, and replaced it by presenting material on the whiteboard (and in some cases the "doc cam", e.g., for larger or more complex diagrams). I have found that using the whiteboard slows the course down, but provides students with more time to take notes and in general more time to "digest" the material. It also provides more opportunities for students to ask questions and for me to adapt material "on-the-fly" based on student questions and feedback during lectures. In general, my experience over the past three years has been that using the board helps engage a much larger percentage of the students in a course because of the slower pace, whereas the much smaller "high achievers" in a course seem to prefer the quicker pace afforded by PowerPoint.

In each of my classes, I have continued to use in-class exercises, which typically consist of students working together in small groups of 2–3 on problems related to topics presented during the lecture. Exercises are usually designed to take between 5–10 minutes, however, in some cases longer exercises are used. I have found the use of in-class exercises to be extremely useful, because they help students reinforce lecture material by "trying it out" and give students the opportunity to work together to "talk through" the material. I have also tried to use exercises to help structure the class, e.g., ideally by lecturing for 20–30 minutes (depending on whether a quiz is given), then 5-15 minutes for exercises, and then 20-30 minutes of follow-up lecture (often interspersed with additional exercises). Exercises also allow me to more directly interact with students individually during class (by talking with groups during the exercises), and in general I have found that students ask more questions during and after exercises compared to when I am just lecturing. In each course offering, I have tried to introduce increasingly more in-class exercises into my lectures. I have found this to be much easier the second and third time a course is taught. I have also started using "warm-up" exercises in some courses (e.g., CPSC 121) in which I pose questions to students—often as a mix of review and challenge questions—to help reinforce material and to introduce new topics.

These warm-up exercises typically ask students to work out a problem on their own, and then discuss their result with a neighbor—again, to try to stimulate discussion and to give students an opportunity to "talk through" the material. Based on a KEEN workshop I attended in the summer of 2012 at St. Louis University, I also began developing "problem-based learning" (PBL) exercises that I have used within my CPSC 121 courses. PBLs consist of students working in teams on a real-world engineering problem or scenario. The units I developed were spread over a week period, in which the first day was used for brainstorming solutions, the second day was used for algorithm design, and the third day was used for algorithm implementation (within the lab). Time was set aside in each class for students to describe their results on a sheet of flip-board paper and to give a short (1 minute) presentation to the class on their work. The PBL exercises seemed to engage students, requiring them to work together in a small team to apply what they have learned in the course to relatively difficult problems. I plan to continue to develop and use more of these problem-based learning exercises within CPSC 121 as well as other computer science courses.

Frequent Short Quizzes. In each of my classes, I typically give a short (5–10 minute) quiz at the beginning of a class once per week. I believe quizzes are valuable for three reasons: (1) they help students become familiar with the types of questions I ask on exams; (2) they provide students with early feedback on their understanding of material (and in a test-like setting); and (3) they provide me with early feedback on how well students understand the material, allowing me to re-explain and review topics as needed. I always return students' graded quizzes on the lecture following the quiz and also go over answers to the quiz questions at the beginning of the class period.

Group Projects. In each of my upper-division courses I try to assign small-to-medium sized group projects. All projects include a project proposal to ensure students start early and so that I can provide early feedback, intermediate deliverables (again, for feedback and to ensure groups stay on track), writing code within their group, a project presentation (between 5–15 minutes, depending on the class size), and a project report. Group projects are assigned in addition to regular homework assignments. I have found group projects to be valuable to students for three main reasons: (1) they provide students with an additional opportunity to apply the skills they have learned in the class to a larger project (requiring the use of multiple topics covered in the course); (2) they provide students with additional experiences working within a team (which is increasingly important within computer science, especially for those students who will graduate and go on to work as professional software developers); and (3) they provide students with additional practice communicating technical content, both orally in their project presentations and written in their project reports. I typically grade these projects on multiple factors: each intermediate deliverable is graded, the presentation is graded on a fixed set of criteria (provided to the students as part of the assignment), the quality of the final report and code written, overall difficulty of the project (i.e., effort of the team), and the contributions made by each group member in the project (based on self-evaluation and evaluation by other team members). CPSC 491 and 492 are project-based classes. In each of these courses, I provide students with a number of intermediate deliverables (assignments) to keep them on track and to receive frequent feedback on their work. I also have begun doing more peer-review within these courses, which gives students practice evaluating their peers, provides valuable information on individual student contributions, and helps to keep students engaged and interested in the other team's projects. When doing peer reviews, I ask students to rate work on a set of pre-defined criteria and to provide "plus/delta" comments, i.e., to list at least two things that were done well (the pluses) and at least two things that could be changed (the

delta). A major challenge I've faced in peer review is that students often will give only "A's". To help with this situation, I often grade the reviews (at least initially) to help motivate students to provide honest assessments. I also spend time in class discussing issues surrounding peer assessment and how becoming good at assessment will help them throughout their career.

This semester in CPSC 462 (a special topics course on Data Mining), I have tried a new approach in which students work in pairs on homework assignments but individually on a project. This is largely an experiment. Homework assignments are larger than typical individual assignments and require students to implement the algorithms and approaches discussed in class (with various hints on how to solve the problems also given in class). Students were told early and frequently in the course that they will need to understand the implementation work to adapt for their final (individual) project, which requires them to apply the tools and techniques to a real-world dataset, to give a short presentation to the class, and to write a technical report on their results. Because there are 40 students in the course, with large and detailed programming assignments, the idea was to treat these assignments more like lab-type work, and then require them to demonstrate their understanding through a larger individual project. It often takes thirty minutes or more to grade a single, large programming assignment. With such large courses, it can be extremely time consuming to grade these assignments, and so the hope is that students are motivated to perform within their small group so that they can do well on the individual project (while making it feasible to spend time grading each assignment).

Homework Assignments. In each of my courses, I assign weekly homework (sometimes longer assignments for upper division courses). For most of my courses, homework consists of short to medium-sized programming assignments or technical problem sets. More recently, I have also begun assigning written homework in addition to regular assignments. The main goal of the written homework is to motivate students to do the assigned readings. These assignments typically require students to do the assigned reading, reflect on what they have read (by formulating at least two questions they have and two things they found surprising or interesting in the reading), and answer a small number of content-specific questions. I've found that students will often not do the readings without the required written part of the reading assignments. For programming assignments, my general approach is to give students the opportunity to practice using the techniques discussed in lectures while at the same time challenging them to think about the material beyond just the examples given in class. This means that to complete the assignments, students must think about the concepts or techniques being discussed, and apply and combine them in different ways to a new problem. Based on student feedback, most students find the homework difficult (since it requires thinking), but very beneficial to their learning. I have also found that a few students are not used to having to go through this type of learning process, and instead have a difficult time because they expect homework to largely focus on recitation (e.g., via fill-in-the-blank or highly directed worksheets in which students simply follow a predefined set of steps to solve a problem). A simple example of the difference would be teaching students about multiplication, addition, and formulas, and then giving them a problem that requires developing a new formula composed of these operators versus showing them an example of a specific formula and then asking them to recreate the formula but with different arguments. I've found that students presented with the former, but that expect the latter, often feel they are not given sufficient information to complete assignments (even though they have been given the information they need).

Grading Schemes. Over the past five years, the way I have assessed students has changed. Prior

to joining Gonzaga, I was primarily involved with teaching graduate courses within a quarter system (as opposed to semesters). My courses initially involved two exams (a midterm and a final), with homework making up 50% of the final grade. Many students expressed concern with only having two exams (each worth a fairly large portion of the final grade), and after speaking with colleagues, I began reorganizing my courses to include more exams, each worth a smaller percent of the final grade. The general grading scheme I use now for most of the courses I teach consists of quizzes worth 15–20%, three mid-semester exams worth 20–30%, a final exam worth 20%, homework assignments worth 30–35%, and (in upper division courses) a team project worth 10–15%. For example, in CPSC 121, assignments are worth 35%, quizzes are worth 15%, each mid-semester exam is worth 10%, and the final exam is worth 20%. I've found this scheme to work well with more frequent exams and where quizzes plus exams make up over half of the final grade.

Retention and Relevance. Because of the importance and availability of computing technology today and its role in many college-aged students' lives, it is easy for students to think that the problems of interest in computer science are "solved". This viewpoint is shared by many students studying engineering, who are required to take a computer science course (CPSC 121) in their first year at GU. Like many fields today, computation plays an increasingly important role in engineering, requiring in many cases students to be able to think in terms of algorithmic solutions to problems, to be able to write programs to analyze and visualize data, and in some cases to write entirely new software applications for their field. Bridging this gap in first-year engineering students is especially difficult because often they don't have a clear sense of what engineering is. Similar issues exist with students in other, non-engineering disciplines, where learning computer science techniques could be valuable for their careers. To help address some of these issues in my own courses, I am interested in incorporating approaches to help make computer science more relevant to engineering students and in general to help retention of computer science and engineering majors. I recently received a grant through the KEEN program at GU to purchase 18 personal "scribbler" robots[1] and to develop a set of teaching modules around the robots. The goal of this work was twofold: (1) to develop a set of assignments that provide a more tangible and "tactile" programming experience for engineering majors to help increase engagement and student-perceived relevance for the course; and (2) to allow students to work towards a number of the KEEN objectives focused on developing an entrepreneurial mindset. As part of this work, I developed seven supplemental assignments involving the robots. In each assignment, students worked in small teams to solve problems in basic robot control (e.g., autonomously navigating robots around obstacles). Each team had their own robot that they programmed using Python to solve control problems. The last two assignments involved competitions: the first required students to "race" their robots around obstacles; and the second required students to play a game of "team tag" (where some robots were taggers and others were racers). There was also one assignment focused on brainstorming product ideas and performing market research related to applications of personal robots. Each robot assignment was in addition to the normal programming assignments given in the class. While there were a few issues with using the robots in the lab, the use of robots made a significant difference in terms of student perceptions of computer science and the relevance of the course. I plan to use the robots again within my CPSC 121 courses in the future, and also to further explore ways to help make computer science more relevant and engaging to engineering and non-engineering students. In the last two years, some of the adjunct instructors have used the scribblers and modules within their

[1] see www.betterbots.com

courses (including Matthew Bell and Bruce Worobec this semester).

Office Hours. I have found that because computer science and engineering students at GU take so many courses each semester, it is effectively impossible to schedule a time for office hours where all students can attend. Therefore, I have tried hard in my classes to ensure that students feel comfortable asking to schedule times outside of my official office hours to meet. I also generally try to leave my door open during the day (unless I'm in a meeting), and tell students they are welcome to drop in whenever my door is open. While a number of students come to my office, I've found that many students do not take advantage of office hours. Those students who do use office hours seem to have a much easier time in the course. Similarly, I also have made an effort to be available and to quickly respond to student questions via email. I am continually surprised, however, at how few students ask questions through email. I have yet to find a good solution to this problem (after trying a number of other technologies to help with this).

Student Supervision Activities. The last three years I have been the instructor for CPSC 491 and 492, which are the courses associated with Senior Design for computer science[2]. There were a total of five senior design projects in the first year, four last year, and six this year. In addition to supervising each project as the faculty advisor, I also hold regular class hours both semesters (3 hours per week) in which I lecture on software engineering techniques, meet with student teams, and have students perform various exercises in teams (such as quality assurance and testing, defining weekly tasks, and giving practice presentations). During in class exercises, I typically talk with teams about their projects (e.g., if they finish the exercise early or I discuss an aspect of their project related to the exercise). These informal discussions help me to stay on top of the teams progress. This year, because of the large number of teams and students, I have begun meeting outside of the senior design course with students. These meetings are scheduled every two weeks (three teams per week) for a half hour, and are meant to supplement the in-class discussions with each team. I also supervised 16 undergraduate research students over the last five years—more details concerning my work with these students is provided under Professional Development.

Student Evaluations and Grades. I have included all numerical and written student evaluations for each of the courses I have taught at GU. Although I taught courses at a variety of levels and to different audiences (computer science majors, engineering students, and arts and sciences students), the numerical evaluations are fairly consistent. The following table summarizes the median scores of the 10 numerical questions on the evaluation forms for each of the courses I've taught over the past five years (except one section of CPSC 491 and the sections of UNIV 099). Also included is the average class grade percentage for each course.

[2]Different than the other engineering disciplines, the CS Department assigns one faculty advisor for all senior design projects, which are organized through the two courses (491 ad 492).

Table 1: Median student evaluation scores (max is 7) with average grade (percent) for each course.

	Courses																					
Questions	F09		S10		F10		S11		F11			S12		F12*			S13		F13			S14	
1	6	6	6	6	6	6	6	6	6	6	6	6	6.5	6	6	6	6	6	6	6	6	7	6
2	6	6	7	7	7	7	7	7	7	7	7	7	7	6	7	7	6	7	6	6	7	7	7
3	7	6	7	7	7	7	7	7	6	6	6	7	7	6	7	7	6	7	7	7	6	7	7
4	6	6	7	7	6	6	6	5	5	6	6	6	6.5	6	6	6	6	6	6	6	5	7	7
5	6	5.5	7	6	6	6	6	6	5	6	6	6	6.5	6	7	7	6	6	6	7	6	7	6
6	6	6	7	7	7	7	7	7	7	7	7	7	7	7	7	7	6	7	6	6	7	7	7
7	6	6	7	7	7	7	7	7	6	6.5	7	6	7	6	7	7	6	6	6	7	7	7	7
8	6	6	7	6	7	7	5	7	6	7	7	7	7	5	7	7	5.5	6	6.5	7	7	7	7
9	6	6	7	6	6	6	6	7	6	6	6	6	7	6	7	7	5	6	7	7	6	7	7
10	6	6	7	6	6	6	6	6.5	6	6	6	6	7	6	7	7	5	6	6	6	6	7	7
Avg Grade	86	85	83	85	80	83	81	83	78	82	83	85	83	77	78	82	81	76	83	85	81	86	78

As shown in the table, approximately 45% of the median scores were a 7, 51% were a 6, and 4% were a 5. Similarly, average grade percentages of students (which are heavily skewed depending on the number of students in the course since I do not grade on a curve) ranged from 76% to 86% (C to B letter grades). Evaluation scores are not shown for CPSC 491 taught in the Fall of 2012. This course is the Senior Design project course, which was broken into five sections—one section for each design team. Unfortunately, only one of the five groups had their evaluations "activated" on Zagweb. Thus, only four of the 16 students provided evaluations (which also explains some of the written comments in CPSC 492, the second semester of Senior Design). Instead, the scores for 499 (a course that accompanies 491) were used, which had seven (out of 16) respondents. In general, while the written evaluations vary across the courses above, I feel they are overall very positive and provide considerably more context surrounding the scores. The UNIV 099 course does not use numerical evaluations, only written evaluations, which are provided with the rest of the written course evaluations.

Analysis and Refinement. One of the most important lessons I've learned at Gonzaga is that teaching a course well requires teaching it many times (and improving upon how it is taught each time). I have been able to teach a number of courses multiple times, and each time I have made various refinements and adjustments. One of the most valuable instruments for me has been student feedback (both through speaking with students about the courses and through written course evaluations). While not as time consuming as developing a course for the first time, improving course organization, quizzes, homework assignments, exams, etc., takes a good deal of effort. However, I feel it is extremely important and also rewarding to engage in the work of refining and tuning a course as well as to experiment and adopt new methods into a course to improve student engagement and student learning. I always look forward to being able to further refine and develop my courses. In general, I feel that my teaching has improved considerably since joining Gonzaga. However, I feel that there is still considerable room for improvement, and especially in terms of developing and using more techniques to engage students and to foster increased student participation in lectures, and by continually working on the overall organization of my courses. One area I would like to particularly work on is providing the right "amount" of material in courses—I often have a tendency to provide too much information as opposed to holding back some material and taking the extra time to ensure students learn the essential topics really well (in contrast to many topics superficially).

2 Professional Development

My research interests are broadly in the areas of database systems, conceptual modeling, and data integration with applications motivated by data management challenges in the life sciences. Focusing on scientific data management provides many interesting research challenges and is an opportunity to make contributions within and outside of computer science. My recent research is primarily aimed at developing *scientific workflow technology* for modeling and automating scientific data analyses, *data provenance* techniques for tracking the lineage of data products, and approaches for representing *observational data semantics* to provide scientists with tools to perform large-scale data discovery and integration. Since joining Gonzaga, my professional development activities have included paper writing, grant writing, attending research meetings, and contributing to scientific working groups. More details on these activities are described below.

Research and Development Contributions. Over the last five years, my research has primarily been driven by collaborative projects funded through external research grants. This work has led to a number of contributions including: new modeling and optimization techniques for dataflow-based scientific workflow systems; new approaches for storing and retrieving (e.g., querying and browsing) data provenance generated from workflow systems; and a system called ObsDB for annotating and querying heterogeneous ecological data sets representing field observations. An overview of these areas of research are discussed in more detail in the invited journal paper listed below (paper 29), which is also included in the professional development materials. The worklfow modeling and optimization techniques as well as the provenance approaches were implemented and distributed through the Kepler Scientific Workflow System (http://www.kepler-project.org). Kepler is a collaborative project that distributes an maintains an open-source and extensible scientific workflow system used within a number of other research efforts. One of the larger projects that uses Kepler as well as the provenance extensions we developed is the CAMERA project (http://camera.calit2.net/). Within CAMERA, Kepler workflows perform various analyses within their web portal, which is used by a large number of microbial and ecological scientists. The modeling and optimization approaches we developed have also been used within other research projects including metagonomic studies at UC Davis and a data pipeline system developed at the Stanford Linear Accelerator Center (SLAC). The work on ObsDB provided prototypes that have been incorporated into other systems, including the Morpho metadata editor and Metacat data repository system, which is distributed through the National Center for Ecological Analysis and Synthesis (NCEAS). Both systems are deployed within a number of ecological research sites, including the Long-Term Ecological Research (LTER) Network. More recently, my work has focused on using formal logic languages and frameworks to help taxonomists and systematists align, reason about, and integrate biological taxonomies. Through the Euler project, we are working closing with a taxonomist at ASU on developing software tools for the broader taxonomic community. A recent workshop paper (paper 27, also included in the professional development materials) describes the prototypes we have developed in more detail. This work has been especially interesting for me because it combines more theoretical computer science topics (e.g., automated theorem proving) with new tools being developed by the computer science community to efficiently solve constraint-based reasoning problems (in particular, using so-called "answer-set programming" systems). Finally, I have been involved in a number of smaller projects while at GU, which have helped to both fund undergraduate student research and allow me to explore other research areas including data analytics and data mining. My hope is that some of this work will also lead to larger externally funded research projects in the future.

Refereed Research Papers. The following is a chronological (by year) list of the peer-reviewed articles that were either published or accepted for publication since I joined Gonzaga. Each publication includes brief descriptions of the role I played in its production and the publication venue. Copies of the following papers are also provided within this packet. Please note that in the field of computer science, conference proceedings are published as books, and conference publications are generally considered the same as journal publications in other fields. This is especially true in the subfield of database management, where papers in many conferences generally have a higher citation count and considerably lower acceptance rates than those in similar journals. In many cases, workshop proceedings (typically organized within larger conferences) are also published as books (e.g., through Springer's Lecture Notes in Computer Science series).

1. M. Anand, **S. Bowers**, B. Ludäscher. 2009. A navigation model for exploring scientific workflow provenance graphs. *Proceedings of the 4th International Workshop on Workflows in Support of Large-Scale Science* (WORKS), ACM Press, ISBN: 978-1-60558-717-2.

   – I supervised Manish Anand (a PhD student at UC Davis) with this work, developed (with help from Manish) the navigation model, and wrote the paper. WORKS is part of the Supercomputing conference, and is a yearly workshop devoted to issues surrounding scientific workflow technology.

2. D. Zinn, **S. Bowers**, T. McPhillips, B. Ludäscher. 2009. Scientific workflow design with data assembly lines. *Proceedings of the 4th International Workshop on Workflows in Support of Large-Scale Science* (WORKS), ACM Press, ISBN: 978-1-60558-717-2.

   – I helped supervise Daniel Zinn (a PhD student at UC Davis), participated in the development of the formal worklfow model presented, and did approximately half of the writing of the paper.

3. D. Thau, **S. Bowers**, B. Ludäscher. 2009. Merging sets of taxonomically organized data using concept mappings under uncertainty. *Proceedings of the 8th International Conference on Ontologies, DataBases, and Applications* (ODBASE), LNCS 5871, pp. 1103–1120, ISBN 978-3-642-05150-0.

   – I helped supervise Dave Thau (PhD student at UC Davis), made contributions to the approach presented, and wrote and edited the majority of the sections of the paper together with Dave. ODBASE is one of the major conceptual modeling conferences.

4. D. Zinn, **S. Bowers**, S. Köhler, B. Ludäscher. 2010. Parallelizing XML data-streaming workflows via MapReduce. *Journal of Computer and System Sciences* (JCSS), 76(6):447–463.

   – Similar to (2) above. I did a considerable amount of editing and rewriting of the second version of the paper for final acceptance. JCSS is a top-tier computer science journal, primarily organized around special topics (in this case, scientific workflow systems).

5. **S. Bowers**, J. Madin, M. Schildhauer. 2010. Owlifier: Creating OWL-DL ontologies from simple spreadsheet-based knowledge descriptions. *Ecological Informatics*, 5(1):19–25.

   – I implemented the version of the system described in the paper as well as wrote the paper. This journal is focused on applications of computer science (data management issues and computational approaches) to ecology.

6. M. Anand, **S. Bowers**, B. Ludäscher. 2010. Techniques for efficiently querying scientific workflow provenance graphs. *Proceedings of the ACM International Conference on Extending Database Technology* (EDBT), pp. 287-298, ISBN 978-1-60558-945-9.

   – I developed the approaches described in the paper and did the majority of writing. Manish implemented the system and performed the benchmarks. EDBT is a top-tier database conference.

7. **S. Bowers**, J. Kudo, H. Cao, M. Schildhauer. 2010. ObsDB: A system for uniformly storing and querying heterogeneous observational data, *Proceedings of the IEEE International Conference on e-Science*, pp. 261–268, IEEE Computer Society, ISBN 978-1-4244-8957-2.

   – I supervised Jay Kudo (a Gonzaga undergraduate student) and wrote the paper. Jay implemented the system described in the paper. The e-Science conference is one of the larger conferences focused on applications of computer science to scientific data and computation problems.

8. M. Anand, **S. Bowers**, B. Ludäscher. 2010. Provenance browser: Displaying and querying scientific workflow provenance graphs. *Proceedings of the IEEE International Conference on Data Engineering* (ICDE), pp. 1201-1204, IEEE Computer Society, ISBN 978-1-4244-5444-0.

   – I wrote this demo paper and the initial implementation of the system described. Manish implemented additional capabilities into the system and performed benchmarks. ICDE is a top-tier database conference.

9. D. Zinn, **S. Bowers**, B. Ludäscher. XML-based computation for scientific workflows. *Proceedings of the IEEE International Conference on Data Engineering* (ICDE), pp. 812-815, IEEE Computer Society, ISBN 978-1-4244-5444-0.

   – My role was similar to paper (2) in this demo paper. I primarily edited the paper.

10. P. Missier, C. Goble, S. Dey, A. Sarkar, B. Shrestha, B. Ludäscher, **S. Bowers**, I. Altintas, M. Anand. 2010. Linking multiple workflow provenance traces for interoperable collaborative science. *Proceedings of the ACM International Workshop on Workflows in Support of Large-Scale Science* (WORKS), ACM Press, ISBN 978-1-60558-717-2.

    – This work was a joint effort as part of the DataNet Provenance Working Group summer internship program. Paolo Missier and I did the majority of the writing with help from Bertram Ludäscher. Dey, Sarkar, and Shrestha (graduate students) developed the implementation described in the paper and were supervised by Paolo, Bertram, and myself.

11. M. Anand, **S. Bowers**, I. Altintas, B. Ludäscher. 2010. Approaches for exploring and querying scientific workflow provenance graphs. *Proceedings of the International Provenance and Annotation Workshop* (IPAW), Revised Selected Papers, LNCS 6378, pp. 17-26, ISBN 978-3-642-17818-4.

    – My role in this paper was similar to (1). I developed the query language, which was implemented and benchmarked by Manish. I wrote the majority of the paper. IPAW is one of two workshops dedicated to issues in data and workflow provenance.

12. I. Altintas, M. Anand, D. Crawl, **S. Bowers**, A. Belloum, P. Missier, B. Ludäscher, C. Goble, P. Sloot. 2010. Understanding collaborative studies through interoperable workflow provenance. *Proc. of the International Provenance and Annotation Workshop* (IPAW), Revised Selected Papers, LNCS 6378, pp. 42-58, ISBN 978-3-642-17818-4.

– This paper was primarily initiated by Ilkay Altintas (a PhD student at the time). Together with Manish, the three of us equally contributed to the writing and editing of the paper. Manish also implemented under my guidance much of the underlying system experiments.

13. D. Thau, **S. Bowers**, B. Ludäscher. 2010. Towards best-effort merge of taxonomically organized data. *Proceedings of the International Workshop on New Trends in Information Integration* (NTII), ICDE Workshops, pp. 151–154, IEEE Press, ISBN: 978-1-4244-6522-4.

    – My role in this short paper was similar to paper (3). NTII was a workshop held in conjunction with ICDE.

14. **S. Bowers**, H. Cao, M. Schildhauer, M. Jones, B. Leinfelder, M. O'Brien. 2010 A semantic annotation framework for retrieving and analyzing observational datasets. *Proceedings of the Workshop on Exploiting Semantic Annotations in Information Retrieval* (ESAIR), pp. 31-32, ACM Press, ISBN 978-1-4503-0372-9.

    – This short paper was written by myself together with Huiping Cao.

15. I. Altintas, M. Anand, T. Vuong, **S. Bowers**, B. Ludäscher, P. Sloot. 2011. A Data Model for Analyzing User Collaborations in Workflow-Driven eScience. *International Journal of Computers and Their Applications* (IJCA) 18(3):160–179.

    – My role in this paper was similar to paper (12). This was an extended version of that work.

16. H. Cao, **S. Bowers**, M. Schildhauer. 2011. Approaches for Semantically Annotating and Discovering Scientific Observational Data. *Proceedings of the International Conference on Database and Expert Systems Applications* (DEXA), LNCS vol. 6860, pp. 526–541, ISBN 978-3-642-23087-5.

    – I supervised Huiping Cao (a postdoc at UC Santa Barbara) together with Mark Schildhauer. This was primarily collaborative work with Huiping, who implemented the approaches and benchmarking described in the paper under my guidance. I did a majority of the writing with help from Huiping. DEXA is one of the major conceptual modeling conferences.

17. M. Agun, **S. Bowers**. 2011. Approaches for Implementing Persistent Queues within Data-Intensive Scientic Workows. *Proceedings of the IEEE World Congress on Services* (SERVICES), pp. 200-207, IEEE Computer Society, ISBN: 978-0-7695-4461-8.

    – I supervised Michael Agun, who was a Gonzaga undergraduate student. I wrote the paper and developed many of the ideas underlying the paper with Michael, who implemented the approaches and performed benchmarking of the approaches. This is one of the larger conferences focused on software services, where our paper was part of a session on workflow technology.

18. L. Dou, D. Zinn, T. McPhillips, S. Köhler, S. Riddle, **S. Bowers**, B. Ludäscher. 2011. Scientific Workflow Design 2.0: Demonstrating Streaming Data Collections in Kepler. *Proceedings of the International Conference on Data Engineering* (ICDE), pp. 1296-1299, IEEE Computer Society, ISBN 978-1-4244-8958-9.

    – This demo paper was based on work I did with Tim McPhillips on extending the stream-based computational model in Kepler. I was responsible for many of the underlying ideas and formalisms (together with Tim), which were implemented primarily by Lei Dou and Sven Köhler.

19. W. Saunders, **S. Bowers**, M. O'Brien. 2011. Protégé Extensions for Scientist-Oriented Modeling of Observation and Measurement Semantics. *Proceedings of the 8th International Workshop on OWL: Experiences and Directions* (OWLED), CEUR Workshop Proceedings, vol. 796.

    – I supervised Wesley Saunders (an undergraduate student at Gonzaga). I wrote the paper, which describes an extension to a popular ontology editor for simplifying the creation of observational ontologies. Wesley and I worked together to define the extensions, which he implemented. Margaret O'Brien is an ecologist who tested and used the system to implement an ontology focused on ocean studies at the Santa Barbara Coastal LTER site. OWLED is a popular workshop within the semantic web community focused on the OWL description-logic ontology language.

20. B. Leinfelder, **S. Bowers**, M. O'Brien, M. Jones, M. Schildhauer. 2011. Using Semantic Metadata for Discovery and Integration of Heterogeneous Ecological Data. *Proceedings of the Environmental Information Management Conference* (EIM), pp. 92–97 (doi:10.5060/D2NC5Z4X).

    – I wrote and edited a large portion of the paper, which describes a semantic annotation editor created within a popular metadata tool used by ecologists. Ben Leinfelder implemented the tool, whose interface was designed together by each of the co-authors. EIM is a conference focused on tools and techniques for ecological information management.

21. H. Cao, **S. Bowers**, M. P. Schildhauer. 2012. Database Support for Enabling Data-Discovery Queries over Semantically-Annotated Observational Data. *Transactions on Large-Scale Data- and Knowledge-Centered Systems* vol. 6, pp. 198-228, ISBN: 9783642341786.

    – My role on this paper was similar to paper (16). This is an extended version of that paper. TLDKS is a newer Springer journal that spans the conceptual modeling and data management communities.

22. M. K. Anand, **S. Bowers**, B. Ludäscher. 2012. Database Support for Exploring Scientific Workflow Provenance Graphs. *Proceedings of the 24th International Conference on Statistical and Scientific Database Management (SSDBM)*, LNCS vol. 7338, pp. 343-360, ISBN: 9783642312342.

    – My role on this paper was similar to papers 11, 6, and 1. SSDBM is one of the major database conferences, with a focus on scientific applications.

23. **S. Bowers**, T. McPhillips, and B. Ludäscher. 2012. Declarative Rules for Inferring Fine-Grained Data Provenance from Scientific Workflow Execution Traces. *Proceedings of the International Provenance and Annotation Workshop (IPAW)*, Revised Selected Papers, LNCS vol. 7525, pp. 82–96, ISBN 978-3-642-34221-9.

    – I developed the approaches, implementation, and evaluation described in this paper and also wrote the paper. The work stemmed from conversations I had with Tim McPhillips, who also discussed the ideas with me as they evolved.

24. S. Dey, S. Köhler, **S. Bowers**, B. Ludäscher. Datalog as a Lingua Franca for Provenance Querying and Reasoning. *Proceedings of the 4th USENIX Workshop on the Theory and Practice of Provenance (TaPP)*, 2012.

    – This work stemmed from another DataNet summer internship program and was based on our prior work on provenance. Saumen Dey led the writing and implementation, and I helped edit the paper.

25. P. Missier, B. Ludäsher, S. Dey, M. Wang, T. McPhillips, **S. Bowers**, M. Agun, and I. Altintas. Golden Trail: Retrieving the Data History that Matters from a Comprehensive Provenance Repository. *International Journal of Digital Curation* 7(1):139–150, 2012.

    – This work also stemmed from the DataNet summer internship program, which involved Michael Agun (an undergraduate student from Gonzaga). I was involved in student supervision, and also helped write and edit the paper.

26. **S. Bowers**, R. Englin, C. Fonseca, P. Jewell, L. Joplin, P. Mosca, T. Pacheco, J. Troxel, and T. Weeks. 2013. Lightweight Ontology-Based Tools for Managing Heterogeneous Observational Data. *Proceedings of the International Workshop on Semantics for Biodiversity* (S4BioDiv), CEUR Workshop Proceedings, vol. 979, ISSN 1613-0073.

    – I supervised each of the co-authors on the paper (all of whom were undergraduate students at Gonzaga). The work described in the paper was carried out (under my supervision) over the summer of 2012 by the students. I wrote the paper.

27. M. Chen, S. Yu, N. Franz, **S. Bowers**, and B. Ludáscher. 2013. Euler/X: A Toolkit for Logic-based Taxonomy Integration. *Proceedings of the International Workshop on Functional and and (Constraint) Logic Programming*. (Available at arXiv.org: http://arxiv.org/abs/1402.1992).

    – This system paper describes extensions to the implementation created by Dave Thau. I primarily played a supervisory role (as a project Co-PI) to Mingmin Chen and Shizuo Yu (PhD students at UC Davis), and also helped in the writing and editing of the paper.

28. **S. Bowers**, K. Yerion. 2013. Programming Personal Robots within an Introductory Computer Science Course for Engineering Majors. *Journal of Computing Sciences in Colleges*, 29(1):133–139.

– This paper describes the robotics approach and results used in the Fall of 2012. The paper was jointly written with Kathie Yerion.

29. R. Walls, J. Deck, R. Guralnick, S. Baskauf, R. Beaman, S. Blum, **S. Bowers**, P. Buttigieg, N. Davies, D. Endresen, M. Gandolfo, R. Hanner, A. Janning, L. Krishtalka, A. Matsunaga, P. Midford, N. Morrison, E. Tuama, M. Schildhauer, B. Smith, B. Stucky, A. Thomer, J. Wieczorek, J. Whitacre, J. Wooley. 2014. Semantics in Support of Biodiversity Knowledge Discovery: An Introduction to the Biological Collections Ontology and Related Ontologies. PLoS One 9(3):e89606, doi: 10.1371/jour-nal.pone.0089606.

    – This paper came out of a workshop I attended on biodiversity semantics. I played a minor role in the paper, providing mainly comments, suggestions, and edits on the first submission. I also attended the workshop where many of the ideas presented originated. PLoS One is an international peer-reviewed, open access journal. It is primarily focused on science and medicine research, although it also publishes papers in bioinformatics. The journal is very well known, but someone controversial because it is one of the first mainstream open-access journals.

30. S. Dey, S. Köhler, **S. Bowers**, B. Ludäscher. 2014. Computing Location-Based Lineage from Workflow Specifications to Optimize Provenance Queries. 2014. *Proceedings of the International Provenance and Annotation Workshop (IPAW)*, LNCS, to appear.

    – This paper describes a new approach for statically analyzing workflow specifications to pre-compute lineage information. This work was primarily done by Saumen Dey (a PhD student at UC Davis). However, I made significant contributions to the writing of the paper, the research, and presented the paper at the IPAW workshop in Cologne Germany.

31. M. Chen, S. Yu, P. Kianmajd, N. Franz, **S. Bowers**, B. Ludäscher. 2014. Provenance for Explaining Taxonomy Alignments (Extended Abstract). *Proceedings of the USENIX Workshop on the Theory and Practice of Provenance*, LNCS, to appear.

    – This paper describes joint work performed on the Euler project. This is a short paper that describes our approach for combining white-box and black-box provenance. I contributed to the underlying research within Euler.

32. T. Song, S. Dey, **S. Bowers**, B. Ludäscher. 2014. Improving Workflow Design Using Abstract Provenance Graphs (Extended Abstract). *Proceedings of the USENIX Workshop on the Theory and Practice of Provenance*, LNCS, to appear.

    – This papers describes the work of Tianhong Song's PhD proposal (at UC Davis), and extends prior work on using workflow graph specifications (based on COMAD) to infer provenance relationships. This is also a short paper. My role was primarily as an advisor to Tianhong and we had multiple detailed discussions of the work.

33. M. Chen, S. Yu, N. Franz, **S. Bowers**, B. Ludäscher. 2014. A Hybrid Diagnosis Approach Combining Black-Box and White-Box Reasoning. *Proceedings of the International Web Rule Symposium (RuleML)*, LNCS vol. 8620, pp. 127–141.

– This paper goes into more detail about the provenance approaches used within Euler. This work was primarily led by Mingmin Chen, however, the project PIs (Bertram Ludäscher and myself) made contributions both to the research and to the writing and editing of the paper. I helped edit multiple drafts and also with the submitted version of the work.

34. N. Franz, M. Chen, S. Yu, P. Kianma, **S. Bowers**, B. Ludäscher. 2014. Reasoning over Taxonomic Change: Exploring Alignments for the *Perelleschus* Use Case. *PLoS One*, In press.

    – This paper was recently accepted for publication in PLoS One. The paper was written by Nico Franz and uses the Euler framework to analyze a specific taxonomic use case. I made contributions to the tools used within the study, but did not contribute to writing the paper. Nico is actively involved in the Euler project as a taxonomist (domain expert), including our weekly research and development video conference calls.

Additional Publications. In addition to the above, I also wrote the following invited journal paper.

35. **S. Bowers**. Scientific Workflow, Provenance, and Data Modeling Challenges and Approaches. *Journal on Data Semantics*, 1(1):19-30, 2012, ISSN:18612032.

Finally, I served as a co-editor for the 2011 and 2012 SSDBM conferences. This work involved managing the process of collecting camera-ready copies of papers, ensuring all papers met the LNCS style guidelines, editing some papers, putting together the camera-ready copy of the book (front matter, table of contents, etc.), and working with Springer to correct any errors, etc. This work resulted in the following two books published by Springer in the Lecture Notes for Computer Science (LNCS) series.

26. J. Cushing, J. French, **S. Bowers** (Editors). *Proceedings of the 23th International Conference on Scientific and Statistical Database Management (SSDBM)*, LNCS vol. 6809, Springer, 2011, ISBN:3642223508.

27. A. Ailamaki, **S. Bowers** (Editors). *Proceedings of the 24th International Conference on Scientific and Statistical Database Management (SSDBM)*, LNCS vol. 7338, Springer, 2012, ISBN:9783642312342.

Publication Metrics. Although somewhat controversial, an often-used measurement of a researcher's productivity and impact regarding publications is the h-index (similar to an "impact factor"), which is the largest number n of a researcher's papers that each have at least n citations. Google Scholar[3] gives my h-index score as 29 (i.e., 29 papers have at least 29 citations). In addition to the h-index, Google Scholar also reports the following metrics regarding my publications: my i10-index is 60 (i.e., 60 papers have at least 10 or more citations); I have eight papers with over 100 citations each (the top three have 176, 157, and 146 citations); I have 2,999 total citations across all publications; and since 2009 (when I joined GU) I have an h-index of 25, an i10-index of 49, and 2,116 citations.

[3]See http://scholar.google.com/citations?hl=en&user=-_n_uGcAAAAJ

External Grants. The following are external grants that I've worked on since joining Gonzaga. Note that the first two were awarded prior to my joining Gonzaga, but the entire subcontract amount was spent at GU. The subcontracts were spent primarily on student research support, followed by faculty summer support, travel (for conference and workshop presentations as well as project meetings), and equipment (student computers).

1. **A Community-driven Scientific Observations Network to achieve Interoperability of Environmental and Ecological Data** (*Co-PI*, GU Subcontract), National Science Foundation, #0753144, total award amount $750,000 (8/2008–7/2011), with National Center for Ecological Analysis and Synthesis (UCSB), Arizona State University, Monterey Bay Aquarium Research Institute, Rensselaer Polytechnic Institute.

2. **Semantic Enhancements for Ecological Data Management**, (*Co-PI*, GU Subcontract), National Science Foundation, #0743429, total award amount $599,999 (8/2008–7/2011), with National Center for Ecological Analysis and Synthesis and Marine Science Institute (UCSB).

3. **A Dynamic Web-Based Application for Exploratory Analysis of De-Identified Spokane Community Clinical Data Repository (SCCDR)** (*PI*), Institute for Systems Medicine, GU subcontract amount $21,936 (11/2010–9/2011).

4. **III:Small: (Euler) A Logic-Based, Provenance-Aware System for Merging Scientific Data under Context and Classification Constraints** (*Co-PI*, GU Subcontract), National Science Foundation, total award amount $439,000, with UC Davis (10/1/2011–9/30/2014).

5. **Development of the ProvEx Provenance Browser**, GU Subcontract for $2,009 from University of New Mexico as part of the DataONE National Science Foundation Award (www.dataone.org). This money was used to pay an undergraduate student to work during the academic year.

Internal Grants. The following are internal grants (funded through GU) that I have received over the last five years.

6. **Exploring Efficient Buffering Strategies for Data-Intensive Scientific Workflows** (*PI*), Gonzaga University Faculty Research Award, $1,500, (9/2010–5/2011). This grant funded Michael Agun to work on workflow optimization research for a part of the academic year. The work led to a publication (paper 17 above) as well as a Spokane Intercollegiate Research Council (SIRC) presentation.

7. **An Extensible Framework for Generating Interactive Data Visualizations on the Web** (*PI*), Gonzaga University Faculty Research Award, $1,500, (9/2011–5/2012). This grant funded an undergraduate student for part of the academic year. This work extended and generalized the SCCDR visualizations deployed on the web.

8. **Gonzaga University KEEN Proposal for Developing and Adopting Robotics Modules into CPSC 121 (Computer Science I)** (*PI*), GU internal funding through the KEEN Foundation Grant, $10,494 (1/1/2011–12/1/2013). This grant was through the larger KEEN award at GU.

9. **Data-Intensive Approaches for Multi-Dimension, Large-Scale Validation of Rate My Professor Data** (*PI*), Gonzaga University Faculty Research Award, $1,500, (9/2012–5/2013). This grant funded an undergraduate student for part of the academic year. The results of this work led to a SIRC presentation and we plan to write a paper on the findings in the future.

Pending Proposals. The following proposal was submitted in in August of 2014 (this summer), and is currently under review at the National Science Foundation.

10. **Collaborative Research: ABI Development: An Integrated, Scalable Service for Resolving Taxonomic Concept Provenance** (*Co-PI*, GU Subcontract), National Science Foundation, $477,113, with Arizona State University, under review.

Funded Student Research. The following is a short description of the undergraduate students' work that I have funded through the above research projects over the last five years.

- *Michael Agun.* Michael worked with me over a two-year period on optimization approaches for dataflow process networks. Michael joined UC Santa Barbara as a PhD student in computer science in the Fall of 2013.
- *Nathan Hollister.* Nathan worked for one semester on extending the Owlifier Java application. Nathan has graduated and is working as a software developer at Boeing.
- *Jay Kudo.* Jay Kudo worked with me during the summer of 2010 on the initial version of ObsDB. Jay has since graduated from GU and is working at Accenture.
- *Wesley Saunders.* Wes worked with me during the summer of 2010 and the 2011 academic year developing extensions for the Protégé ontology editor. Wes has since graduated from GU and is now working at Boeing.
- *Thomas Scheffer.* Thomas worked with me over a summer as a software developer on the Semtools project. He focused on extending and refactoring portions of the ObsDB system developed by Jay Kudo. Thomas graduated from GU last year.
- *Douglas Coulson.* Doug worked with me over two academic years and one summer on SCCDR and a GU research council grant. Doug graduated from GU last Spring.
- *Josie Hunter.* Josie worked with me during the academic year and over a summer on the SONet project. She developed approaches for analyzing attribute schemes in large data repositories. I am hoping to submit a grant proposal to the NSF ABI program next summer based on this work. Josie finished her graduate studies at Oregon State University and now works as a professional software developer.
- *Paul Jewell.* Paul worked over one academic year building a framework with myself and Dr. Chris LaSota (Physics) to support the evolution of cellular automata systems. After this work, Paul worked for one summer and an academic year on the Euler project. Paul graduated in the Spring and is now working at Amazon.

- *Jacob Troxel and Tyler Weeks.* Jacob worked for one semester on the SONet project (developing semantic annotation support for ObsDB). He continued this work with Tyler over the summer of 2012. Jacob also worked on the ProvEx project during the previous academic year. Jacob is still a student at GU and Tyler graduated two years ago and is working as a software developer at Moss Adams in Spokane.

- *Patrick Mosca, Lauren Joplin, and Carlos Fonseca.* Patrick, Lauren, and Carlos worked over a summer developing "markdown" languages and support for representing ontologies within the ObsDB system. This work was funded through the SONet and Semtools NSF grants. Carlos was an exchange student from Brazil. Patrick also worked with me during the academic year on a GU Research Council grant. Patrick and Lauren have both graduated from GU. Patrick is a software developer at Apple, Inc. and Lauren is a software developer at F5 Networks in Seattle, WA.

- *Tyler Pacheco and Riley Englin.* Tyler and Riley worked during the summer of 2012 on the SONet project, developing query and analysis support within the ObsDB system. Both Tyler and Riley graduated last year. Riley is pursuing a graduate degree from Eastern Washington University and Tyler is working as a software developer at General Dynamics.

- *Corwin Bryan.* Corwin began working for me two summers ago on the Euler project. He continued working with me this the following year, and took an internship last summer at the Pacific Northwest National Lab in Richland, WA. Corwin is still a student at GU.

- *Dylan Kaai.* Dylan continued work on the Euler project, where he extended Corwin's work on a specialized reasoning algorithm for taxonomic alignment and reasoning. Dylan also implemented techniques for provenance and generating all possible worlds from an alignment. He and Corwin won the best poster award at CCSC-NW this year on their work together.

Meetings and Talks. Over the last five years I participated in a number of external meetings related to professional development. The following lists the major meetings and talks I attended.

- *SONet Community Workshop (Santa Barbara, CA), Oct. 2009.* As part of my work on the SONet project, I helped organize a community workshop at UC Santa Barbara. This meeting hosted 20 participants from various institutions (including the Long Term Ecological Research network, NASA, OGC, the iPlant consortium, IRI, the Cornell Lab of Ornithology, and the Tetherless World Project at RPI). The goals of the workshop were to discuss approaches for modeling observational data, develop use case requirements for an interoperable observation-based data framework, and define possible demonstration projects for SONet. I also presented work on the OBOE observational model at the workshop, and lead breakout groups on interoperability issues surrounding core observation models.

- *Workshop on Workflows in Support of Large-Scale Science (Portland, OR), Nov. 2009.* I attended WORKS, which was held at Supercomputing 2009. Manish Anand (a former PhD student) presented our paper at the workshop on provenance.

- *DataONE Working Group Planning Meeting (Davis, CA), Feb. 2010.* I participated in a planning meeting hosted at UC Davis to organize a provenance working group within the NSF DataONE project. In this meeting, I helped draft a working group proposal (which was

accepted and funded through DataONE). Working groups participate in DataONE annual meetings as well as regular video conferences.

- *SONet Plant Trait Modeling Workshop (Santa Barbara, CA), May 2010.* I helped organize and participated in this SONet workshop held at UC Santa Barbara, which involved members of different projects focused on leveraging plant-trait data. I gave a talk at this meeting on recent updates to the OBOE observational model.

- *Data Conservancy Ontology Workshop (Ithaca, NY), Jul. 2010.* This meeting was held at the Cornell Ornithology Lab as part of the Data Conservancy project (funded through the NSF/DataNET program). I gave a talk on the OBOE observational model. The purpose of the meeting was to explore collaborations between the Data Conservancy project and the SONet project.

- *DataONE All Hands Meeting (Albuquerque, NM), Oct. 2010.* As a member of the Provenance Working Group, I participated in this meeting and also give a talk on provenance support and applications within the Kepler Scientific Workflow System.

- *e-Science (Brisbane, Australia), Dec. 2010.* I presented the ObsDB paper at this week long technical conference.

- *Eastern Washington University, May 2010.* I gave an invited talk on provenance in scientific workflows at the EWU Computer Science Department as part of their "Silver Anniversary" Colloquium Series.

- *JWG-ODS Workshop (Santa Barbara, CA), Apr. 2011.* I helped organize and participated in this workshop at UC Santa Barbara which involved members of different NSF projects focused on leveraging observational ontologies to support discovery and integration of scientific data. This meeting was a follow-up to previous SONet workshops. I gave a talk at this meeting on recent updates to the OBOE observational model and applications of OBOE.

- *OWLED Workshop (San Francisco, CA), Jun. 2011.* I presented our paper on Protégé extensions at the workshop, which was part of the Semantic Technologies Conference.

- *DataONE Working Group Meeting (Davis, CA), Jun. 2011.* I attend this DataONE provenance working group meeting hosted at UC Davis. The goal of the meeting was to develop a set of initial use cases and a general scientific-workflow provenance model for the Provenance Working Group. We also outlined a set of deliverables for the DataONE summer internship program. This meeting consisted of approximately 15 participants from various institutions within the U.S. and the U.K.

- *IEEE SERVICES Conference (Washington D.C.), Jul. 2011.* I presented our paper on workflow optimization techniques at the Scientific Workflow session of the conference.

- *DataONE All Hands Meeting (Albuquerque, NM), Oct. 2011.* I attended the Provenance Working Group meeting at the DataONE All Hands Meeting. The working group met for two days to further refine use cases and the provenance model.

- *CCSC-NW Conference (Richland, WA), Oct. 2011.* I attended the Northwest Region Consortium for Computing Sciences in Colleges Conference, which primarily focused on talks and tutorials relating to computer science education.

- *Principles of Provenance Seminar (Dagstuhl, Germany), Feb. 2012).* I attended this week-long computer-science research seminar at the Leibniz Center for Informatics, located in Dagstuhl, Germany. The Dagstuhl seminars are well-known within computer science as one of the premier venues for informatics where researchers come together to exchange knowledge and to discuss research findings. Attendance at all Dagstuhl Seminars are by invitation only. I contributed at the seminar to one of three invited tutorials. Our tutorial was on provenance technology within scientific workflows, and a short write-up of the tutorial was published as part of the Dagstuhl report series:

  > B. Ludäscher, **S. Bowers**, P. Missier. Tutorial: Provenance in Scientific Workflows. In J. Cheney, A. Finkelstein, B. Ludäscher, S. Vansummeren (editors), *Principles of Provenance (Dagstuhl Seminar 12091)*, Dagstuhl Reports 2(2):84-113, 2012. (http://drops.dagstuhl.de/opus/volltexte/2012/3507/)

- *Semantics of Biodiversity Workshop (Lawrence, KS), May 2012).* I was invited to attend this workshop at Kansas University. The goal of this workshop was to identify potential ontologies, ontology frameworks, and common areas of development in terms of semantic approaches for biodiversity informatics. This workshop was led by John Deck (UC Berkeley) and Barry Smith (SUNY Buffalo).

- *DataONE Provenance Working Group Meeting (Santa Barbara, CA), Jun. 2012).* I attended this research meeting at UC Santa Barbara as a member of the Provenance Working Group. The meeting focused on techniques and approaches for leveraging and extending provenance frameworks to support ecological and biodiversity applications. This work led to the small grant to develop a provenance browser as part of the DataONE data management system.

- *International Provenance and Annotation Workshop (Santa Barbara, CA), Jun. 2012).* I presented a paper on provenance inference rules at this workshop. The DataONE provenance working group meeting was co-located with this workshop.

- *International Conference on Scientific and Statistical Databases (Chania, Crete, Greece), Jun. 2012).* I presented our paper on mixing navigation, querying, and browsing provenance information at this conference.

- *St. Louis University KEEN Faculty Development Workshop (St. Louis, MI), Aug. 2012.* I attended this week-long workshop at SLU on active and problem-based learning approaches. My attendance at the workshop was sponsored by the KEEN GU grant.

- *DataONE All Hands Meeting (Albuquerque, NM), Sep. 2012).* I attended this meeting as a member of the Provenance Working Group. This was largely a working meeting that was a follow up to the meeting in the summer. I also attended the semantics working group at the DataONE all hands meeting.

- *Euler All Hands Meeting (Davis, CA), Sept. 2012).* I attended this meeting at UC Davis, which was a working meeting for the NSF-funded Euler research project.

- *CCSC-NW Conference, (Evergreen, WA), Oct. 2012* . I attended the Northwest Region Consortium for Computing Sciences in Colleges Conference. I was also the student poster chair for the conference.

- *KEEN Winter Conference (Tempe, AZ), Jan. 2013*. I attended the KEEN winter conference (all hands meeting) at ASU, which consisted of a number of speakers discussing KEEN education topics, breakout meetings, and a tour of the College of Technology and Innovation at ASU.

- *Extended Semantic Web Conference (Montpellier, France), May 2013*. I attended this week-long computer-science research conference, where I presented a paper at the Semantics for Biodiversity workshop on ObsDB.

- *International Provenance and Annotation Workshop (Cologne, Germany), June 2014*. I attended this workshop where I presented our paper on location-based lineage in scientific workflows.

In addition to the above, I participated in various weekly phone and video conference meetings as part of the NSF projects I have been involved in since while at GU (SONet, Semtools, Euler). I also was a mentor for the NSF DataONE summer internship program (funded through the DataONE project). This involved weekly meetings for two months over the summer for two years. Finally, I served as an NSF Panelist for the National Science Foundation in March, 2014. As a member of the review panel, I reviewed a number of proposals for the NSF IIS Medium grant program.

3 Advising

I have been an academic advisor for computer science majors for three of my five years at GU. I currently have over 40 advisees—although this number includes advisees for a faculty member on sabbatical this year. In general, I try to apply the same approach to advisees as with office hours, i.e., by encouraging my advisees to feel that they are more than welcome to stop by and discuss any academic issues, opportunities for internships, to discuss career plans, etc. In teaching the UNIV 099 course, I have had the opportunity to think much more about advising and how to help students that are struggling academically. I feel I have improved in this area, i.e., identifying issues with students who are struggling, but I still have a long way to go.

For each advisee I maintain an electronic advising form (used by all CS faculty) and fill it in each semester after they have taken their classes and received their final grades. I meet with new freshman advisees three times in the Fall Semester: at the initial meeting prior to when classes begin, a few weeks before midterm grades (to individually introduce myself and to discuss the advisee process and class registration), and then again when I give them their midterm grades and help them with their schedules. During their second semester I meet with them twice again: just prior to midterm grades to check in with them and to get them thinking about their schedules, and then to give them their midterm grades and discuss schedules. During their sophomore year, I typically only officially meet with advisees to give them their midterm grades. During the Fall semester I encourage my advisees to attend the Engineering Career Fair, and in particular, to start thinking about internship opportunities. For most students, internships are not until their junior year, however, it is important for them to begin thinking about doing internships early, especially given their importance when they apply for jobs in their senior year. In each meeting I typically

give the students a printed copy of the filled-in advising form. Since the recent changes in the CS curriculum (which provides 12 general credits), I have had many more students interested in pursuing minors. For these students, I often meet with them to go over a rough "four year plan". While this takes a fair bit of time, I have found this really helps students to see the bigger picture of how their classes will fit together and that doing a minor is feasible.

For the past four years I have served as the faculty advisor for Upsilon Pi Epsilon (UPE), which is the GU computer science honor society. The GU chapter of UPE was formed in 2010, and has been active since. The main activities supported through UPE have been to provide tutoring services for computer-science students, to host guest speakers from industry, and together with the ACM student chapter (now the computer science club) to organize a series of practice programming competitions at GU. Over the past five years we have sent 15 teams (3 students per team) to the yearly ACM programming competition. Over the last two years, the teams have done very well, placing 2nd in the Eastern region of the Pacific Northwest competition two years in a row. This past year, the GU chapters of UPE and ACM have been jointly run through the "computer science club", and so this last year I have served as the de facto faculty advisor for ACM as well. As a faculty advisor for UPE and the computer science club, I feel my primary role is to simply be available for the students to help organize and attend activities. This more hands-off approach is largely based on my own experiences as an undergraduate student actively involved with the ACM—the students get the most out of these clubs when they drive them, and not when they are driven by the faculty advisor. One exception is the ACM programming competition, which I have taken the lead on organizing for the last few years.

In addition to "official" academic advising duties, I also believe working with students on research projects is an opportunity to advise students. Advising students in research projects involves a mix of mentoring, career guidance (e.g., with respect to discussing career goals and how the research work fits into these goals and/or helps them to achieve these goals), and education (i.e., guiding students through the material related to the research topic). I strongly believe that providing research experiences to undergraduate students can be an extremely valuable part of their overall undergraduate education experience, and can offer unique learning experiences that are not possible within the classroom setting. Furthermore, working closely with a faculty member on research can provide advising and mentoring opportunities that supplement the more standard academic advising interactions, e.g., by being able to discuss the students work together with their career goals.

4 Academic Citizenship and Service

Since joining GU, I have been involved in service to the Computer Science Department, service to the School of Engineering and Applied Science (SEAS), University-wide committee work, and professional service to the broader computer-science community.

Departmental Service. In addition to attending CS Department faculty meetings, I have also been involved in various tasks associated with the CS ABET accreditation process, including managing the Alumni Survey, designing two new Design Advisory Board Surveys, managing the ETS Major Field Test for Computer-Science majors, and in general participating in strategic discussions related to ABET accreditation. Over the past two years, I have also been involved in helping Dr. Yerion with the Bachelor of Arts in Computer Science degree. I was also the faculty advisor for

Upsilon Pi Epsilon and the Computer Science Club (as described above).

School-Wide Service. My service to SEAS includes serving as Associate Director of the ICCSL (i.e., the compute cluster), as SEAS representative on the Mission Advisory Council (Fall 2011–Fall 2012), as the Chair of the Herak Center Space Committee (2012–2013), and as a member of the Herak Lab Renovation Committee (2012-2013). (Note that these two committees are now part of the larger development and programming work on the Integrated Science and Engineering building.) I am also actively involved in the Center for Engineering Design and Entrepreneurship (CEDE) as faculty advisor for Computer Science Senior Design projects and as a faculty representative on the CEDE Faculty Committee. Finally, I have taught four sections of the one-credit UNIV 099 course (one per semester for the past two years). Each section was primarily composed of Engineering students who are struggling academically (on academic probation). This course covers a range of topics including goal setting, time management, learning styles, study skills (listening, reading, note taking, studying for exams), and taking exams. The material is tailored to engineering students by focusing the material on strategies for more technical courses/programs. A large part of the class involves student self reflection (through metacognitive journal entries and in-class exercises) as well as hands-on "assignments" where students try out the methods discussed in class and then write about their effectiveness. I have enjoyed teaching this course because I feel that it provides a useful service to both the students taking the course and to the school as a whole. Although this semester I am not teaching a section, I will be teaching another section in Spring 2015.

University-Wide Service. I am a member of the Academic Council Policy and Planning Committee. This committee has been active, and includes work on examining and making recommendations regarding policies in the catalog, revising and aligning procedures for Academic Council Advisory Committees, and revisions to forms that require academic advisor signatures. I was also elected to the Patent Committee, and participated as a member of the Sponsored Research Office's Grant Advisory Board (which reviewed and made recommendations to policies concerning external funding). I became one of the two SEAS representatives on the Faculty Senate in the Fall of 2013. I was also elected in the Spring to the commencement awards committee.

General Committee Membership and Service. Since joining Gonzaga I have served as a program committee member and reviewer for a number of computer science conferences, workshops, journals, and other professional organizations including:

1. The International Conference on Scientific and Statistical Database Management (SSDBM 2011, 2012), Proceedings Editor, and Program Committe Member (SSDBM 2010, 2014).

2. The International Conference on Conceptual Modeling (ER 2010, 2011, 2012, 2013, 2014), Program Committee Member

3. The International Conference on Web Information Systems (WISE 2011, 2012), Program Committee Member

4. The International Conference on Business Process Management (BPM 2010, 2011), Program Committee Member

5. The IEEE International Conference on Web Services (ICWS), Program Committee Member, 2012

6. The IEEE International Conference on e-Science (e-Science 2010, 2013), Program Committee Member
7. The IEEE International Conference on Embedded Software and Systems (ICESS 2010), Program Committee Member
8. International Conference on Computation Science (ICCS), PC Member for the Workshop on Advances in the Kepler Scientific Workflow System and Its Applications, 2012
9. The Annual Conference on Digital Government Research (dg.o 2010, 2011), Program Committee Member
10. The Annual Consortium for Computing Sciences in Colleges Northwest Regional Conference (CCSC-NW 2012, 2013, 2014), Paper Reviewer and Student Poster Chair.
11. The ACM High Performance Distributed Computing Symposium (HPDC'2010), Program Committee Member
12. The IEEE International Workshop on Scientific Workflows (SWF 2010, 2011), Program Committee Member
13. The International Provenance and Annotation Workshop, Program Committee Member, 2014.
14. The International Workshop on Data Semantics (DataSem'2010), Program Committee Member
15. The International Workshop on Workflow Management in Service and Cloud Computing (WMSC'2010), Program Committee Member
16. The International Workshop on Managing Data Quality for Collaborative Science (MDQCS'2010), Program Committee Member
17. The International Workshop on Workflow Models, Systems, Services and Applications in the Cloud (CloudFlow), PC Member, 2012
18. The Journal of Data Semantics, Co-Editor in Chief, 2011–2013, Editorial Board Member, 2006–2010.
19. The Future Generation Computing Systems (FGCS) Journal, Reviewer, 2012.
20. The Journal of Web Semantics, Reviewer, 2011.
21. The Journal of Computers and their Applications, Reviewer, 2011.
22. IEEE Transactions on Knowledge & Data Engineering (TKDE), Invited Reviewer, 2010
23. Data & Knowledge Engineering Journal (DKE), Invited Reviewer, 2010
24. ACM Computing Surveys (CSUR), Invited Reviewer, 2010
25. Journal of Universal Computer Science (JUCS), Invited Reviewer, 2010

26. W3C Provenance Incubator Group, Invited Expert (2009–2010)

27. NSF ABI Panel, Reviewer, 2011, and IIS Medium Panel, 2014.

5 Overall Self-Assessment

The greatest challenge for me over the past five years at Gonzaga has been juggling an increasing teaching and service load with research activities. Teaching especially has involved a considerable amount of work in terms of course preparation, development, planning, grading, and so on. I also found teaching the Senior Design course very challenging, since this is not a traditional course in that it is heavily project based and requires advising many distinct design projects. Teaching Senior Design has been a learning experience for me, but very rewarding and enjoyable. One factor that has led to increase workload over the past couple of years has been the sudden increase in computer science majors. Many of our upper division courses have 40 or more students, which introduces a number of additional challenges in terms of preparation and time spent grading, etc. In reflecting upon the last five years, I feel that I have been very productive, especially given the shift in workload (more teaching, and more students per class) and responsibilities (within the department and school). However, maintaining a strong research portfolio requires a considerable amount of time and especially within a field such as computer science, where collaborations with faculty at other institutions and service to the broader community are critical. One area where I have had to "cut back" my time is in reviewing and program committee work, and by trying to more narrowly focus my collaboration efforts. Obtaining external funding is also a major challenge at a school like Gonzaga, especially since maintaining collaborations and connections to the broader community requires a considerable amount of travel (e.g., to conferences and meetings). I have also found that involving undergraduate students in research, while very rewarding (and one of the main reasons I wanted to work at a school like Gonzaga), does not always lead to research results. One reason for this is that students often have little time during the academic year outside of their courses to pursue research activities because of the heavy academic loads required in engineering. Many students also don't have the foundation to carry out research activities (even many of the "best" students), and thus, considerable time must be spent getting students "up to speed" to make visible progress. Finally, I have found that students do not have the technical writing skills needed to describe their research and results in a form that is professional and publishable. I think a significant challenge for me in the future will be striking a balance between maintaining a productive research program (including providing students with undergraduate research opportunities) and the time needed to be a successful teacher and to make contributions in university service. With this said, however, I very much look forward in the future to having the opportunity to refine my teaching skills, to contribute to the Computer Science Department and SEAS, and to continue to work closely with students on research and design projects. After teaching at Gonzaga for the last five years, I have many new ideas and insights which I hope I have the opportunity to apply in the coming years by receiving tenure at GU.