esa

ECOSPHERE

INNOVATIVE VIEWPOINT

Enhancing collaboration between ecologists and computer scientists: lessons learned and recommendations forward

Cayelan C. Carey,[1,†] Nicole K. Ward,[1] Kaitlin J. Farrell,[1] Mary E. Lofton,[1] Arianna I. Krinos,[1] Ryan P. McClure,[1] Kensworth C. Subratie,[2] Renato J. Figueiredo,[2] Jonathan P. Doubek,[1] Paul C. Hanson,[3] Philip Papadopoulos,[4] and Peter Arzberger[5]

[1]*Department of Biological Sciences, Virginia Tech, Blacksburg, Virginia, USA*
[2]*Electrical and Computer Engineering, University of Florida, Gainesville, Florida, USA*
[3]*Center for Limnology, University of Wisconsin-Madison, Madison, Wisconsin, USA*
[4]*San Diego Supercomputer Center, University of California-San Diego, La Jolla, California, USA*
[5]*Pacific Rim Applications and Grid Middleware Assembly (PRAGMA), University of California-San Diego, La Jolla, California, USA*

Citation: Carey, C. C., N. K. Ward, K. J. Farrell, M. E. Lofton, A. I. Krinos, R. P. McClure, K. C. Subratie, R. J. Figueiredo, J. P. Doubek, P. C. Hanson, P. Papadopoulos, and P. Arzberger. 2019. Enhancing collaboration between ecologists and computer scientists: lessons learned and recommendations forward. Ecosphere 10(5):e02753. 10.1002/ecs2.2753

Abstract. In the era of big data, ecologists are increasingly relying on computational approaches and tools to answer existing questions and pose new research questions. These include both software applications (e.g., simulation models, databases and machine learning algorithms) and hardware systems (e.g., wireless sensor networks, supercomputing, drones and satellites), motivating the need for greater collaboration between computer scientists and ecologists. Here, we outline some synergistic opportunities for scientists in both disciplines that can be gained by building collaborations between the computer science and ecology research communities, with a focus on the benefits to ecology specifically. We also identify past contributions of computer science to ecology, including high-frequency environmental sensor technology, advanced supercomputing capacity for ecological modeling, databases for long-term and high-frequency datasets, and software programs for ecological analyses, to anticipate future potential contributions. These examples highlight the power and potential for further integration of computer science technology and ideas into the ecological research community. Finally, we translate our own experiences working together as a team of computer scientists and ecologists over the past decade into a conceptual framework with recommendations for supporting productive collaborations at the interface of the two disciplines. We specifically focus on how to apply best practices of team science for bridging computer science and ecology, which we advocate will substantially benefit ecology long-term.

Key words: big data; computational ecology; computer programming; cyberinfrastructure; information technology; interdisciplinary; modeling; quantitative literacy; software; team science.

Received 26 March 2019; accepted 10 April 2019; final files received 29 April 2019. Corresponding Editor: Judy Cushing.

† **E-mail:** cayelan@vt.edu

The Rise of Computationally Intensive Ecological Research Requires Collaborations with Computer Scientists

Ecologists are increasingly using computational approaches to tackle research questions in the era of big data (Soranno and Schimel 2014, Durden et al. 2017). These techniques require both software applications and hardware systems that extend beyond the desktop environment, and thus expertise in computer science. For example, ecologists are increasingly archiving and sharing their data via curated repositories that enable data discovery, reuse, and citation (e.g., via the Dryad Digital Repository; Environmental Data Initiative, EDI; and Knowledge Network for Biocomplexity, KNB); using computer programming in R, Python, C, FORTRAN, and other languages to conduct data management and visualization (Valle and Berdanier 2012, Hampton et al. 2015); deploying high-frequency wireless sensors for environmental monitoring (e.g., the Global Lake Ecological Observatory Network, GLEON and National Ecological Observatory Network, NEON); and analyzing datasets with computationally intensive software programs and models (Lai et al. 2019).

While many ecologists today regularly use computational tools in their research workflows, we anticipate that the increasing size and complexity of ecological datasets will require an even greater and more widespread integration of computational approaches into ecology in the near future. Given the growing collection of long-term monitoring data, high-frequency sensor data, bioinformatics and molecular data, remote sensing data, Earth system model output data, and many other big datasets, ecologists are increasingly conducting data-driven computational science (Hampton et al. 2013, 2017). As an example, a recent survey found that the number of papers published in ecological journals that reported the use of the R statistical language increased fivefold from 2007 to 2018 (Lai et al. 2019).

The accelerating pace of accumulation from new data streams is accompanied by the need for new approaches and tools to store, manage, analyze, and visualize data (Hampton et al. 2013, 2015, 2017), necessitating a greater integration of computer, computational, and information science and engineering (subsequently referred to as computer science) into the common practice of ecology. Moreover, we expect that the demand for both computer science expertise and new hardware and software tools for ecological applications will continue to grow rapidly, given the environmental challenges ecologists are tackling. For example, climate change, biodiversity loss, and habitat fragmentation are complex issues that require intensive modeling and data analyses at multiple spatial and temporal scales (Evans 2012).

Given these research needs, we propose that increasing collaboration between ecologists and computer scientists has the potential to substantially accelerate the advancement of both disciplines, with particularly important benefits for ecology. Similar to the genomic revolution in molecular biology that occurred in the 1990s with the integration of computer science algorithms to query and sort genetic databases, thereby enabling previously impossible genome mapping and subsequent discoveries in intracellular processes and protein structural organization (Mitchell et al. 1990, Fickett 1996, Landweber and Kari 1999, Mac Dónaill 2006), we expect that groundbreaking innovation similarly likely lies ahead with future integration of computer science into ecology. Below, we identify potential mutual benefits of greater collaboration between ecologists and computer scientists, provide recommendations for how to develop productive collaborations among researchers from these two disciplines, and finally share suggestions for how to advance computer science–ecology collaboration.

Increasing Collaboration Between Computer Scientists and Ecologists Will Benefit Both Disciplines

Benefits to ecology

As a result of both recent innovations in computer science and increasing access to new technologies, ecologists are broadening their data acquisition methods to include in situ sensors, satellites, crowdsourcing, and citizen science. These diverse data streams are rapidly expanding the landscape of information available to understand ecological phenomena because they cross traditional ecosystem boundaries, span multiple time scales, and represent fundamentally new kinds of observations. While these new data hold much promise for accelerating the pace of discovery in ecology, they will remain largely untapped without further integrating the expertise of computer scientists to help translate these data into ecological understanding.

While big data and computational challenges are widespread throughout ecology (Hampton et al. 2015), the study of lake water clarity provides a useful example to illustrate how computer science technologies and approaches can overcome these challenges and generate new ways to collect, process, and analyze ecological data. Historically, water clarity was measured by manually sampling a lake with a Secchi disk on the weekly to yearly scale, resulting in thousands of observations being collected annually in the United States (e.g., Lottig et al. 2014). Now, new optical sensors developed by computer and electrical engineers provide highly sensitive measures of water quality every second, equating to over 30 million observations per year from one sensor and potentially billions of observations per year collected by multiple sensors on a single lake (Carey et al. 2012, Weathers et al. 2013). Consequently, the volume and variety of data on lake water clarity and other water spectral characteristics are growing exponentially, as these data are indicator metrics of water quality (Watras et al. 2015, Birgand et al. 2016). Water clarity can also be observed from sensing technologies via aircraft and satellites, which greatly expand the geographic coverage of lakes that can be studied globally (Lee et al. 2018). Citizen scientists are further increasing the spatial coverage of observations by measuring water clarity with mobile devices (e.g., the Lake Observer smartphone app; lakeobserver.org, Graham et al. 2011). All of these aforementioned technologies —new optical sensors, satellite sensors, smartphone apps—are uniquely enabling the collection of new ecological data.

While technologies informed by computer science have enabled the rapid accumulation of water clarity big data (as defined by the 5Vs: data velocity, volume, value, variety, and veracity; Demchenko et al. 2014), transforming those data more rapidly into new ecological knowledge and understanding of lake ecosystems require further computer science collaboration. In particular, major challenges that remain include transmitting ecological data from sensors to discoverable and accessible repositories, converting raw data to meaningful ecological variables (e.g., translating electrical voltages from fluorescent sensors into phytoplankton biomass concentrations; Roesler et al. 2017), extracting usable information from complex and diverse data sources, and connecting patterns in the data to ecological processes (Lee et al. 2018). These challenges create many opportunities for collaboration between ecologists and computer scientists, including the development of cyberinfrastructure (e.g., virtual private network software and cloud computing) that is adaptable to a variety of data streams from a diversity of environmental observatories and facilitates findable, accessible, interoperable, and re-usable data (FAIR data; Wilkinson et al. 2016); the adoption of software techniques and technologies within the ecological research community (e.g., modeling in R); the development of computing power that scales with the size of the data and the demands of the models (Subratie et al. 2017, Turner and Carpenter 2017); and the melding of data-driven models with theory-based ecological models that facilitate both data mining and interpretation (Karpatne et al. 2017). These computing approaches, which all represent significant breakthroughs in the realm of computer science, are beginning to be used by ecologists, but are not yet well-integrated into ecological research (Porter et al. 2005, Seidl 2017).

There are many other examples of collaborations between computer scientists and ecologists that highlight the value of integrating computational tools and methods into ecological research. Research teams of computer scientists and ecologists have generated datasets and applications that track bird migrations at the global scale (eBird; ebird.org) which would not have been possible without the co-development of user-friendly applications and computational infrastructure to manage large remote sensing datasets (Horton et al. 2018). Ecologists and computer scientists have also teamed up to use artificial intelligence to optimize wildlife corridors in computer-modeled landscapes (St. John et al. 2018). Finally, ecologists and computer scientists are using new machine learning methods (sensu Peters et al. 2014) to increase our understanding of global forest biodiversity and productivity (Liang et al. 2016). All of these examples highlight how computer science–ecology collaborations can result in the development of tools that address ecologically relevant questions that span multiple spatial and temporal scales.

Benefits to computer science

Collaborating with ecologists can also provide many benefits to computer scientists. Computer scientists and engineers, and especially experimental computer system researchers, are motivated by the challenges that arise in the design and implementation of applications. Fundamentally, computers are problem-solving machines, and the problems computers are able to solve are revealed by end users through their interactions with applications. Ecology as a discipline tackles many complex challenges that motivate problem-solving; thus, considering the future application needs of ecologists can drive computer science innovation and provide opportunities to tackle new problems using new approaches, resulting in novel systems, software, and publications. For example, ecologists on our team needed to run lake ecosystem model simulations to predict water clarity but did not have the computing infrastructure to do so efficiently. After many discussions, computer scientists on our team created a simple user interface for submitting thousands of simulations of water clarity, which helped to both improve the ecologists' modeling efficiency and develop a novel computing approach for accessing a Web service through the R statistical environment (Subratie et al. 2017).

For computer scientists who work at the interface between applications and systems—for example, on the design and implementation of the various layers of software that are required by computer cyberinfrastructure—collaboration can naturally lead to scientific advances. For example, by focusing on how to effectively make computational resources (hardware and software) more easily accessible to a new community of users, a computer scientist can tackle the design of computer systems from a different perspective, inspiring novel ideas.

Finally, in addition to enabling computer science innovation, tackling ecology questions can provide inherent motivation for computer scientists to collaborate with ecologists because solving environmental problems can profoundly affect the well-being of individuals and society. Following the example above, the development of computing infrastructure for running lake water clarity simulations has allowed our team to run more simulations than would have otherwise been possible to examine the interacting effects of climate and land use change on lake ecosystem dynamics. This modeling has revealed a strong effect of certain weather variables on water quality, which has substantial implications for lake management (Snortheim et al. 2017). Thus, this computer science–ecology collaboration has provided tangible benefits for researchers in both disciplines: The ecologists have gained an improved understanding of lake water quality, and the computer scientists pioneered new computing methods, while advancing the study of a critical ecosystem service upon which our society depends.

Lessons Learned and Recommendations for Productive Collaborations at the Computer Science–Ecology Interface

When working across any set of different disciplines, there will likely be bumps along the road for computer scientists and ecologists collaborating together, which may be exacerbated by a less-established history of collaboration between these research communities. Recent ecology-focused papers on interdisciplinary collaboration (Cheruvelil et al. 2014, Goring et al. 2014, Baker 2015, Read et al. 2016, Cheruvelil and Soranno 2018) and general recommendations for best practices in the team science literature (Bennett et al. 2010, Boerner et al. 2010, Disis and Slattery 2010, Falk-Krzesinski et al. 2010, Bennett and Gadlin 2012, National Research Council 2015) rarely discuss collaborations between ecologists and computer scientists, which suggests that these interdisciplinary collaborations may be less common.

The science of team science is rooted in the identification of common challenges and solutions for collaborative teams (National Research Council 2015). While working together on multiple projects as a team of ecologists and computer scientists over the past decade, we have seen first-hand how these challenges, including disciplinary integration, goal alignment, and team member interactions (sensu National Research Council 2015), can apply to collaboration between the two disciplines.

Here, we collate our experiences into a conceptual framework and guide for how to develop and maintain productive collaborations at the computer science–ecology interface (Fig. 1). We mapped these strategies onto a pyramid, with each tier representing successive stages of computer science–ecology collaboration that build on each other. Lower tiers must be maintained as teams move up the pyramid in the pursuit of achieving novel research outcomes, and should be revisited throughout the research process to enhance collaboration. The content of the pyramid emerged from our team's collaborative experiences and is informed by interdisciplinary research and team science theory (Mathieu et al. 2000, Repko 2008, Salazar et al. 2012, Morse 2015, National Research Council 2015, Stern and Coleman 2015); the subsequent subsections below describe each of the tiers. While these lessons learned and recommendations are by no means limited to computer science–ecology collaborations, we think that our experiences provide a distinct perspective in navigating this particular type of cross-disciplinary research.

Appreciating disciplinary differences and unique disciplinary contributions

The underlying foundation to successful communication and alignment of project goals among all group members in the interdisciplinary collaboration pyramid (Fig. 1) is formed when collaborative computer science–ecology research groups acknowledge, examine, and appreciate differences in intellectual tradition, work culture, and methodological approaches that may exist due to discipline-specific training and experience (Repko 2008). As a starting point for such a discussion, we analyzed a National Research Council (2005) report that was written to motivate collaboration between computer scientists and biologists. This report identifies several similarities and differences in the theoretical frameworks and cultural practices of computer scientists and biologists. For example, the work of computer scientists is often guided by the rules of mathematics, leading many computer scientists to focus on developing abstract, generalizable algorithms (National Research Council 2005). In contrast, the work of ecologists is often constrained by context- and system-dependent empirical data, resulting in an emphasis on data-driven analyses (National Research Council 2005). These fundamentally different approaches can frame how computer scientists and ecologists view the use of models and data in their research.

To develop authentic, productive, and rewarding interdisciplinary collaborations, it is essential that researchers in each discipline fully appreciate the scope of research conducted by members of the other domain, their expertise in the field, and the value of their perspective to the collaboration (Fig. 1; Repko 2008). For ecologists, appreciating computer science collaborators as full members of the project, rather than as technical support or information technology consultants, is critical for establishing mutually rewarding scientific collaborations. In this respect, ecologists must understand and acknowledge the research goals of their computer science colleagues, rather than simply assuming that their expertise is providing a service to an ecological research team in the form of increased computing capacity. Specifically, it is inappropriate to view computer science as merely programming and data management, which are important skills in computer science but not the intellectual core of the discipline (National Research Council 2005). Rather, it is helpful for ecologists to appreciate the heterogeneity and breadth of computer science, which spans theory to experiments, hardware and software to networking, computing and algorithms to data, and systems to components.

Moreover, while the historical contributions of computer science to the life sciences have already had profound impacts, the contributions of the life sciences (including ecology) to computer science have been narrower in scope (National Research Council 2005). For example, while computer science contributions to life science have provided the foundation for new sub-fields such as bioinformatics and computational biology, applications of biological phenomena to computer science, such as swarm intelligence and neural networks, have provided valuable new

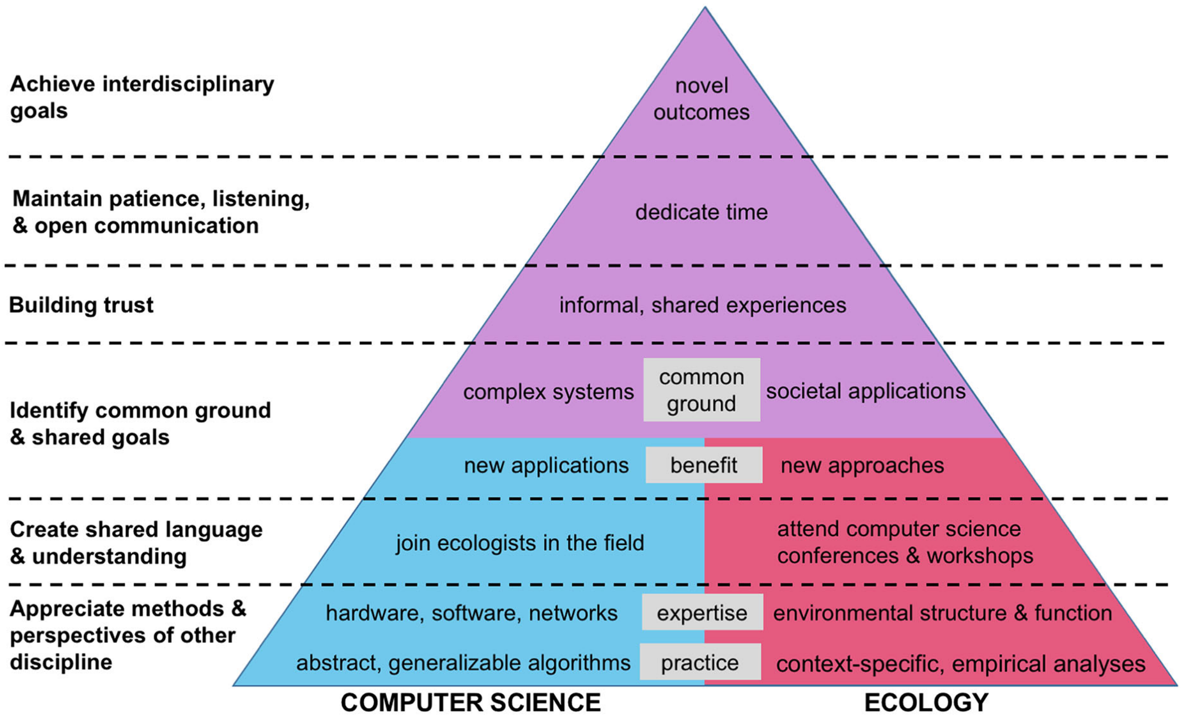

Fig. 1. A conceptual framework for computer science–ecology collaborations, based on the authors' experiences and informed by interdisciplinary research and team science theory (Mathieu et al. 2000, Repko 2008, Salazar et al. 2012, Morse 2015, National Research Council 2015, Stern and Coleman 2015). Each tier of the pyramid, from the foundation (appreciate methods and perspectives of the other discipline) to the apex (achieve interdisciplinary goals), represents successive stages for enhancing interdisciplinary collaboration. We note that collaboration is not linear, and thus, it is important to continually maintain and revisit lower tiers throughout the research process. The text in white to the left of the pyramid summarizes the strategy of each tier to enhance collaboration. The blue and red sections represent examples, activities, descriptions, or actions specific to the individual disciplines of computer science and ecology, respectively; purple tiers represent shared interdisciplinary perspectives, approaches, and outcomes; and the text in gray describes categories for computer science and/or ecology considerations and descriptions. For example, much of the expertise of computer science is in hardware, software, and networks, whereas much of the expertise of ecology is in environmental structure and function.

tools but have not (yet) led to new sub-disciplines within computer science. Thus, the onus is on life scientists—here, our focus is on ecologists—to understand the research goals of their computer scientist colleagues and ensure that computer science–ecology collaborations are of equal value to both parties.

Creating shared language and understanding

Creating a common language and cross-disciplinary understanding is the second tier in computer science–ecology team formation (Fig. 1). Shared understanding of the research problem will help teams realize and understand the mutual benefits for team members from both disciplines (National Research Council 2005, Repko 2008, Salazar et al. 2012). Consequently, it is critical that team members know enough about the other discipline to be able to acknowledge and more fully appreciate the equal value of contributions from both fields of study.

Participating in training activities in the other discipline can build understanding among computer scientists and ecologists. For example, we found that bringing computer scientists from our team into the field to collect lake water samples and troubleshoot water quality sensors helped improve their understanding of the workflow of the field aspect of ecological research, and ultimately helped shape the computational tools used by the group. Likewise, having the ecologists in our collaborative group attend conferences and workshops centered around computer science and information technology (e.g., meetings of the Pacific Rim Applications and Grid Middleware Assembly [PRAGMA] community; www.pragma-grid.net) provided both critical insights into how some computer scientists approach research questions, and an opportunity to learn discipline-specific language to improve communication within our team. Building a common vocabulary while explicitly exploring how each discipline approaches the research process can help computer science–ecology teams surmount challenges when they inevitably arise (O'Rourke and Crowley 2013).

Identifying common ground and shared goals

The third tier to enhancing interdisciplinary collaboration between computer scientists and ecologists is identifying shared goals, which starts by recognizing the important similarities between the two disciplines (Fig. 1). Finding common ground is a critical part of this stage of collaboration (Repko 2008). At a high level, the disciplines of computer science and ecology are both focused on exploring and understanding complex systems (National Research Council 2005). In addition, research from both disciplines has profound societal implications, whether by changing the tools we use to live and work (e.g., mobile phones and computers) or by informing our understanding of the planet and environmental policy in an era of global change (e.g., developing air pollution standards and climate change temperature targets; Kiesler et al. 1984, Nurullah 2009). These similarities could lead to a shared familiarity in working with complex systems and the understanding of—and motivation to do—research that impacts society (Moor 1984, Alpert et al. 2003), both of which could enable productive research between computer scientists and ecologists. Frequently revisiting the initial goals developed at the onset of a computer science and ecology collaboration will ultimately increase the likelihood of mutually beneficial research outcomes.

Problem concept mapping and shared mental mapping can help interdisciplinary collaborators identify shared research goals (Mathieu et al. 2000, Morse 2015, National Research Council 2015). This exercise entails project team members working together to outline their understanding of the research problem or question to be addressed, including all relevant components and contributing factors. Initially, every team member's own mental map will be slightly different, with larger differences usually occurring between team members from different disciplines (Morse 2015). The process of integrating each team member's mental map into one shared conceptual map of the research problem will increase shared understanding among team members, and importantly, across disciplines.

Developing new shared goals in computer science–ecology collaborations can be facilitated by deliberate efforts to bring computer scientists and ecologists together, such as workshops and funding opportunities that support joint projects. Co-developing research proposals opens new collaborative avenues in each respective discipline to researchers in the other. In addition to the obvious benefit of increased access to research funding, the process of writing these proposals and participating in joint workshops has helped us spur closer collaborations by increasing the breadth and depth of our shared research vision. For example, this publication was catalyzed at a joint workshop of computer scientists in PRAGMA and freshwater ecologists in GLEON that was held to improve the computing capacity for lake water quality modeling. Other opportunities for facilitating collaboration exist through co-organizing workshops and special sessions at Ecological of Society of America (ESA) conferences and scientific working groups at the National Center for Ecological Analysis and Synthesis (NCEAS). In particular, NCEAS is working actively to enable collaborations between ecologists and computational scientists and use innovative computational methods to improve our understanding of ecology (nceas.ucsb.edu/data-science).

Building trust

The fourth tier to enhancing interdisciplinary collaboration in our pyramid is building trust between interdisciplinary team members (Fig. 1), which is required for maintaining long-term partnerships (National Research Council 2005, Salazar et al. 2012, Stern and Coleman 2015). Outside of the field or laboratory, social interactions among members of interdisciplinary teams can help ensure lasting collaborations (Read et al. 2016) by building trust (Salazar et al. 2012, Stern and Coleman 2015). Informal, shared experiences build social capital among team members and develop relationships that will help the team overcome conflicts and disciplinary divides that are bound to emerge during the process of cross-disciplinary research (Cheruvelil et al. 2014). In our team, we have found that a lack of familiarity can initially lead to communication barriers, whereby individuals may not voluntarily voice their opinion during a group discussion when they do not know the other team members well. However, these types of challenges are easier to surmount when informal activities help build camaraderie—whether through sharing a meal or a laugh. Interpersonal bonding through informal interactions is particularly valuable for developing collective communication competence and trust, which are both essential for maintaining high-performing teams (Thompson 2009, Salazar et al. 2012, Cheruvelil et al. 2014, Read et al. 2016).

Maintaining patience, active listening, and open communication

The fifth tier of the interdisciplinary collaboration pyramid (Fig. 1)—creating a culture of patience, listening, and open communication—is necessary for the maintenance and evolution of interdisciplinary teams long-term (Brown et al. 2015). While all interdisciplinary research requires patience, the disciplinary differences and less-established history of collaboration between ecology and computer science may necessitate the cultivation of a particular ethos of patience in building effective computer science–ecology teams. Because many ecologists lack familiarity with the research goals and multiple research areas of computer scientists, there may be a lack of understanding of the constraints associated with computer science–ecology collaborations. For example, the ecologists in our team were generally unaware of the financial cost of running many computationally intensive simulation models using virtual networks and distributed servers: there was the initial expectation that an infinite number of model simulations could be run using the computer scientists' infrastructure.

In addition, discipline-specific workflows and jargon may slow the development of effective collaborations. However, meeting often and in person, when feasible, can help overcome these hurdles, as "just being in the room" can provide real-time clarification of jargon or discrepancies, enhancing team effectiveness (National Research Council 2015). For example, as described above, we collaborated in developing a Web service platform (GRAPLEr, www.graple.org) to provide distributed computing tools for modeling lake ecological processes (Subratie et al. 2017). At a recent meeting, we discovered an inconsistency in how the ecologists and computer scientists understood the workflow of model output. Because of this discrepancy, ecological interpretation of the model output would have been incorrect, despite the software performing correctly. If not for the active listening and open communication among the team members, the time cost of identifying and correcting the discrepancy could have been expensive and may not have been realized until after the project was completed.

Through all these steps, ecologists must learn about the research pursuits that are exciting for their computer science collaborators. Active, interested listening proves invaluable in building these authentic connections across disciplines. In our experience, the most rewarding computer–ecology collaborations are those in which researchers in both disciplines are committed to long-term, deep collaborations that use cutting-edge computing tools to tackle complex ecological challenges. While these collaborations may take multiple years to produce a product, which underscores the transactional costs of interdisciplinary research (Brown et al. 2015), we believe they ultimately lead to better science, new and better tools and methods to conduct that science, rapid access to constantly evolving data tools and computational infrastructure, and novel research outcomes (the top of the interdisciplinary collaboration pyramid in Fig. 1).

Enabling Greater Collaboration Among Ecologists and Computer Scientists Requires New Perspectives and Training Initiatives

When forming computer science–ecology collaborations, we emphasize the importance of engaging with computer scientists early in the process of technology development, as open communication among collaborators about what technical functionality is helpful (for the ecologist) vs. what advances computer science research questions and is feasible (for the computer scientist) can accelerate the pace of science on both ends. Early engagement may also allow, where practical, for human-centered design (HCD) of computer science tools with ecologists as a designated end user (Maguire 2001, Jaimes et al. 2007). Because development of new computer science tools is inherently an iterative process, ecologists need to be patient and willing to try out technology products that may not be fully developed, as this end-user testing and feedback is essential for product refinement in computer science. Finally, it is important that all collaborators receive credit appropriate to their discipline for participating in the project (Goring et al. 2014). For example, ecologists are often assessed on the number of peer-reviewed journal articles they have published, whereas papers published in prestigious peer-reviewed conferences or patents are also relevant metrics of academic success for computer scientists.

A solid computational skill set among ecologists—such as programming and math skills—will enhance collaborative productivity between ecologists and computer scientists by improving communication (Rentsch et al. 2010, 2014). In our experience, computer code or community standards such as FAIR for datasets provide a common language for communicating our research across the two disciplines. For example, our team has encountered multiple situations in which collaboratively going through each other's code was far more effective for communicating our research objectives than trying to explain them verbally. Well-developed quantitative skills and computational literacy not only allow ecologists to better understand computer science research, but also help ecologists communicate their own research goals to improve collaboration with computer scientists.

While we are certainly not the first group to underscore the importance of computational and quantitative training for students in ecology (e.g., Barraquand et al. 2014, Carey et al. 2015, Touchon and McCoy 2016, Klug et al. 2017, Farrell and Carey 2018), we suggest that one previously overlooked benefit to such training is that computational ecologists are more likely to recognize the benefits of, and successfully engage in, collaborations with computer scientists. Consequently, in addition to a strong curriculum in mathematics, we recommend that undergraduate ecology students learn basic skills in a programming language (e.g., R, Python, C++, and Java). Furthermore, in our experience, integrating training modules that analyze big data into undergraduate ecology courses has successfully built students' computational literacy at a range of experience levels (www.MacrosystemsEDDIE.org; Carey et al. 2015, Carey and Gougis 2017, Klug et al. 2017, O'Reilly et al. 2017), especially because the modules can be easily adapted for a wide range of classrooms. Ecology undergraduate and graduate students who become familiar with computing techniques early in their careers will not only be able to tackle new ecology research questions, but also be able to gain increasingly essential skills in data science, programming, and computing. These computational skills will benefit them in their future careers, regardless of whether they are in academia, industry, or government (Mellody 2014).

Conclusions

There is a clear need with multitudinous benefits for closer interaction among computer science and ecology. For the ecologists in our group, collaborating with computer scientists has provided some of the most fruitful and exciting research opportunities experienced throughout our careers. While we cannot claim that our experience will be universal, we advocate that developing collaborative relationships with computer scientists has the potential to greatly advance ecology as a discipline.

Acknowledgments

Cayelan C. Carey conceived and designed the conceptual framework with Nicole K. Ward, Kaitlin J. Farrell, and Mary E. Lofton, and all authors contributed to writing and provided feedback. None of the authors have any conflict of interest. This work was catalyzed by a PRAGMA-GLEON workshop at the PRAGMA 32 conference in Gainesville, Florida, USA, and builds upon several years of collaboration among team members. We gratefully acknowledge support from

National Science Foundation (NSF) grants ACI 1234983, ACI 1339737, CNS 1737424, CNS 1527415, EF 1702506, and ICER 1517823. We thank Grace Hong for facilitating the PRAGMA-GLEON workshop, the larger PRAGMA community for supporting this work, and previous and current GRAPLE students, especially Jaikrishna Sukumar, Saumitra Aditya, Vahid Daneshmand, Alexandra Gerling, and Kathleen Hamre.

Literature Cited

Alpert, P., A. Keller, S. Airame, W. K. Lauenroth, R. V. Pouyat, H. A. Mooney, K. H. Rogers, and C. M. Breen. 2003. The ecology-policy interface. Frontiers in Ecology and Evolution 1:45–50.

Baker, B. 2015. The science of team science. BioScience 65:639–644.

Barraquand, F., T. H. G. Ezard, P. S. Jørgensen, N. Zimmerman, S. Chamberlain, R. Salguero-Gómez, T. J. Curran, and T. Poisot. 2014. Lack of quantitative training among early-career ecologists: a survey of the problem and potential solutions. PeerJ 2:e285.

Bennett, L., and H. Gadlin. 2012. Collaboration and team science: from theory to practice. Journal of Investigative Medicine 60:768–775.

Bennett, L. M., H. Gadlin, and S. Levine-Finley. 2010. Collaboration & Team Science: a field guide. National Institutes of Health No. 10-7660. National Institutes of Health Office of the Ombudsman, Center for Cooperative Resolution, Bethesda, Maryland, USA.

Birgand, F., K. Aveni-Deforge, B. Smith, B. Maxwell, M. Horstman, A. B. Gerling, and C. C. Carey. 2016. First report of a novel multiplexer pumping system coupled to a water quality probe to collect high temporal frequency in situ water chemistry measurements at multiple sites. Limnology and Oceanography: Methods 14:767–783.

Boerner, K., N. Contractor, H. J. Falk-Krzesinski, S. M. Fiore, K. L. Hall, J. Keyton, B. Spring, D. Stokols, W. Trochim, and B. Uzzi. 2010. A multi-level systems perspective for the science of team science. Science Translational Medicine 2:49cm24.

Brown, R. R., A. Deletic, and T. H. F. Wong. 2015. How to catalyse collaboration. Nature 525:315–317.

Carey, C. C., R. Darner Gougis, J. L. Klug, C. M. O'Reilly, and D. C. Richardson. 2015. A model for using environmental data-driven inquiry and exploration to teach limnology to undergraduates. Limnology and Oceanography Bulletin 24:32–35.

Carey, C. C., and R. D. Gougis. 2017. Simulation modeling of lakes in undergraduate and graduate classrooms increases comprehension of climate change concepts and experience with computational tools. Journal of Science Education and Technology 26:1–11.

Carey, C. C., P. Hanson, D. A. Bruesewitz, G. W. Holtgrieve, E. L. Kara, K. C. Rose, R. Smyth, and K. C. Weathers. 2012. Organized Oral Session 43. Novel applications of high-frequency sensor data in aquatic ecosystems: discoveries from GLEON, the Global Lake Ecological Observatory Network. Bulletin of the Ecological Society of America 93:100–105.

Cheruvelil, K. S., and P. A. Soranno. 2018. Data-intensive ecological research is catalyzed by open science and team science. BioScience 68:813–822.

Cheruvelil, K. S., P. A. Soranno, K. C. Weathers, P. C. Hanson, S. J. Goring, C. T. Filstrup, and E. K. Read. 2014. Creating and maintaining high-performing collaborative research teams: the importance of diversity and interpersonal skills. Frontiers in Ecology and the Environment 12:31–38.

Demchenko, Y., C. De Laat, and P. Membrey. 2014. Defining architecture components of the big data ecosystem. Page 104–112 *in* Proceedings of the International Conference on Collaboration Technologies and Systems, 2014. IEEE, Minneapolis, Minnesota, USA.

Disis, M. L., and J. T. Slattery. 2010. The road we must take: multidisciplinary team science. Science Translational Medicine 2:1–5.

Durden, J. M., J. Y. Luo, H. Alexander, A. M. Flanagan, and L. Grossmann. 2017. Integrating "big data" into aquatic ecology: challenges and opportunities. Limnology and Oceanography Bulletin 26:101–108.

Evans, M. R. 2012. Modeling ecological systems in a changing world. Philosophical Transactions of the Royal Society of London B: Biological Sciences 367:181–190.

Falk-Krzesinski, H. J., K. Börner, N. Contractor, S. M. Fiore, K. L. Hall, J. Keyton, B. Spring, D. Stokols, W. Trochim, and B. Uzzi. 2010. Advancing the science of team science. Clinical and Translational Science 3:263–266.

Farrell, K. J., and C. C. Carey. 2018. Power, pitfalls, and potential for integrating computational literacy into undergraduate ecology courses. Ecology and Evolution 8:7744–7751.

Fickett, J. W. 1996. Finding genes by computer: the state of the art. Trends in Genetics 12:316–320.

Goring, S. J., K. C. Weathers, W. K. Dodds, P. A. Soranno, L. C. Sweet, K. S. Cheruvelil, J. S. Kominoski, J. Rüegg, A. M. Thorn, and R. M. Utz. 2014. Improving the culture of interdisciplinary collaboration in ecology by expanding measures of success. Frontiers in Ecology and the Environment 12:39–47.

Graham, E. A., S. Henderson, and A. Schloss. 2011. Using mobile phones to engage citizen scientists in research. Eos 92:313–315.

Hampton, S. E., C. A. Strasser, J. J. Tewksbury, W. K. Gram, A. E. Budden, A. L. Batcheller, C. S. Duke, and J. H. Porter. 2013. Big data and the future of ecology. Frontiers in Ecology and the Environment 11:156–162.

Hampton, S., et al. 2015. The Tao of open science for ecology. Ecosphere 6:120.

Hampton, S. E., et al. 2017. Skills and knowledge for data-intensive environmental research. BioScience 67:546–557.

Horton, K. G., B. M. Van Doren, F. A. La Sorte, D. Fink, D. Sheldon, A. Farnsworth, and J. F. Kelly. 2018. Navigating north: How body mass and winds shape avian flight behaviours across a North American migratory flyway. Ecology Letters 21:1055–1064.

Jaimes, A., D. Gatica-Perez, N. Sebe, and T. S. Huang. 2007. Human-centered computing: toward a human revolution. IEEE Computer 40:30–34.

Karpatne, A., G. Atluri, J. H. Faghmous, M. Steinbach, A. Banerjee, A. Ganguly, S. Shekhar, N. Samatova, and V. Kumar. 2017. Theory-guided data science: a new paradigm for scientific discovery from data. IEEE Transactions on Knowledge and Data Engineering 29:2318–2331.

Kiesler, S., J. Siegel, and T. W. McGuire. 1984. Social psychological aspects of computer-mediated communication. American Psychologist 39:1123–1134.

Klug, J. L., C. C. Carey, D. C. Richardson, and R. Darner Gougis. 2017. Analysis of high-frequency and long-term data in undergraduate ecology classes improves quantitative literacy. Ecosphere 8:e01733.

Lai, J., C. J. Lortie, R. A. Muenchen, J. Yang, and K. Ma. 2019. Evaluating the popularity of R in ecology. Ecosphere 10:e02567.

Landweber, L. F., and L. Kari. 1999. The evolution of cellular computing: nature's solution to a computational problem. BioSystems 52:3–13.

Lee, Z., et al. 2018. Global water clarity: continuing a century-long monitoring. Eos 99. https://doi.org/10.1029/2018EO097251

Liang, J., et al. 2016. Positive biodiversity-productivity relationship predominant in global forests. Science 354:aaf8957.

Lottig, N. R., T. Wagner, E. N. Henry, K. S. Cheruvelil, K. E. Webster, J. A. Downing, and C. A. Stow. 2014. Long-term citizen-collected data reveal geographical patterns and temporal trends in lake water clarity. PLoS ONE 9:e95769.

Mac Dónaill, D. A. 2006. Digital parity and the composition of the nucleotide alphabet. IEEE Engineering in Medicine and Biology Magazine 25:54–61.

Maguire, M. 2001. Methods to support human-centered design. International Journal of Human-Computer Studies 55:587–634.

Mathieu, J. E., T. S. Heffner, G. F. Goodwin, S. Salas, and J. A. Cannon-Bowers. 2000. The influence of shared mental models on team process and performance. Journal of Applied Psychology 85:273–283.

Mellody, M. 2014. Training students to extract value from big data: summary of a workshop. National Academies Press, Washington, D.C., USA.

Mitchell, E. M., P. J. Artymiuk, D. W. Rice, and P. Willett. 1990. Use of techniques derived from graph theory to compare secondary structure motifs in proteins. Journal of Molecular Biology 212:151–166.

Moor, J. H. 1984. What is computer ethics? Metaphilosophy 16:266–275.

Morse, W. C. 2015. Integration of frameworks and theories across disciplines for effective cross-disciplinary communication. *In* M. O'Rourke, S. Crowley, S. D. Eigenbrode, and J. D. Wulfhorst, editors. Enhancing communication & collaboration in interdisciplinary research. SAGE Publications, Thousand Oaks, California, USA.

National Research Council. 2005. *In* J. C. Wooley and H. S. Lin, editors. Catalyzing inquiry at the interface of computing and biology. National Academies Press, Washington, D.C., USA.

National Research Council. 2015. *In* N. J. Cooke and M. L. Hilton, editors. Enhancing the effectiveness of team science. National Academies Press, Washington, D.C., USA.

Nurullah, A. S. 2009. The cell phone as an agent of social change. Rocky Mountain Communication Review 6:19–25.

O'Reilly, C. M., et al. 2017. Using large data sets for open-ended inquiry in undergraduate science classrooms. BioScience 67:1052–1061.

O'Rourke, M., and S. J. Crowley. 2013. Philosophical intervention and cross-disciplinary science: the story of the Toolbox Project. Synthese 190:1937–1954.

Peters, D. P. C., K. M. Havstad, J. Cushing, C. Tweedie, O. Fuentes, and N. Villanueva-Rosales. 2014. Harnessing the power of big data: infusing the scientific method with machine learning to transform ecology. Ecosphere 5:art67.

Porter, J., P. Arzberger, H.-W. Braun, P. Bryant, S. Gage, T. Hansen, P. Hanson, C.-C. Lin, F.-P. Lin, and T. Kratz. 2005. Wireless sensor networks for ecology. AIBS Bulletin 55:561–572.

Read, E. K., M. O'Rourke, G. S. Hong, P. C. Hanson, L. A. Winslow, S. Crowley, C. A. Brewer, and K. C. Weathers. 2016. Building the team for team science. Ecosphere 7:e01291.

Rentsch, J. R., L. A. Delise, E. Salas, and M. P. Letsky. 2010. Facilitating knowledge building in teams: Can a new team training strategy help? Small Group Research. 41:505–523.

Rentsch, J. R., L. A. Delise, A. L. Mello, and M. J. Staniewicz. 2014. The integrative team knowledge building strategy in distributed problem-solving teams. Small Group Research. 45:568–591.

Repko, A. F. 2008. Interdisciplinary research process and theory. SAGE Publications, Los Angeles, California, USA.

Roesler, C., et al. 2017. Recommendations for obtaining unbiased chlorophyll estimates from in situ chlorophyll fluorometers: a global analysis of WET Labs ECO sensors. Limnology and Oceanography: Methods 15:572–585.

Salazar, M. R., T. K. Lant, S. M. Fiore, and E. Salas. 2012. Facilitating innovation in diverse science teams through integrative capacity. Small Group Research 43:527–558.

Seidl, R. 2017. To model or not to model, that is no longer the question for ecologists. Ecosystems 20:222–228.

Snortheim, C. A., P. C. Hanson, K. D. McMahon, J. S. Read, C. C. Carey, and H. A. Dugan. 2017. Meteorological drivers of hypolimnetic anoxia in a eutrophic, north temperate lake. Ecological Modelling 343:39–53.

Soranno, P. A., and D. S. Schimel. 2014. Macrosystems ecology: big data, big ecology. Frontiers in Ecology and the Environment 12:3.

St. John, R., S. F. Toth, and Z. B. Zabinsky. 2018. Optimizing the geometry of wildlife corridors in conservation reserve design. Operations Research 66:1471–1485.

Stern, M. J., and K. J. Coleman. 2015. The multidimensionality of trust: applications in collaborative natural resource management. Society & Natural Resources 28:117–132.

Subratie, K. C., S. Aditya, S. Mahesula, R. Figueiredo, C. C. Carey, and P. C. Hanson. 2017. GRAPLEr: a distributed collaborative environment for lake ecosystem modeling that integrates overlay networks, high-throughput computing, and WEB services. Concurrency and Computation: Practice and Experience 29:e4139.

Thompson, J. 2009. Building collective communication competence in interdisciplinary research teams. Journal of Applied Communication Research 37:278–297.

Touchon, J. C., and M. W. McCoy. 2016. The mismatch between current statistical practice and doctoral training in ecology. Ecosphere 7:e01394.

Turner, M. G., and S. R. Carpenter. 2017. Ecosystem modeling for the 21st century. Ecosystems 20:211–214.

Valle, D., and A. Berdanier. 2012. Computer programming skills for environmental sciences. Bulletin of the Ecological Society of America 93:373–389.

Watras, C. J., K. A. Morrison, J. T. Crawford, C. P. McDonald, S. K. Oliver, and P. C. Hanson. 2015. Diel cycles in the fluorescence of dissolved organic matter in dystrophic Wisconsin seepage lakes: implications for carbon turnover. Limnology and Oceanography 60:482–496.

Weathers, K., et al. 2013. The Global Lake Ecological Observatory Network (GLEON): the evolution of grassroots network science. Limnology and Oceanography Bulletin 22:71–73.

Wilkinson, M. D., et al. 2016. The FAIR guiding principles for scientific data management and stewardship. Scientific Data 3:160018.