2024

Developing Partnerships for Academic Data Science Consulting and Collaboration Units

Marianne Huebner
Michigan State University

Laura Bond
Boise State University

Felesia Stukes
Johnson C. Smith University

Joel Herndon
Duke University

David J. Edwards
Virginia Commonwealth University

See next page for additional authors

Authors

Marianne Huebner, Laura Bond, Felesia Stukes, Joel Herndon, David J. Edwards, and Gina-Maria Pomann

WILEY

Developing partnerships for academic data science consulting and collaboration units

Marianne Huebner[1,2] | **Laura Bond**[3] | **Felesia Stukes**[4,5] |
Joel Herndon[6,7] | **David J. Edwards**[8] | **Gina-Maria Pomann**[9]

[1]Center for Statistical Training and Consulting, Michigan State University, East Lansing, Michigan, USA

[2]Department of Statistics and Probability, Michigan State University, East Lansing, Michigan, USA

[3]Biomolecular Research Center, Boise State University, Boise, Idaho, USA

[4]Computer Science, Engineering and Mathematics Department, Johnson C. Smith University, Charlotte, North Carolina, USA

[5]Historically Black Colleges and Universities (HBCU) Data Science Consortium, Atlanta, Georgia, USA

[6]Center for Data and Visualization Sciences, Duke University, Durham, North Carolina, USA

[7]Duke University Libraries, Duke University, Durham, North Carolina, USA

[8]Statistical Sciences and Operations Research, Virginia Commonwealth University, Richmond, Virginia, USA

[9]Department of Biostatistics and Bioinformatics, Duke University, Durham, North Carolina, USA

Correspondence

Marianne Huebner, Center for Statistical Training and Consulting, Michigan State University, East Lansing, Michigan, USA.
Email: huebner@msu.edu

Data science consulting and collaboration units (DSUs) are core infrastructure for research at universities. Activities span data management, study design, data analysis, data visualization, predictive modelling, preparing reports, manuscript writing and advising on statistical methods and may include an experiential or teaching component. Partnerships are needed for a thriving DSU as an active part of the larger university network. Guidance for identifying, developing and managing successful partnerships for DSUs can be summarized in six rules: (1) align with institutional strategic plans, (2) cultivate partnerships that fit your mission, (3) ensure sustainability and prepare for growth, (4) define clear expectations in a partnership agreement, (5) communicate and (6) expect the unexpected. While these rules are not exhaustive, they are derived from experiences in a diverse set of DSUs, which vary by administrative home, mission, staffing and funding model. As examples in this paper illustrate, these rules can be adapted to different organizational models for DSUs. Clear expectations in partnership agreements are essential for high quality and consistent collaborations and address core activities, duration, staffing, cost and evaluation. A DSU is an organizational asset that should involve thoughtful investment if the institution is to gain real value.

KEYWORDS

bioinformatics, biostatistics, image analysis, social sciences

1 | INTRODUCTION

Data science consulting and collaboration units (DSUs) are core infrastructure for research at universities. DSU employees have training in a specific quantitative area such as statistics, epidemiology, bioinformatics and engineering, among others. They consult or collaborate with faculty, research staff or students on topics related to data management, study design, data analysis, data visualization, predictive modelling, preparing reports, manuscript writing or advising on statistical methods. For the DSU staff, this work could result in a role as Co-PI on grant-funded projects, co-authorship on papers and conference presentations, the development of new statistical methodology, new technology or in scholarly products designed to communicate the findings to diverse audiences.

University leadership, PIs of laboratories or PhD advisors may view some of these activities as routine (Desai et al., 2022), but not taking them seriously results in reproducibility issues, suboptimal or even incorrect choices of analyses or interpretation (https://stratos-initiative.org/en; Baillie et al., 2022; Sauerbrei et al., 2014) or questionable research practices such as low statistical power, p-hacking and selective reporting (Bishop, 2019; Held & Schwab, 2020). DSU staff are integral members of research teams often fulfilling a variety of roles such as expert professional, advocate, communicator or leader (Zapf et al., 2019). Responsibilities are broad and can include maintaining the integrity of data collection and analysis, ensuring fidelity to study plan and adherence to participant protections and respecting the interests of those who fund the research (American Statistical Association [ASA], 2022).

In addition to these formalized connections and responsibilities, DSUs can achieve more value as an active part of the larger university network, rather than as a siloed unit. Thus, a thriving DSU needs collaborative partnerships. Such partnerships can be developed with other research service units on campus, individual departments, colleges or with multi-institutional collaboratives. Potential partners seek to provide their members with access to data science expertise or seek to collaborate with units having complementary expertise. While a statement of work (SOW) defines the work for specific research projects (Kumuthini et al., 2020; Peterson et al., 2022), successful partnerships will negotiate rules for engagement, for example, with a Memorandum of Understanding (MOU). This helps clarify what is covered or not covered and provides continuity for collaborations with changes in administrative leadership.

The scope of this paper is that DSUs provide services university wide or across specific colleges, for example, academic statistical consulting units with administrative home in a central Office of Research, biostatistical consulting units for medical colleges or data science consulting units located in Libraries. The administrative home has many implications for mission, budget, operations and sustainability. These units may engage in a multitude of partnerships depending on mission and staffing availability. Collaborative partnerships discussed in this paper focus on university internal services or research activities where at least one principal investigator is a faculty member at the same university as the DSU. This can include partnerships with colleges, departments, university research centres, core services or research teams. It cuts across disciplines including arts, communication, arts, social sciences, health and STEM (science, technology, engineering and mathematics) disciplines. Partnership agreements should clearly define rules for engagement, such as the type of activities staff and faculty will engage in, the duration of the engagement and financial agreements. External contracts are out of scope for this paper. These involve additional layers of university level support (legal for business associate agreements, IT for secure data transfer, Smart IRB, etc.). Many partnerships evolve from informal arrangements and can take time to develop into more formal arrangements. However, informal arrangements will not be discussed in this paper.

The purpose of this paper is to provide guidance on best practices for building partnerships for DSUs who may differ by their administrative homes, number of staff, scope or funding mechanisms. We discuss six general rules for successful partnerships for DSUs including alignment with mission and strategic plans, ensuring sustainability and growth, the importance of communication and of having clear expectations for partnership agreements and to be prepared for the unexpected. To develop guidance for the academic research community about what should be considered when developing partnerships as a DSU, an interest group of academic leaders of DSUs was formed as part of the Data Science Consulting and Collaboration Workshop, May 2023 (https://datascienceacademy.ncsu.edu/data-science-consulting/ds-consulting-highed-workshop/). The lived experiences of the authors represent diverse situations for DSUs with whom they are affiliated, from central services in the Office of Research to Libraries to units within colleges or affiliated with research technology facilities. An example of an affiliation for a DSU is shown in Figure 1 with a diverse set of partnerships. Six rules to identify, develop, sustain or grow partnerships were identified by consensus of the authors and illustrative examples of these rules were discussed for various organizational models of DSUs. DSU collaborations encompass faculty, research staff or students across the university or in specific colleges, with multidisciplinary research teams on grants, or external partners. Some of these DSUs have experiential or teaching components for data science. In Section 2, we describe how DSUs may be situated within an institution and give examples of potential partnerships specific to the mission of those DSUs. In Section 3, we describe six rules that DSUs need to consider when developing and managing successful partnerships. Several examples are given for each rule.

2 | COLLABORATIVE PARTNERSHIPS FOR ACADEMIC DSUS

DSUs typically have a formal management structure which relies on an official chain of command within the college or university. Models of academic DSUs differ by scope, staffing and administrative homes. Staffing can range from a specific data scientist within a unit, to a DSU where
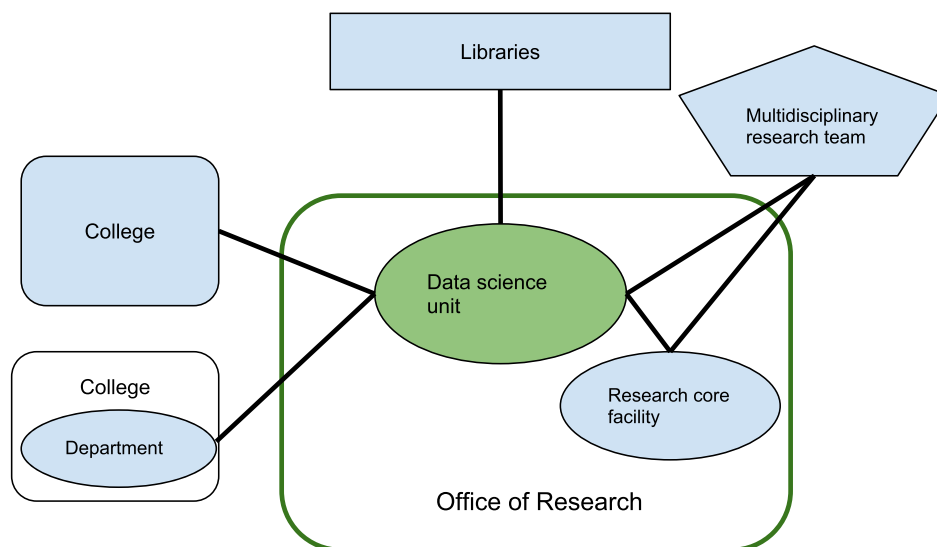
FIGURE 1 Example where the administrative home of the DSU (green) is the Office of Research with a network of partnerships (blue) that contribute funding or joint hires.

students are the primary consultants mentored by a faculty member, or a DSU with full time staff serving all researchers across the university (Gibbons & Freund, 1980; Vance, 2015). The type of model may be determined by preferences and values of institutional leaders and the support they are willing to contribute, including financial, marketing and administrative support. Some amount of institutional support is necessary, to reflect the importance of the service to the university as a whole for active scholarship, providing education in data science skills, and for honing and expanding skills for the DSU staff. University support ensures employee benefits and access to continuing professional education. Sustainable institutional support enables access to DS support not covered by funded activities (Perkins et al., 2016; Vance, 2015). In contrast, many fee-based mechanisms, such as cost centres, have restrictions on how much income can be generated from within-university partners and how it can be used. This can impact how partnerships are structured and can make it difficult to foster long-term collaborations. Policies and regulations imposed by university leadership can facilitate or hinder cross-disciplinary collaborations or collaborations between multiple DSUs. For example, departmental structure may not recognize teaching and/or scholarship outside of its own domain in review and promotion criteria, supervisors may not recognize the value of more time-intensive tasks (e.g., data quality assurance, evaluating analysis methods for their applicability in projects) or colleges with their own DSU may not be open to enabling collaborations with other DSUs.

The mission or purpose varies for different DSUs depending on staffing and types of services provided. DSUs in scope for this paper can range from a single consultant within a core facility, hybrid units with both full time staff and graduate student research assistants, to units with a large number of staff and diversity in scope (Perkins et al., 2016). This determines which administrative homes and partners may be suitable.

Several examples are listed below:

1. If the DSU supports grant proposal development and funded research projects, then the DSU can be a university wide core facility with funding from the Office of Research who may also be the administrative home. In addition, specific colleges may become partners supporting DSU staff with specific skill sets for their faculty research.
2. If the DSU provides expertise and guidance on study design, statistical methods or conducting data analysis and interpreting results to improve quality and impact of research, the DSU can be a university wide core facility with funding from the Office of Research who may also be the administrative home. Alternatively, the DSU may be located in academic units (colleges, departments) with activities related to those academic units.
3. If the DSU trains and mentors graduate students to become expert consultants, then possible partners are the Graduate School or departments who would like to include such education in their graduate programme and provide funding for research assistantships and faculty mentors.
4. If the DSU supports analyses and analytical tools for instrumentation output such as genomics, flow cytometry, mass spectrometry and services, then the unit may be housed in a core facility that provides technical support for such instrumentation. Partners can be those who have an agreement with the core facility for these services or have need for consultation regarding resulting data.
5. Emerging technologies and advanced statistical methods may require specialized skills and sophisticated expertise beyond that of current staff of DS consulting units. Potential partners can be individual faculty who would be available for collaborations on grants and the DSU can broker such connections outside of the DSU. Alternatively, partnerships may require support to cover the cost of training to the DSU consultants. Parker and colleagues pointed out that data-intensive research is a team sport (Parker et al., 2021).

DSUs should build capacity in high demand areas via new hires or professional development of current staff members in response to the strategic goals of the university or specific academic units (Van Tuyl, 2023). Building relationships with university leaders and understanding their goals enable a DSU to form potential partnerships, which increase collaborations and funding opportunities.

Partnerships can have a large impact for the members of the academic unit interested in engaging with a DSU. An established partnership can sustain collaborations through challenges such as disruptions due to the COVID19 pandemic (Figure 2, college 1). When administrative leaders are supportive, it can lead to an increase in research collaborations (Figure 2, college 2). When priorities for academic leaders change, this can result in discontinuation of the partnerships, removing access for their members (Figure 2, college 3). Establishing multiple diverse partnerships can help the DSU be sustainable in the face of changing demands. Academic leaders should realize that enabling or restricting access to a DSU will affect faculty, research staff and students in their unit.

3 | SIX SIMPLE RULES FOR PARTNERSHIPS

3.1 | Rule 1: align with institutional strategic plans

Aligning a DSU with an institution's strategic plan requires a purposeful and coordinated approach. First, the DSU should make every effort to familiarize itself with the institution's strategic plan's goals, priorities and initiatives. Doing so will allow for clear identification of synergies between the unit's expertise and the strategic plan, whether by supporting specific research themes, academic programmes, community engagement or diversity and inclusion efforts. Communication and collaborative planning with key stakeholders are necessary to ensure the unit's mission actually aligns with strategic priorities.

An adaptation of the DSU's activities may be required to better align with strategic goals. For example, if interdisciplinary research is a priority, the unit can offer cross-disciplinary DS training or emphasize support for interdisciplinary research projects. This might involve adding more specialized support related to specific fields (e.g., genomics and spatial/temporal data) or emerging technologies. By participating in the support of research that closely aligns with the institution's strategic priorities, the unit can showcase its impact while advancing the institution's mission. Participation in institution-wide initiatives, events and seminars/symposiums related to the strategic plan will facilitate integration of the DSU into the institution's efforts (Rule 5: Communicate). Measuring and reporting the impact of the DSU's contributions to any strategic initiative are also needed to demonstrate tangible evidence of alignment with the strategic plan. A cycle of collecting feedback and adaptations ensures ongoing alignment.

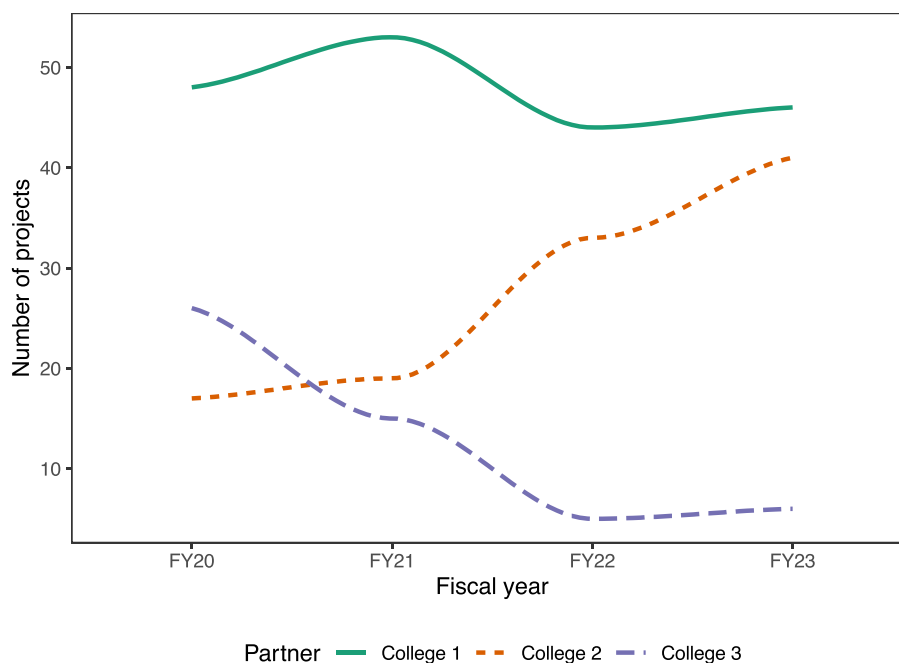**FIGURE 2** Number of projects across fiscal years for three academic units (example provided by CSTAT/MSU). College 1 has a longstanding partnership with dedicated FTE but the research office moved to remote work in FY22, College 2 became a partner and increased their contribution and College 3 discontinued the partnership agreement and, thus, services became less accessible to their faculty and students.

Example 1. At Duke University School of Medicine, researchers have access to the Biostatistics, Epidemiology, and Research Design (BERD) Core, a DSU that connects them with potential quantitative collaborators across the institution. Individuals who lead this resource work with research groups to track the needs for quantitative expertise and collect data to identify unmet collaboration needs. For example, they may track that there is a need to hire someone with expertise in spatial analysis if enough researchers are looking for a collaborator with this expertise. This assessment allows the resource to provide guidance to leadership to align with strategic plans to identify unmet needs and help evaluate what the goals should be for building a collaboration to best meet those needs.

Example 2. Some universities may have multiple decentralized DSUs. One DSU may be located in the medical colleges, another in the College of Agriculture, and yet another in the College of Arts and Sciences. Some DSUs may focus on biostatistics support, some on data analytics and visualization and some on bioinformatics with overlapping skill sets. This decentralization requires duplicated administrative support, and the decentralization could be problematic for faculty and students seeking the best fit of DS support for their research. Each DSU should strive to have a plan for referring projects that do not meet their mission or expertise to other DSUs. However, coordinating efforts among DSUs to centralize information faces roadblocks when this coordination is not seen as central to the mission of the DSU and/or its collaborators. For small units, this can have an isolating effect on staff members of these units.

3.2 | Rule 2: cultivate partnerships that fit your mission

The DSU may be inherently interconnected with many cross-campus units and should be thoughtful about how to produce positive value among them. A suitable place to anchor partnerships is around the unit's purpose and overall objectives. This mission can serve as a guide that will direct how the DSU develops and cultivates relationships and partnerships with its staff, clients and other stakeholders. If the unit does not currently have a formal mission statement, creating one as a first step is strongly encouraged. The unit's approach to partnerships should be systematic, intentional and thoughtful. It can begin to assess ways to best engage with others across the institution and leverage partnerships that are motivated by a shared purpose. The unit may find that in some cases, the partners to be involved have complementary needs/skills, shared goals or both.

The DSU may be approached by clients with requests that currently do not fit the mission or does not currently employ staff with the necessary skills. Whether or not there are criteria within a DSU for determining which partnerships to take on, establishing partnerships outside of the mission leads to 'mission creep' and may require additional personnel and/or skills development. These can dilute the overall effectiveness of the DSU, in absence of careful planning. It is important for DSUs to establish networks and contacts to be able to provide referrals or references to potentially suitable contacts from a myriad of options (Pomann et al., 2023). Referring potential clients to another unit more suited to their project can in turn build trust and respect for your unit.

Informal networks can have a significant role in the success of the DSU. Organizational network interactions, relationships and connections with individuals and other units outside of existing partnerships can contribute positive value for the DSU. Lin (2001) defines such an approach as social capital or an investment in social relations with expected returns that are captured from embedded resources in social networks. These can form the building blocks for new partnerships.

Example 1. A DSU with the mission of supporting grant-funded projects is invited by a multidisciplinary team of faculty to pursue collaborative efforts with the goal of providing services for a research proposal funded by federal or foundation granting agencies. This collaboration has demonstrated a successful, complementary, and mutually beneficial partnership to both the unit and faculty collaborators. The DSU could approach the leading faculty's departments or college administrators for sponsorship to provide overarching funding in a fixed amount, with continued (long-term) or interim (short-term) agreements, or through full-time equivalency salary (FTE) to provide continued activities or to start future endeavours.

Example 2. A DSU with the mission of training and mentoring graduate students to become consultants or collaborators includes graduate student research assistants on their staff. Trained graduate research assistants within the DSU can support the research activities of other graduate students seeking statistical help for their doctoral dissertations. These activities develop skills and experience for these students. These graduate student assistants can also serve as 'ambassadors' for the DSU to explain the available support to other graduate students or explain the value of joining the DSU as a staff member. The DSU could request a dedicated graduate research assistant position directly from The Graduate School to support graduate student success on campus.

3.3 | Rule 3: ensure sustainability and prepare for growth

Financial sustainability of the DSU is the bare minimum for its longevity. A strategic plan within a DSU, aligned with its mission, provides insight into which research areas or disciplines to support, the development of appropriate skill sets, the staffing needs for the unit (e.g., additional FTEs or ad hoc consultants) and budget projections that identify funding sources. We echo comments from Parker and colleagues (2021) that securing champions who recognize the centre from multiple administrative levels is essential. In addition, the strategic plan for the DSU needs to ensure that university-specific priorities are met (see Rule 1). The unit should be aware of how methodologies specific to partner departments and disciplines are changing to anticipate the needs of these clients.

Preparing budget projections and identifying funding sources necessarily go hand-in-hand. An understanding of the budgeting process at your institution is required to know about financial resources. This includes allocations from state funding with specified uses, endowments with or without specified use requirements, income from external sources, universities' policies of overhead allocation from grants and availability of discretionary funding from colleges and academic units. Understanding which funding sources may be available for short- and long-term financing can help the DSU navigate financing current fixed costs and initiatives and anticipate funding for future goals.

It is essential that funding sources are available for university-specific priorities, such as supporting grant-funded research projects or providing training in DS methods to faculty or students across the university. It is also crucial that the university invests in continuing professional development of the DSU staff, supporting unfunded work on grant proposal development to reap future benefits, providing administrative support for managing the unit and other aspects of human resource costs. It is difficult to envision long-term sustainability without explicit funding support from university administration. When DSUs are required to completely cover their operating costs, they may struggle to survive in university settings which can stifle innovation of cross-disciplinary faculty and limit accessibility of services to traditionally marginalized groups. To avoid such unintended consequences, university leadership in partnership with DSUs should make intentional efforts to create service models that are equitable and accessible to all, including those who may need support in overcoming technological or monetary barriers to data services.

> **Example 1.** DSUs may offer educational opportunities for faculty and students. These can take the form of fee-based workshops, course credit classes, web-based on-demand training, or free offerings from the unit. Such offerings range from discipline-specific topics and methodologies to broader programming and software carpentry, or responsible conduct of research. They contribute to the unit's reputation and recognition of expertise and thus may help with sustainability. However, in the experience of the authors such activities do not generate more than a nominal amount of income and they typically require a funding source to cover personnel costs to develop such training.

> **Example 2.** DSUs can receive direct funding from grants, for personnel buy-out and possibly for expenses such as training and computing expenses. In some cases, DSU employees may themselves be principal investigators of extramurally funded grants. These buy-outs provide funding to the DSU. Partnership agreements between the DSU and the funded unit should cover distribution of overhead funds and how related, non-funded activities will be paid for, e.g. new proposal development, post-grant manuscript writing, and any unfunded data sharing agreements. Grant funding alone, without these agreements, does not ensure DSU sustainability, but does provide growth opportunities for the DSU.

> **Example 3.** Sustainability is not only about money. Continuing education and discussions about the health of the centre are critical for DSU staff to remain relevant in their respective areas of expertise. A DSU at one R1 university regularly holds a 'Quality Academy' for full-time staff. These are group activities and discussions aimed to develop best practices and transform statistical collaboration and consulting. Such activities are crucial for long-range planning and ensuring quality services.

3.4 | Rule 4: define clear expectations in a partnership agreement

Successful and sustainable partnerships have clear expectations for the collaborations. Activities, expected outcomes and time allocated to these activities should be negotiated at the start.

While smaller projects can be managed with clear research aims and analysis plans, a more extensive partnership that is longer term and/or covers multiple projects needs clear expectations through formal partnership agreements, contracts or MOUs. A partnership agreement describes the services that are in scope as well as define those that are not covered (Table 1).

If a joint appointment or %FTE of one person staffs the partnership, an MOU should cover annual review and promotions criteria (Simmons, 2011).

A project database that tracks timelines, outcomes and time spent is needed to make sure the work stays aligned with the agreement and regular check-ins ensure satisfaction of all parties (Rule 5: Communicate) or can identify potential problems early (Rule 6: Expect the unexpected).

TABLE 1 Potential topics to be covered in a partnership agreement (not an exhaustive list).

Topic	Content examples
Objectives	Produce competitive proposals for external funding; promote the statistical and computational skills of graduate students
Core activities	Data management, study design, data analysis, data visualization, advice on statistical methods, manuscript writing, repository services, instruction on software, etc.
Grant proposals	Rules about pre-award engagement depending on post-award funding for activities and level of inclusion of DSU staff
Funded grants	Rules about activities covered by the grant or by the partnership agreement
Activities not covered	Coursework and instruction
Initial contact	Request that initial contact is via an online form to identify type of collaboration with a brief project description, key personnel and contact, funding sources (if applicable)
Hours per project or hours per client	Number of hours for students working on dissertations or for faculty and research staff for their research projects; determine whether the number of hours are per project or per individual
Staff assignment to project	Specific requests for staff members versus flexibility for the DSU regarding assignment depending on availability
Evaluation and feedback	Frequency (quarterly, biannually and annual) with predetermined metrics; qualitative feedback from project PIs; agreed upon criteria for annual review and promotion for dedicated staff for the partnership
Cost	Fixed cost, hourly, salary or annual contribution
Length of agreement	1 to 5 years, or continuing if there was a joint hire; shorter durations as mutually required by the consultant and the client

The rapid change in the scope of data science can prove treacherous for DSUs as the understanding of what is 'usual and customary' shift from the original partnership agreements. There can be reasons for not engaging or disengaging from partnerships such as a lack of staff availability, the represented skill set in the unit is not a good fit, unreasonable workload given the compensation or due to different priorities for the DSU or their potential partners.

Example 1. An investigator has funds to pay for omics data analysis and defines the activities and outcomes in general terms for their research aim. During the course of the project the investigator requested work in addition to 'standard reports' for this type of data. To accommodate such requests, flexibility is needed to accommodate additional time and budget. Communicating an estimate for the additional work should take place before this work commences to make sure all parties agree on the new scope.

Example 2. The Duke University BERD Core follows a collaboration model that embeds custom long-term teams of quantitative experts within departments, divisions, centres and so forth throughout the School of Medicine. MOUs outline the amount of faculty and staff effort that will be dedicated to work on the research team annually. MOUs have supported more than 35 fully funded Core staff biostatisticians, as well as partial effort for more than 20 different quantitative faculty in the Department of Population Health who mentor the staff teams. This collaboration unit model was referenced in comparison to a similar partnership model used at Stanford University to develop their Quantitative Sciences Unit (Desai et al., 2022).

Example 3. A partnership agreement was developed with a college partner for 45% FTE of a statistician employed by the DSU. The MOU outlined that annual performance evaluation was to be conducted in line with job expectations for the DSU and feedback from the college partner. A new Dean of the college changed evaluation criteria for appointments in the college, but they were not in line with the evaluation criteria in the DS consulting unit. Conversations occurred to ensure the statistician was evaluated according to the agreed upon criteria.

3.5 | Rule 5: communicate!

Regular communication requires practice and discipline. It is crucial to prepare an 'elevator speech' or a 'pitch' that concisely explains what the data science consulting unit does, its recent successes and future goals and how it aligns with strategic goals. When the opportunity arises, this pitch will facilitate awareness with university leaders as well as bring information to new faculty and new administrators. This needs to be frequently reviewed and updated. Crafting concise and compelling stories takes practice.

Several levels of communication are needed: (1) exchanging information with new faculty and administrators and advertising the services across the university or relevant academic units, (2) building relationships with current stakeholders to report usage and successes and areas for improvement, (3) listening to goals and priorities of other units and being open to collaborations and (4) building informal networks.

DSUs are uniquely positioned to understand the academic enterprise and facilitate connections between groups. In addition to data science skills, the importance of providing leadership training and coaching for DSU staff (Moldoveanu & Narayandas, 2019) cannot be overstated to support the DSU's ongoing partnerships and growth. Leadership qualities including effective communication, management under uncertainty, excellent organizational skills, ongoing self-reflection and ability to interact with university leadership and to build trust and new relationships are important to consider when selecting or grooming leaders for a DSU. Building a leadership team and shared strategic decision making prepares the DSU for unexpected changes—in the field of data science, in university leadership, in staffing changes and in emerging science and technologies.

Communication can take many forms, for example, social media posts, writing reports, giving presentations at resource fairs, building a website with testimonials, creating short videos or conducting surveys. Working with coordinators in specific academic units or graduate programmes facilitates the distribution of information (Oliver et al., 2019) and nurtures DSU champions for the long term (Parker et al., 2021).

Example 1. A report is generated with information about how DSU services are used by faculty, research staff, and students in a college. The level of detail for such partnership success metrics depends on the preferences of the college level administrators in charge of these partnerships. It may be sufficient to report the number of projects stratified by students or faculty, number of grant proposals and information about funded grants and possibly highlight success stories and goals for the DS consulting unit that may benefit the college. When more details are requested, a more in-depth report is provided such as including the names of researchers, their mentors, if applicable, status of the project (ongoing, on hold or finished), name of consultant and cumulative hours used. Separate tables include grant proposals and externally funded projects with faculty name, consultant name, status and hours used. This is based on a project database that logs consulting requests, time sheets and notes about the projects. Such reports provide tangible evidence of alignment with partnership agreements. A cycle of feedback collection and adaptation ensures ongoing satisfaction.

Example 2. DSU representatives (typically the director and a graduate student research assistant) attend the annual graduate student resource fair hosted by The Graduate School. A poster with background about the DSU, flyers with information and QR code (if flyers run out) and a sign-up sheet for the email list are available. The representatives engage with the graduate students and listen to their research plans. If needed, referrals to other services at the university can be made.

Example 3. An elevator speech is a brief, compelling statement about what a DSU does and their valuable contributions to the research enterprise. A DSU could develop multiple elevator speeches depending on the audience and the purpose, namely, for education or for promoting services. Advice on effective speeches are available in the business world, but they may be too generic to apply in consulting contexts (Simpson, 2016). It may be helpful to create an elevator speech about a recent case study including an overview of the setting and challenges that were overcome with the appropriate statistical method without getting into technical details. The key findings may or may not be included. If the purpose of the elevator speech is to convince a college dean to engage in a partnership with the DSU, it is helpful to have usage information from this college available.

3.6 | Rule 6: expect the unexpected!

Outreach, careful planning and deliberate communication play a central role in defining successful data science partnerships. While the previous rules greatly improve the opportunities for a successful partnership, these efforts cannot completely mitigate the evolving, contested space that defines data science (Donoho, 2017). This rapid change in scope of data science can prove treacherous for DSUs as the shifting understanding of what comprises the 'usual and customary' scope of data science work can undermine partnership agreements. For example, changing expectations in the field of data science for the usage of particular software/approaches to visualizations and associated coding could potentially create misunderstandings between the DSU and the partner. Investing in new technologies and developing new skill sets may be an expectation of the DSU by the partners even without funding to cover associated costs. As the scope of data science evolves, it can be difficult to specify the scope of partnerships in advance. Amendments to partnership agreements may be necessary in addition to identifying institutional financial support.

An understanding of budgeting and awareness of risks helps DSUs to be prepared for the unexpected as people, organizations and institutions shift over time. There may be periods of uncertain funding since grant-funded projects can be awarded at unpredictable times and timelines need to be adjusted that will impact all stakeholders. Partnership agreements may end as priorities change or agreed upon work is completed. Successful data science partnerships should account for the dynamic nature of data science and institutional arrangements in which they are formed

by acknowledging these elements of uncertainty and designing strategies to mitigate their impact over the life of the agreement and effect on the staff involved.

Changes in institutional arrangements often undermine partnership agreements. Staff/students change, academic units reorganize with new leadership at the helm and institutional priorities or structures shift. Partnership agreements that are predicated on the understanding of one set of individuals or a particular set of institutional arrangements can flounder as new participants review the partnership agreements without the benefit of the original shared understanding. These types of changes may not be pre-specified in the agreements and can lead to ad hoc solutions. Clear expectations for topics outlined in Table 1 may mitigate some effects of such expected or unexpected changes.

Best practices for DSUs include encouraging innovation on the DS team, checking partnership expectations regularly and reevaluating agreements and strategic plans. Suggested approaches for this are given.

1. Encourage innovation on the data science team: Data science teams should be encouraged to gain expertise with new tools and methods as part of the partnership agreement. This can benefit the partner organization in accessing the latest tools and methodologies, and it enables the DSU to expand its skillset and adapt to changing trends. While a lack of staffing, time or funding may prevent experimentation, DS teams that embrace experimentation are less likely to be surprised as the data science environment evolves or may even become obsolete.

2. Check expectations regularly: Partnership agreements focus on particular outcomes or a defined set of work and should include provisions for evaluating the work or reviewing the agreement before the final outcome. Meeting regularly during the project period will help ensure that that the effort remains in alignment with the project or determine if adjustments are needed. For agreements that impact data science staff evaluations, these regular meetings are essential to ensure that data science teams receive credit for their efforts.

3. Reevaluating agreements and strategic plans: Staffing changes or organizational changes should immediately trigger a review of the partnership agreement. Existing templates for transition plans to ensure smooth handover of projects can be calibrated for specific agreements. Adding terms to the initial agreement that specify how DSUs and partners negotiate staffing and organizational changes can reduce uncertainty and provide mutually agreed means of terminating the agreement in the event of these changes.

Example 1. Gene expression analysis demonstrates how data scientists must remain active in continuing education and their ability to adapt to new tools and technologies. While PCR-based methods have been used widely for decades to analyse a small number of preselected genes, gene expression quantification has advanced from microarrays (for hundreds of pre-identified genes) to RNA-seq (whole transcriptome shotgun sequencing) and to single-cell RNA-seq (identification of expression to specific cells). Data scientists must be sufficiently familiar with the rapidly evolving technologies and biological understanding to perform correct and meaningful analyses. They should prepare for advances in spatial transcriptomics, exosome profiling, micro-RNA sequencing and their applications to personalized medicine, to name only a few just in 2023. The recognition and guidance of good experimental design persists, however, and remains critically important.

Example 2. A DSU was at capacity with the number of research projects that can be handled with current staffing. One PhD level staff member resigned due to finding a fully remote job in the pharmaceutical industry. The DSU hired additional graduate student research assistants, prioritized projects and communicated clearly about expected delays when responding to requests. The director engaged with partners to identify demand for data science skills, possibly not sufficiently represented in the unit, and discussed opportunities for joint hires at the PhD level.

Example 3. Many DSU processes are built on top of technology information systems such as web applications for building and managing surveys, databases that are used by researchers from multiple sites and shared computing clusters or repositories that store, organize and share project files. At times, there are planned and unplanned system outages, data breaches or unanticipated shutdown or failure of systems and equipment that could disrupt the DSUs operational workflow. There may be intangible costs that are associated with losing system access such as increased stress on staff, partner dissatisfaction, or missed deadlines. Coordinating with IT Support services ahead of time and adhering to data backup policies may help mitigate negative effects of unplanned technology-based failure.

Example 4. The programming languages used in data science continually change as languages, community coding norms and analytic tools evolve to reflect the preferences of the research and computing communities. At a minimum, disagreements within partnerships over code, tools and methods can add friction to a partnership. Successful partnerships will attempt to scope the methods and tools in advance while reviewing possible changes during the life of the partnership. In general, all partners should also consider how these tools and methodology choices potentially impact the ability of others outside of the partnership to understand the research.

4 | DISCUSSION

This paper provides six rules to identify, develop, sustain or grow partnerships providing short- and long-term benefits to both the DSU and the partner. These six rules are intertwined: Building partnerships that fit the mission (Rule 2) requires effective communication with stakeholders (Rule 5) and alignment with strategic plans (Rule 1). Defining clear expectations at the outset of the partnership agreement (Rule 4) mitigates, but does not prevent, unforeseen issues (Rule 6). Ensuring sustainability and preparation for growth (Rule 3) is the result of effectively managing the other five rules, and Rule 6, expect the unexpected, is embedded in everything a DSU does. Some metrics used to measure the success of the partnership are described in Rule 5, but metrics should also be used to ensure alignment with the institution's strategic plan, justify sustainability and identify areas for growth, and provide consistent tools for communicating value beyond the DSU. A carefully designed project database enables a DSU to extract the necessary information and allows more complex metrics (e.g., total project collaboration hours, total mentoring hours and partner comments) to be disseminated. A successful DSU will revisit these rules, and their partnerships, on a continuing basis.

These rules are based on consensus of the authors, who come from a variety of backgrounds and institutions where they support data science activities. Regardless of the size of the unit or the university, we emphasize that a DSU is an organizational asset that should involve thoughtful investment if the institution is to gain real value. The number of data scientist roles is projected to grow 36% between 2021 and 2031 (Bureau of Labor Statistics, US Department of Labor, 2023), and organizations, including higher education institutions, are maximizing value using data science to leverage data for decision-making, research and strategic practices (Ranjan & Foropon, 2021). Higher education has become an increasingly competitive environment in that students compete for admission, academics compete for intellectual merits and institutions compete for students, faculty and external funding (Altbach et al., 2009). If implemented correctly, DSUs can give universities a competitive advantage that results in an increase in the performance and productivity of all other units.

Data science has steadily evolved into a core role in data-focused organizations. The scope and definition of work in this space continue to transform as new technologies, software and methods continue to redefine the concept of 'data science'. Sustaining and growing a DSU is a challenging task, but the rewards could include more widespread use of responsible data science practices across the university (Parker et al., 2021). DSUs provide core infrastructure for research for grants and contracts and student success among other types of support (Pruitt & Vance, 2022).

There are many benefits of building partnerships for the DSU and for their partners. Partnerships introduce more stability in the funding commitments for a DSU. Social capital for DSUs and their partners emerge from such collaborations with individuals or with academic units (Kezar, 2014). Shared values can advance innovative scientific endeavours and provide contextual learning opportunities and motivate professional development. When access to DS services is available for members of the partners (e.g., students and faculty of a particular college), they will find DS support quickly, increasing high quality research output. University leadership in partnership with DSUs should make intentional efforts to create service models that are equitable and accessible to all, including those who may need support in overcoming technological or monetary barriers to data services. Funding for data science support may differ across disciplines, for example, STEM disciplines compared to Humanities and may be more difficult to obtain for traditionally marginalized groups.

In conclusion, this paper helps with providing a better understanding of the important contributions that DSUs provide and challenges they might face. Guidance for identifying, developing and maintaining successful partnerships for academic DSUs is formulated in six rules. Illustrative examples for each rule may serve to alert DSU staff and their academic partners to helpful practices or potential pitfalls. While these rules are not exhaustive, they are derived from experiences in a diverse set of DSUs rather than focus on one DSU model.

ORCID

Marianne Huebner https://orcid.org/0000-0002-9694-9231
Laura Bond https://orcid.org/0000-0003-4975-1019
Felesia Stukes https://orcid.org/0000-0002-7773-1421

Joel Herndon (ID) https://orcid.org/0000-0001-9995-9040

Gina-Maria Pomann (ID) https://orcid.org/0000-0002-6324-3953

REFERENCES

Altbach, P. G., Reisberg, L., & Rumbley, L. E. (2009). Trends in global higher education: Tracking an academic revolution. 2009 World Conference on Higher Education - The New Dynamics of Higher Education and Research for Societal Change and Development, Paris, France. https://unesdoc.unesco.org/ark:/48223/pf0000183219

American Statistical Association (ASA). (2022, February 1). Ethical guidelines for statistical practice. https://www.amstat.org/your-career/ethical-guidelines-for-statistical-practice

Baillie, M., le Cessie, S., Schmidt, C. O., Lusa, L., Huebner, M., & Initiative, for the T. G. "Initial D. A." of the S. (2022). Ten simple rules for initial data analysis. *PLoS Computational Biology*, *18*(2), e1009819. https://doi.org/10.1371/journal.pcbi.1009819

Bishop, D. (2019). Rein in the four horsemen of irreproducibility. *Nature*, *568*(7753), 435. https://doi.org/10.1038/d41586-019-01307-2

Bureau of Labor Statistics, U.S. Department of Labor. (2023, September 2). Occupational outlook handbook, data scientists. https://www.bls.gov/ooh/math/data-scientists.htm

Desai, M., Boulos, M., Pomann, G. M., Steinberg, G. K., Longo, F. M., Leonard, M., Montine, T., Blomkalns, A. L., & Harrington, R. A. (2022). Establishing a data science unit in an academic medical center: An illustrative model. *Academic Medicine : Journal of the Association of American Medical Colleges*, *97*(1), 69–75. https://doi.org/10.1097/ACM.0000000000004079

Donoho, D. (2017). 50 years of data science. *Journal of Computational and Graphical Statistics*, *26*(4), 745–766. https://doi.org/10.1080/10618600.2017.1384734

Gibbons, J. D., & Freund, R. J. (1980). Organizations for statistical consulting at colleges and universities. *The American Statistician*, *34*(3), 140–145. https://doi.org/10.2307/2683870

Held, L., & Schwab, S. (2020). Improving the reproducibility of science. *Significance*, *17*(1), 10–11. https://doi.org/10.1111/j.1740-9713.2020.01351.x

Kezar, A. (2014). Higher education change and social networks: A review of research. *The Journal of Higher Education*, *85*(1), 91–125.

Kumuthini, J., Chimenti, M., Nahnsen, S., Peltzer, A., Meraba, R., McFadyen, R., Wells, G., Taylor, D., Maienschein-Cline, M., Li, J.-L., Thimmapuram, J., Murthy-Karuturi, R., & Zass, L. (2020). Ten simple rules for providing effective bioinformatics research support. *PLoS Computational Biology*, *16*(3), e1007531. https://doi.org/10.1371/journal.pcbi.1007531

Lin, N. (2001). Building a network theory of social capital. In R. Dubos (Ed.), *Social capital. Theory and research* (1st ed). New York: Routledge. https://doi.org/10.4324/9781315129457

Moldoveanu, M., & Narayandas, D. (2019, April). The future of leadership development. *Harvard Business Review*. https://hbr.org/2019/03/the-future-of-leadership-development

Oliver, J. C., Kollen, C., Hickson, B., & Rios, F. (2019). Data science support at the academic library. *Journal of Library Administration*, *59*(3), 241–257. https://doi.org/10.1080/01930826.2019.1583015

Parker, M. S., Burgess, A. E., & Bourne, P. E. (2021). Ten simple rules for starting (and sustaining) an academic data science initiative. *PLoS Computational Biology*, *17*(2), e1008628. https://doi.org/10.1371/journal.pcbi.1008628

Perkins, S. M., Bacchetti, P., Davey, C. S., Lindsell, C. J., Mazumdar, M., Oster, R. A., Peduzzi, P. N., Rocke, D. M., Rudser, K. D., & Kim, M. (2016). Best practices for biostatistical consultation and collaboration in academic health centers. *The American Statistician*, *70*(2), 187–194. https://doi.org/10.1080/00031305.2015.1077727

Peterson, R. A., Hochheimer, C. J., Grunwald, G. K., Johnson, R. L., Wood, C., & Sammel, M. D. (2022). Reaping what you SOW: Guidelines and strategies for writing scopes of work for statistical consulting. *Stat*, *11*(1), e496. https://doi.org/10.1002/sta4.496

Pomann, G.-M., Truong, T., Boulos, M., Ebony Boulware, L., Brouwer, R. N., Curtis, L. H., Kapphahn, K., Khalatbari, S., McKeel, J., Messinger, S., O'Hara, R., Pencina, M. J., Samsa, G. P., Spino, C., Zidanyue Yang, L., & Desai, M. (2023). Needles in a haystack: Finding qualitative and quantitative collaborators in academic medical centers. *Academic Medicine : Journal of the Association of American Medical Colleges*, *98*(8), 889–895. https://doi.org/10.1097/ACM.0000000000005212

Pruitt, T. R., & Vance, E. A. (2022). Statistics and data science collaboration laboratories: Engines for development. In *Promoting statistical practice and collaboration in developing countries*. Chapman and Hall/CRC.

Ranjan, J., & Foropon, C. (2021). Big data analytics in building the competitive intelligence of organizations. *International Journal of Information Management*, *56*, 102231. https://doi.org/10.1016/j.ijinfomgt.2020.102231

Sauerbrei, W., Abrahamowicz, M., Altman, D. G., le Cessie, S., Carpenter, J., & STRATOS initiative. (2014). STRengthening analytical thinking for observational studies: The STRATOS initiative. *Statistics in Medicine*, *33*(30), 5413–5432. https://doi.org/10.1002/sim.6265

Simmons, E. H. (2011). Making interdisciplinarity possible. https://www.insidehighered.com/advice/2011/08/19/making-interdisciplinarity-possible

Simpson, D. (2016). "Going up?" A sport psychology consultant's guide to the elevator speech. *Journal of Sport Psychology in Action*, *7*(2), 109–120. https://doi.org/10.1080/21520704.2016.1182091

Van Tuyl, S. (Ed.). (2023). *Hiring, managing, and retaining data scientists and research software engineers in academia: A career guidebook from ADSA and the US-RSE*. Zenodo. https://doi.org/10.5281/zenodo.8329337

Vance, E. A. (2015). Recent developments and their implications for the future of academic statistical consulting centers. *The American Statistician*, *69*(2), 127–137. https://doi.org/10.1080/00031305.2015.1033990

Zapf, A., Huebner, M., Rauch, G., & Kieser, M. (2019). What makes a biostatistician? *Statistics in Medicine*, *38*(4), 695–701. https://doi.org/10.1002/sim.7998

AUTHOR BIOGRAPHIES

Marianne Huebner is the Director of the Center of Statistical Training and Consulting (CSTAT) and a Professor of the Department of Statistics and Probability at Michigan State University, East Lansing, MI. CSTAT reports to the Office of Research and Innovation and has full time staff members as well as graduate students research assistants who collaborate with faculty, research staff and graduate students across the university and provide statistical support to external clients. DSU website: https://cstat.msu.edu/

Laura Bond is an Associate Research Professor in the Biomolecular Research Center at Boise State University. In this position, she consults with graduate students, faculty and staff in general statistical methodology and statistical analysis of bioinformatic data (RNA-seq and mass-spectrometry-based proteomics), in addition to conducting statistical analyses on her own extramurally funded projects. She oversees one part-time Research Assistant who conducts and advises on bioinformatic analyses. DSU website: https://www.boisestate.edu/brc/

Felesia Stukes is an Associate Professor of Computer Science at Johnson C. Smith University (JCSU) and a Co-Director of the HBCU Data Science Consortium. She has led two federally funded programmes to develop research capacity in data science. Dr. Stukes also directs the JCSU Social Computing Lab (SCL) within which students and research faculty can apply their data science skills to address social challenges facing the campus and regional communities. DSU website: http://hbcu-dsc.org/

Joel Herndon is the Director of the Center for Data and Visualization Sciences (CDVS) in Duke University Libraries, Durham, NC. The centre provides data science, data visualization, data management and digital mapping consulting, instruction and project engagement for the Duke community. The CDVS staff are part of the Duke Libraries and report through the Vice Provost for Library Affairs to the Office of the Provost. DSU website: https://library.duke.edu/data

David Edwards is a Professor of Statistics and Chair of the Department of Statistical Sciences and Operations Research (SSOR) at Virginia Commonwealth University. SSOR houses the Statistics and Analytics Consulting Lab, a unit that provides consulting services to enhance the educational experiences and research projects of university faculty and students. The lab is directed by a tenured faculty member and involves one to two PhD students each semester to assist with consulting services. DSU website: https://ssor.vcu.edu/consulting/

Gina-Maria Pomann is an Assistant Professor of Biostatistics and Bioinformatics and the Director of the Biostatistics, Epidemiology and Research Design (BERD) Methods Core at Duke University, Durham, NC. The BERD Methods Core is a team of over 40 student, staff and faculty biostatisticians and data scientists in the Department of Biostatistics and Bioinformatics who collaborate with clinical and/or translational researchers. DSU website: https://biostat.duke.edu/research/biostatistics-epidemiology-and-research-design-berd-methods-core