

Collective Praxis, Collaborative Publishing: The Case of the Data-Sitters Club

LEE SKALLERUP BESSETTE, KATHERINE BOWERS, MARIA SACHIKO
CECIRE, QUINN DOMBROWSKI, ANOUK LANG, AND ROOPIKA RISAM

Abstract: The Data-Sitters Club (<https://datasittersclub.github.io/site/index.html>) is a pedagogical resource consisting of “books” that explore themes and methods from digital humanities and computational text analysis in colloquial and accessible ways, using Ann M. Martin’s Baby-Sitters Club series as the corpus for analysis. In this article, the collective of six scholars who form the project’s core—and who were themselves fans of Martin’s books as teenagers and young adults—elaborate on the ways that the Data-Sitters Club’s composition, content, technical design, workflows, and writing processes push back against the competitive, individualistic tendencies in humanities research that foreclose new possibilities for thinking, working, and creating within the academy while also creating space for, and taking seriously, the importance of failure in iterative research practices. We present the project as one model for an alternative way of imagining scholarly communication that prioritizes the collaborative and therefore a relational, rhizomatic approach to academic life in place of a hierarchical one circumscribed by the practices of the university today. Following calls from critics such as Kathleen Fitzpatrick, Sandy Grande, and Bethany Nowviskie to commit to collectivity, reciprocity, and mutuality above competition and individualism, we put forward the Data-Sitters Club as an example of how scholarly practices might be reimaged through collaborative publishing.

Keywords: Data-Sitters Club, Baby-Sitters Club, digital humanities, computational text analysis, pedagogy

In academic environments, the act of publication is deeply embedded in a system marked by individual competition, precarity in an ever-shrinking professorial job market, and the relentless drive to publish or perish. In the humanities, where single-authored research outputs have long reigned supreme, disciplinary norms converge with the material realities of the academic landscape and labor markets of the 21st century

to disincentivize collaborative research and privilege the lone researcher. But so much is lost in both scholarship and teaching in the university when competition takes priority over collaboration and the individual over the collective. In fact, it compromises the possibility of imagining new forms of scholarly communication and pedagogies—and even new ways of being within the university. These new futures are only possible by creating spaces that honor the messy, speculative, and process-based nature of research. The working practices of the Data-Sitters Club (DSC), the team behind the eponymous digital humanities (DH) project, is one such example of a collective that is trying to reimagine scholarly practices through collaborative publishing.

In the mid-1980s, Ann M. Martin launched the Baby-Sitters Club (BSC), a middle-grade book series that chronicles the adventures of a group of middle school girls living in suburban Connecticut. They have diverse interests but share a love of baby-sitting. Martin (and her team of ghostwriters) wrote hundreds of the wildly popular books across a main series and multiple spin-offs. The franchise spawned everything from a TV series to a feature film to a board game to a fan club. Reprints, a new graphic novel series, and a Netflix deal revived the BSC over the last decade for a new generation, while inciting much nostalgia among its first readers, many of whom are now parents themselves. In the words of Marisa Crawford and Megan Milks, the BSC “birthed an entire generation of loyal—dare we say obsessive—readers.”¹

Among these obsessive readers are this article’s authors, a group of colleagues who work in and around DH, have varying opinions on the methods of computational textual analysis, and share a love for the BSC. Based on these shared interests, in 2019 we launched the Data-Sitters Club, a pedagogical resource intended to explore themes and methods in computational textual analysis in colloquial and friendly ways using the BSC books as the corpus for analysis.² According to Quinn, who was responsible for the great idea to pursue this project in the first place, “The DSC is, in fact, an amazing potential laboratory for computational textual analysis. . . . There’s unexplored material for a whole research agenda in the area of cultural analytics, which could really use a teenage girl lit counterbalance to its superheroes and sci-fi. But what if a project did more than just analyze these books? What if it actually walked through the whole process?”³

In addition to its intended outcomes and value as an outlet for pedagogical research, our collaboration has shown how the Data-Sitters Club’s composition, content, technical design, workflows, and writing processes push back against the competitive, individualistic tendencies in humanities research that foreclose new possibilities for

1. Marisa Crawford and Megan Milks, “Introduction: We Are the Baby-Sitters Club,” in *We Are the Baby-Sitters Club: Essays and Artwork from Grown-Up Readers* (Chicago: Chicago Review Press, 2021), ix.

2. The Data-Sitters Club, <https://datasittersclub.github.io/site/>.

3. Quinn Dombrowski, *DSC #1: Quinn’s Great Idea*, the Data-Sitters Club, November 7, 2019, <https://doi.org/10.25740/jf827gc7731>.

thinking, working, and creating a new mode of university life. It also creates space for and takes seriously the importance of failure in iterative research practices. Our work, therefore, offers one model for an alternative way of imagining scholarly communications that prioritizes the collective and therefore a relational, rhizomatic approach to academic life in place of a hierarchical one circumscribed by the foregoing practices of the university today.

Collaborative Praxis and Redefined “Success”

In her book *Generous Thinking*, Kathleen Fitzpatrick argues that scholars must push back against the inherently competitive, individualistic, and critical nature of scholarship in the academy—and in how academics interact with one another. An environment of this nature, she argues, is inherently antithetical to the very mission of higher education.⁴ With scholarly models that largely reward single-authored work, the humanities are particularly susceptible to a cutthroat, rather than collaborative, research culture. Consequently, Sandy Grande has powerfully proposed that those of us working within higher education must “refuse the university”—resist allowing the “inducements” that university life offers, such as awards, grants, publications, and other rewards, to overdetermine how we work with one another and the communities to whom we are accountable.⁵ Instead, Grande calls us to commit to collectivity, reciprocity, and mutuality.

The Data-Sitters Club is, in many ways, a different paradigm for scholarship that tries to heed Fitzpatrick’s call and Grande’s vision. While all members of the DSC have published in traditional academic genres—such as peer-reviewed, single-authored articles, book chapters, and monographs—we purposefully sought *not* to work in these genres here. Instead, we developed a hybrid genre that is not quite as pithy as a blog post but certainly not as formal as an article or a chapter. We call it a “Data-Sitters Club book.” We were primarily driven by the goals of helping colleagues see the values and challenges of computational textual analysis, making visible the kinds of collaborations that take place in DH research and can be possible in the humanities more broadly, and nurturing the kind of work we enjoyed, even though there is little professional benefit for any of us by the metrics of academic “success.”

Quinn built the project’s first imperfect corpus and also gathered its members: Roopsi, a scholar of DH and postcolonialism who also has encyclopedic knowledge of

4. Kathleen Fitzpatrick, *Generous Thinking: A Radical Approach to Saving the University* (Baltimore: Johns Hopkins University Press, 2019), 1–4.

5. Sandy Grande, “Refusing the University,” in *Toward What Justice? Describing Diverse Dreams of Justice in Education*, edited by Eve Tuck and K. Wayne Yang (New York: Routledge, 2018), 50–52.

the BSC series from childhood reading; Katia, a scholar of Russian literature and genre theorist who spent way more time than she probably should have reading BSC books as a kid; Maria, a children's literature and media studies scholar with DH interests; Anouk, a scholar who works at the intersection of modernism, contemporary literature, and DH; and Lee, a comparative literature scholar and DH specialist. This core group came together first on Twitter, and then on Zoom, and eventually in person through writing these "DSC books"—to date 21 main series guides and five "multilingual mysteries" that chronicle the adventures of the group as it applies DH methods and tools to the BSC corpus. We want to take both the BSC books and our own DSC books seriously but also not take ourselves too seriously. It would violate the spirit and ethos of both sets of books (BSC and DSC) if we were to try to fit this square peg of a subject into the round hole of traditional academic publishing. In creating our DSC books, we were determined to develop a genre reflective of our collaborative practices and the sense of play that animates our work together.

The core Data-Sitters meet synchronously a few times a year, usually over Zoom, to set a loose schedule for the upcoming months. We pitch book ideas (as well as tongue-in-cheek titles based on the original series that go along with them), consider potential collaborators, and discuss any additions we have to our ever-expanding corpus. We also have a WhatsApp group chat where we share our progress (or lack thereof when life gets in the way). One of the challenges is that our group lives in three disparate time zones (separated by an ocean and eight hours), making it challenging to find synchronous times to work together. So, instead, we rely primarily on asynchronous collaboration channels.

Each book is spearheaded by one or two core members, who may or may not have an external collaborator in mind. But even if each book "belongs" to a core member, various members participate in the editing (and even co-writing) process, providing feedback, advice, and even asides that are relevant to the questions that book asks. We initially work in Google Docs to allow for easier collaboration and commentary (Lee began writing this section of our article in Google Docs while Anonymous Unicorn looked on . . . then Katia edited under the benevolent gaze of Anonymous Gopher and Anonymous Otter . . . and now Roopsi is all alone, wishing an Anonymous Nyan Cat would drop by).⁶ This also allows us to easily work across time zones and integrate guest writers.

While lead authors primarily compose each book, we try as much as possible to be non-hierarchical in our approach to collaboration. Expertise and a desire to learn are more important than the degrees or institutional affiliation of a Data-Sitter. And guest status does not preclude joining the larger group; Lee initially was invited as a guest because of her expertise in Québécois literature and translation, but her interests and

6. With an eye to future-proofing this article against technological change, we're referring to the practice within Google Docs for un-authenticated viewers to be represented by a randomly assigned animal, visible at the top of the interface.

curiosity were broad enough that she became a part of the core group. Guest Data-Sitters are typically co-authors on the work, or for those with a more limited engagement, they are credited at the end, as Ann M. Martin would credit ghostwriters.

This project privileges the expertise of guest Data-Sitters who bring valuable perspectives from all ranks of the academy, and this enables us to engage with many topics that extend far beyond the expertise of the core group. Our guests have included graduate students (Annie Lamar); librarians and administrators (Glen Layne-Worthey, Heather Froehlich, Rachael Samberg); undergraduate students (Cadence Cordell, Sathvika Anand); scholars in fields including computer science (Xanda Schofield), DH (Elisa Beshero-Bondar), literature (Mark Algee-Hewitt, Anastasia Salter, Dainy Bernstein, Shelley Staples), and law (Matthew Sag, Erik Stallman); postdoctoral researchers (Isabelle Gribomont); and even primary school students (Sam and Paul Dombrowski, whose user testing of an emulated CD-ROM game from the 1990s provided a perspective that none of the adults on the team could have offered). We also provided mentorship and professional development for Cadence (known to the Data-Sitters as the daughter of a fellow digital humanist, Ryan Cordell), who interned with the DSC through funding provided by Mount Holyoke College. Cadence investigated the Ann M. Martin papers at Smith College—an archive that core team members were eager to explore but had been unable to visit due to our full-time jobs, families, and a pandemic.

The active participation of colleagues, friends, and even a few family members (as in the case of Quinn's children Sam and Paul) in this project reveals that the primary organizing principle of all collaboration is, ultimately, relationships, which can and do frequently buck academic hierarchies in intellectual as well as social life. These relationships are largely the product of years of engagement in the DH community, both in person and—perhaps more crucially—on DH Twitter as it existed in the 2010s. This network exists not as a result of institutional prestige but through ongoing investment in one another's lives and work in this space—that is, from relationality. The community recognition that comes from a high degree of connectedness within this network is in some cases entirely separate from our standing in our respective institutions, where many of us are constrained by staff roles and hierarchies.⁷

7. DH as a subfield is unique in that one can be highly respected in DH but also work in a professional context with significant constraints. For example, a DH-er might, by virtue of not being a faculty member, be limited in their agency, such as autonomy over their work and the opportunity to be a principal investigator (PI) on grants—regardless of their reputations in the field. Or they might work as a faculty member in an institution where there is little support or funding for DH, so their reputation in the field is in spite of their job, not because of it. This has been the case for a number of us. But, inevitably, the DSC, like other DH projects, exists within the context of an academic subfield where reputations matter. Therefore, it is impossible to disaggregate the recognition the project has received and the interest of guest editors in working with us from the composition of our team. We do not underestimate the impact that the composition of the DSC has on attention and collaboration due to reputational capital. However, because we have felt the constraints of our own institutions, we are especially invested in leveraging the impact of our work to make the DSC an expanding network of collaborators, particularly those who may have encountered their own challenges in their workplace.

In giving author credits to all collaborators, no matter what their “status” within academic hierarchies, we are also mindful of differentiating ourselves from the model outlined by Rachel Mann as the prevailing one for faculty-student collaborations, in which graduate students on a DH project do the computational work but are discouraged from publishing findings; often they are credited in footnotes or a mention on a website rather than as an author or co-author.⁸ Because the genres of academic publication most highly valued in the humanities are single-authored articles and books, Mann observes that collaboration as practiced in this model risks sidelining members of the project team who are not faculty in PI or co-PI positions and excluding them from the very scholarship they have contributed to producing. Not only is this inequitable, but it also neglects the social dimension of collaborative working. As Bethany Nowviskie reminds us, DH work happens “within complex networks of human production” that “require close and meaningful human partnerships.”⁹ Despite attention to the cyberinfrastructures and scholarly workflows that underpin it, digital humanists are, Nowviskie suggests, “fashioning ever closer, more intimate and personalized systems of production.”¹⁰ The way the DSC books scrupulously foreground the valued expertise brought by a particular Data-Sitter—rather than the model that Mann holds up to scrutiny—reinforces our shared commitment to the centrality of the social and interpersonal aspects of our work.

The pedagogical content of the project is another way in which the Data-Sitters Club reveals the narrowness of prevailing notions of academic “success.” The scholarship of teaching and learning has not been sufficiently valued as a form of knowledge production in the academy, cast aside for disciplinary content. This is especially ironic since teaching is a core function of universities. In prioritizing collaboration and pedagogy over the inducements of higher education, the DSC project aligns with Henry Giroux’s call to produce forms of education that are “inherently political,” with our practices embracing Giroux’s insistence on the importance of “relentlessly question[ing] the kinds of labor, practices, and forms of teaching, research, and modes of evaluation” that scholarship enacts.¹¹ Furthermore, the content of our books has liberatory aims: to teach students digital literacy, which helps them become what Giroux terms “border crossers,” individuals who have the capacity to “think dialectically, comparatively, and

8. Rachel Mann, “Paid to Do but Not to Think: Reevaluating the Role of Graduate Student Collaborators,” in *Debates in the Digital Humanities 2019*, edited by Matthew K. Gold and Lauren F. Klein (Minneapolis: University of Minnesota Press, 2019), 268.

9. Bethany Nowviskie, “Evaluating Collaborative Digital Scholarship (or, Where Credit Is Due),” *Journal of Digital Humanities* 1, no. 4 (2012), <https://journalofdigitalhumanities.org/1-4/evaluating-collaborative-digital-scholarship-by-bethany-nowviskie/>.

10. Nowviskie, “Evaluating Collaborative Digital Scholarship.”

11. Henry A. Giroux, “Critical Pedagogy in the Age of Fascist Politics,” *Policy and Practice* 37 (2023): 171.

historically.”¹² To accomplish this, the DSC draws inspiration from the affordances of digital scholarship articulated by Nowviskie, looking beyond finished objects and products authored by single scholars, to instead attend to process—the topic of every DSC book. We aim to intervene in what Nowviskie describes as “the systems of production and of reception in which digital scholarly objects and networks are continuously made and remade” and the “evolving and continuous series of transformative processes” in which quality is to be found.¹³ How well we are doing this both within the established DH community—and, importantly to many of us, beyond that relatively restricted group—is something that the DSC is always discussing, rethinking, and attempting to refine.

Critical Defamiliarization by Design

The approaches we embrace in the Data-Sitters Club resonate with Anna Kornbluh’s contributions to the concepts of mediation and immediacy. For Kornbluh, mediated culture has the ability to foster a form of critical defamiliarization—what she terms “imaginatively break[ing] with the merely given”—by drawing attention to how messages are shaped by medium, genre, form, and other ways they are articulated, rather than by their mimetic delivery to a reader or viewer.¹⁴ This attentiveness to mediation contrasts with what Kornbluh terms “cultural immediacy,” a style that surfaces extreme affect, delegitimizes and disavows its involvement in processes of mediation, and, in its place, substitutes the “auto-authority of presence.”¹⁵ For Kornbluh, mediation “evokes the social process of making representation, connections, and meaning.”¹⁶ These processes are also at the heart of what the DSC books seek to do, as they chart the mistakes and the wrong assumptions, the dead ends and the roads-not-taken of doing digital scholarship. This focus on the friction in the process becomes especially important for the whizzy tech that can accompany some DH work: shiny digital things which present themselves to a non-statistically informed reader or viewer as self-evident truth. Such objects do not escape mediation, of course, but they can take on the seductive sheen of immediacy, as Kornbluh defines it, and thereby help to propagate an orientation towards the internet and digital platforms that have become so much a part of daily life that they do not register as needing to be critically appraised or even interpreted.

12. Giroux, “Critical Pedagogy,” 171.

13. Nowviskie, “Evaluating Collaborative Digital Scholarship.”

14. Anna Kornbluh, *Immediacy, or The Style of Too Late Capitalism* (London: Verso, 2024), loc. 15 of 233, Adobe Digital Editions.

15. Kornbluh, *Immediacy*, loc. 55 of 233.

16. Kornbluh, *Immediacy*, loc. 15 of 233.

Style, for Kornbluh, is one of the key ways through which cultural objects can foreground their mediated qualities and thereby alert readers and viewers to the fact that, as mediations, they are “composites of language, composites of images, compositions of meaning, composed ideas.”¹⁷ The highly stylized way in which the DSC books are presented thus functions as a way of convoking, and invoking, the collective of writers behind them: the books thus participate in what Kornbluh describes as the “catalyzing [of] representation itself as a medium of collectivity.”¹⁸ The website where the DSC books are hosted is a relative newcomer in the field of scholarly and academic-adjacent publishing initiatives that are thriving in digital spaces outside mainstream publishing platforms designed to maximize shareholder profits; the scholarly preprint archive arXiv.org, first established in 1991 and widely used in the scientific disciplines, has a longer history of defamiliarizing the way scholarly papers were conventionally published and distributed and thereby serves to make the mediating processes of scientific journals more obvious. As Kornbluh observes, “residual, emergent, and extra forms of cultural production activate rather than evacuate mediation.”¹⁹ We aim to subvert the highly technologized aesthetics characteristic of white, male DH projects, matching ours to the Web 1.0 style of the 1990s when the BSC books were published and deploying design choices that read as “femme” to align with the BSC books’ intended audience. Our candy-pink site background, the playful way we introduce authors, and the “admonitions” when other voices chime in to a book all serve to generate the productive friction of mediation and thereby draw attention to the situated and contingent ways in which books are put together, digital infrastructures are designed and maintained, and knowledge is constructed.²⁰ With the goal of evoking nostalgia for the youth series literature section in a public library from the 1980s and ’90s—complete with its colorful pastel rainbows of book covers—we decided to present the DSC books on the website in a grid, with a prominent visual symbol for each. It seemed like an obvious choice to modify the actual BSC book covers to serve this function. While our earliest books used the classic 1990s American covers, as the project has evolved, we have begun to choose from a wider range of options—including a cover from a graphic novel adaptation, a UK edition cover, and a recent French translation cover—when the details of those images better fit the topic of our writing.

Like the BSC books themselves, which are notoriously formulaic, Data-Sitters Club books are typically structured around a kind of narrative arc: we introduce a situation or a problem, then walk the reader through our attempts to address the instigating issue.

17. Kornbluh, *Immediacy*, loc. 173 of 233.

18. Kornbluh, *Immediacy*, loc. 173 of 233.

19. Kornbluh, *Immediacy*, loc. 173 of 233.

20. We use Markdown to publish the prose portions of the Data-Sitters Club books; in that jargon, “admonitions” are call-out boxes.

Sometimes we step away for an interlude where we directly confront our frustrations and failures along the way or try to explain technical or legal concepts that are at play. Finally, we wrap things up with some kind of conclusion—which is often, in reality, a non-conclusion gesturing towards the challenges of doing this kind of work or other paths we could explore. How much time we devote to each step of this journey varies wildly, and some books diverge entirely from this format, such as *DSC #13: Goodbye, Friends, Goodbye*, in which we memorialize three recently deceased colleagues through an offering of memories from different Data-Sitters. As a result of these different narrative journeys, our books vary significantly in length (fig. 1), from a little over 3,000 words (*DSC #1: Quinn's Great Idea*, *DSC #3: The Truth About Digital Humanities Collaborations (and Textual Variants!)*, and *DSC #M1: Lee and the Missing Metadata*) at the shortest to nearly 50,000 words (*DSC #9: The Ghost in Anouk's Laptop*). For our longer books (those over 12,000 words), the extended word count is due to excerpts of text generated by a large language model (*DSC #9*) or the output of code we have run. Publishing our work on our own website, in a form that we have devised ourselves, means that we can write what we feel moved to write, without angsty over how it aligns with a journal's accepted genres and word counts—or the degree to which code and its output can be supported alongside academic prose. At the same time, recognizing that this freedom for us might limit our readership and accessibility to some of the DH-curious but nonexpert folks we would also like to reach, self-publishing online gives us the ability to pilot new forms of communicating our work, such as the new TL;DR series that the DSC started in 2024.

Across all of the DSC books, we strive to make the scholarship we are doing, as well as our techniques and approaches, legible. As Cadence puts it in *DSC #17: Cadence's Archives Mystery*, “But I also think a lot of first-time archival researchers aren't prepared for what archival research will be like—in my experience, research seminars for undergrads typically help students learn to research for secondary sources in a library's collection, and don't include any perspective on what archival research might be like.”²¹ Expand that to any DH approach and anyone new to the field. We aim to make visible and accessible not just what we discovered, but how we did it. In all of our books we make very plain the processes we followed, step-by-step, to get where we were trying to go. Sometimes those processes are very technical (see *DSC #4: AntConc Saves the Day!*), but just as often, they are very personal (see *DSC #13: Goodbye, Friends, Goodbye*). Very often the path or processes we follow lead . . . nowhere, or at least not at all where we thought they would.²² But even these destination-less journeys may have value for some

21. Cadence Cordell, *DSC #17: Cadence's Archives Mystery*, the Data-Sitters Club, September 21, 2022, <https://doi.org/10.25740/sd796vb8535>.

22. In *DSC #M4: Isabelle and the Missing Spaghetti-O's*, our early efforts to train an entity recognition model on food failed, leading us to resort to WordNet word lists instead. See Lee Skallerup Bessette, Quinn Dombrowski, and Isabelle Gribomont, *DSC #M4: Isabelle and the Missing Spaghetti-O's*, the Data-Sitters Club, July 22, 2021, <https://doi.org/10.25740/vz142ty4818>.

Word count of Data-Sitters Club books



Figure 1. Word counts of Data-Sitters Club books.

readers, normalizing the reality that even though sometimes there *is* no output, we grow as scholars in the process undertaking the work.

Genre-Bending and Its Uses

We are scholars who are parents, coaches, teachers, horse girls, sewists, weavers, musicians, polyglots, activists, and incorrigible readers. All those parts of our identities contribute to the processes that give rise to our work. While not every facet of our selves comes through in each DSC book, we work to preserve our unique voices, make plain our biases and blind spots, and let our enthusiasm (and disappointment!) shine through. For this reason, we adopt the BSC series model of identifying each voice in a book: to not only make visible those differences and strengths but underscore them. In scholarship since the 1990s, especially in feminist and queer theory and the qualitative social sciences, it has become common for researchers to identify their positionality in relation to their topic(s) of study, providing human context for their entryway into

(and pathways through) a given project. Instead of claiming an impossible objectivity through silence on these matters, addressing positionality shines light on the person and conditions behind the work—as well as the debates and tensions within our collaboration that can be productive in their own right. The tensions are real.²³ Roopsi and Maria, the two women of color and least enamored with computational textual analysis methods, are prone to raising concerns about the research questions we pursue and, at times, the conclusions we draw, particularly regarding race. As we try to negotiate the tensions productively, we do so guided by the idea that identifying the social and political contexts that produce our identities, interests, methodological preferences, and access of various kinds strengthens rather than undermines scholarship.

How we approach the DSC books, and what we ultimately end up writing about, defies genre. We talk about DH methods, pedagogy, power structures, collaboration, material history, processes, and failures, as well as our own experiences and connections to the topics and each other. Such topics are not always welcome in traditional academic publishing, but this has made the experiment of the DSC all the more personally and intellectually valuable to the group. Such freedom from the constraints of genre also applies to the media we use to communicate our work: in addition to words (whether natural language or coding language), the DSC books incorporate a number and variety of visuals that would not be possible in print journals. In addition to the screenshots and data visualizations that one would expect from a DH pedagogy resource, childhood and college-era photos, pictures of us from our Zoom meetings, animated GIF memes, screenshots of social media posts, and other images contribute to text that takes on a diary or scrapbook vibe. This collecting reinforces a broader point that we try to make throughout the DSC book series: that humanities research, including computational text analysis, is fundamentally *personal*.

The nature of some of these images—particularly the animated GIFs—is not without consequences. Much as we relish the digital publication platform as writers, we acknowledge that it is not always the best medium for learners. When a colleague asked for book recommendations for a university DH center’s collection of pedagogical materials, Quinn quipped, “You could print out the Data-Sitters Club.” That initial moment of sarcasm rapidly grew into the seed of a new sub-project: reimagining DSC books as zines that could reach different audiences than the web version. While the zine initiative is still in its early phase, it has already led to some important reflection on the components of DSC books—their length, mix of code and prose, color images, and inherently digital elements like the animated GIFs—and how these would need to be transformed

23. See Lee Skallerup Bessette et al., *DSC Super Special #1: The Data-Sitters Debate at Dartmouth*, the Data-Sitters Club, October 2, 2024, <https://datasittersclub.github.io/site/dscss1.html>, for a near-transcript of the group having one of these discussions in its first in-person meeting.

for compatibility with a low-cost printed medium like the zine. DSC books also share some political resonances with zines, which, as Janice Radway observes, are “not only about writing and reading but also about community formation and social intervention,” themes that emerge with regularity in DSC books as well.²⁴ However, it is in their use, circulation, and connective countercultural powers that zines express their full community-creating abilities; we have yet to see where our DSC books and printed zines will go, who will take them up, and what they will do with them from there.

Technical Specifications as Responsive Mediation

On a technical level, each DSC book is a Markdown file with text and images (for books with no code) or a Jupyter Notebook that interweaves text, images, code, and code output, which is then transformed into HTML. These web pages have the trappings of scholarly objects: each has an assigned DOI with associated metadata, and a PDF version is stored in the Stanford Digital Repository within a collection for the DSC that itself has a MARC record available via WorldCat.²⁵ Most of the books are polyvocal: even when there is a single primary narrator, we use “admonitions” to insert commentary or side points from other collaborators. Sometimes there is dialogue with a guest Data-Sitter; other times, there is an unmediated look into our process, either with screenshots of comment threads from the Google Docs where we draft our work, or near-verbatim transcripts of conversations we’ve had as a group. For books with multiple narrators, we typically center those individual voices by treating the narrator as the highest-level division within the text, with topical sections and subsections nested below.

While some aspects of our website—such as the gallery view of books—have remained constant, changes to the technical underpinnings have shifted what a Data-Sitters Club book is, moving it subtly from something in the realm of a blog post to something closer to a chapter in a textbook crossed with a diary entry. The DSC began as a Jekyll site (built on a workflow for static site generation), but after *DSC #8: Text-Comparison-Algorithm-Crazy Quinn* required the extensive use of Python code, it became clear that we needed a more robust platform for publishing texts in the spirit of Donald Knuth’s *Literate Programming*. As Knuth describes it:

24. Janice Radway, “Zines, Half-Lives, and Afterlives: On the Temporalities of Social and Political Change,” *PMLA* 126, no. 1 (January 2011): 142, <https://doi.org/10.1632/pmla.2011.126.1.140>.

25. See <https://purl.stanford.edu/cp667df5882>; and <https://search.worldcat.org/title/1426042415>.

Let us change our traditional attitude to the construction of programs: Instead of imagining that our main task is to instruct a computer what to do, let us concentrate rather on explaining to human beings what we want a computer to do.

The practitioner of literate programming can be regarded as an essayist, whose main concern is with exposition and excellence of style. Such an author, with thesaurus in hand, chooses the names of variables carefully and explains what each variable means. He or she strives for a program that is comprehensible because its concepts have been introduced in an order that is best for human understanding, using a mixture of formal and informal methods that reinforce each other.²⁶

Jupyter Book had already gained traction in data science circles as exactly such a platform for turning Jupyter notebooks into publications that could be viewed and downloaded as both executable code and printable PDFs. It was a good fit for our needs, with one exception: relatively weak documentation for how to modify the aesthetics of the resulting website. Keeping our pink color scheme was non-negotiable. With some assistance from Philip Allfrey, a digital humanist and book historian, we were able to style our Jupyter Book, and as of this writing, the DSC is the most colorfully implemented project in the Jupyter Book gallery. These changes are driven by a commitment to responsive mediation, designed to respond iteratively to content, form, and audience.

Failure as Method

One product of our public process is that we often write about failure. When we do research, often the messy drafts, the attempts to get the argument just right through analysis, are hidden by the polished façade of the publication that appears in print. Some element of trial and error is always part of DH methods. The step-by-step nature of the DSC book medium means that our failures, as well as our successes, are, by necessity and by design, part of the process and therefore part of the publication.

DSC #11: Katia and the Sentiment Snobs presents a good example of the way failure can productively inform practice.²⁷ In *DSC #11*, Quinn and Katia set out to understand whether sentiment analysis is a good method for literary studies. The book begins with a conversation between the two. Katia, a relatively new digital humanist, had heard that sentiment analysis could be useful in analyzing some aspects of her corpus, while

26. Donald E. Knuth, "Literate Programming," *Computer Journal* 27, no. 2 (1984): 97–111, <https://doi.org/10.1093/comjnl/27.2.97>.

27. Katherine Bowers and Quinn Dombrowski, *DSC #11: Katia and the Sentiment Snobs*, the Data-Sitters Club, October 25, 2021. <https://datasittersclub.github.io/site/dsc11.html>

Quinn, who has significantly more experience and disciplinary history knowledge, cautioned against using the method based on issues surrounding its use within the history of DH.²⁸ Together, they decided to determine whether these issues are still relevant by writing a DSC book using the method.

DSC #11 details their attempts to analyze several BSC books using sentiment analysis. Through this process, they experienced one failure after another. First, they tried out existing tools for doing sentiment analysis but discovered that these tools did not return results with enough nuance to be useful. Digging deeper, they discovered that the issue lurked in the lexicon scores assigned by these tools, but maybe, Katia and Quinn thought, the issue was that those tools are not designed to analyze literary texts. They turned to Syuzhet, a package specifically designed to examine sentiment in connection with plot in literary works. Syuzhet gave more data than the numbers generated by the first tools they tried, but, as Quinn wrote in *DSC #11*:

Syuzhet offers the humanities scholar an interpretive path forward from those numbers. And you might feel so relieved to have some framework for taking the next step with your analysis that you don't think too hard about the sleight-of-hand it uses to get there. That sleight-of-hand is the phrase "sentiment-derived plot arc." There's an implicit claim here that sentiment scores, like the ones we've been looking at, are a plausible—even a *good*—way to derive a plot arc. If you've been following along this far, you should be feeling uneasy after seeing how those sentiment scores (often fail to) capture the actual sentiment in a text. But those terrible sentiment scores aren't even the biggest problem here: even if the sentiment scores were better, what is the connection between sentiment and plot anyway?

The further Katia and Quinn went down the sentiment analysis method path, the more failure they faced. At this point, three tools had failed, the lexicon those were based on had failed, and now it seemed like even the method itself was prone to failure. But was this failure the fault of the tool or the method?

Quinn and Katia turned to Mark Algee-Hewitt (given the DSC honorary title of Associate Member after the BSC's parlance) for advice. Sagely, he told them, "There's some good work going on in things like event and scene detection that are plot related. . . . But this is a moment where I think that common critique of DH is right: plot is one of those things where it's just far more complex than the proxies we're using. Sentiment analysis is much too blunt a tool."

28. This history is summarized more fully in *DSC #11*. In short, the Syuzhet sentiment analysis and plot arc tool generated a dialogue, which Matt Jockers and Annie Safford conducted in a series of blog posts in 2015. This exchange was pivotal in shaping the discourse around this method for a certain generation of DH scholars.

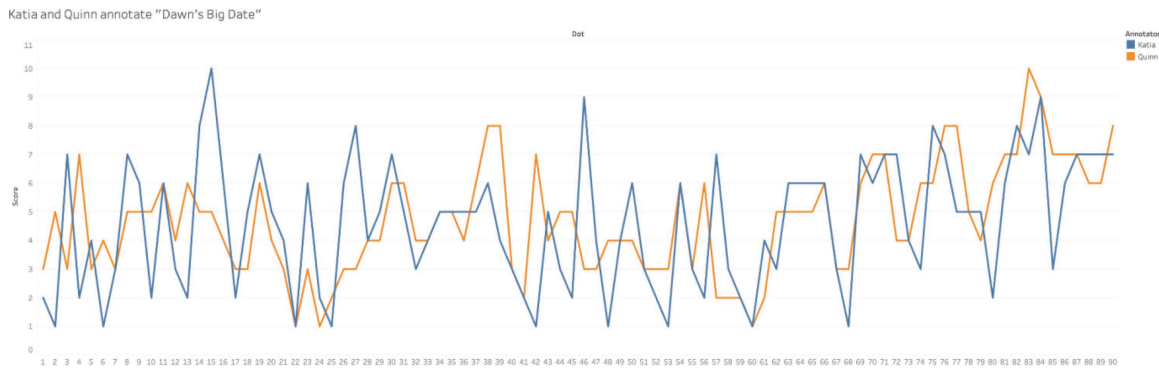


Figure 2. Visualization of hand-coded sentiment scores of *Dawn's Big Date*.

Mark was wise, but Katia and Quinn were still curious about the method. By this point in the process, all the tools had failed, so they set off to create their own sentiment arcs. By trying the method out by hand, they could eliminate any clouding of the results caused by the automated sentiment analysis tool. Seeking to substitute human reading for machine reading, Quinn and Katia read and “scored” chunks of text for several books in the series, providing sentiment scores (positive, negative, neutral) on a comparative scale with at least six scored points in each book’s 15 chapters. They then plotted their respective sentiment arcs on a graph (fig. 2). Looking at one of them, *Dawn's Big Date*, they thought they had a similar, possibly viable result.

Figure 2 maps Quinn and Katia’s hand-coded sentiment arcs over each other to visualize their similarity. They seem similar, but if you look more closely, you see that sometimes they diverge drastically, where one Data-Sitter’s very highest score on the chart will match the other’s lowest. They realized after another talk with Mark that this result was actually symptomatic of yet another failure. Quinn reflected:

At that moment I realized I couldn’t have my impressionistic, squishy, human-reader cake and eat it computationally, too. I was trying to do this graph the human way, with just a little bit of concession to formal constraints (by having a consistent number of squares for each chapter). And there’s nothing inherently wrong with doing things that way! But how you collect your data *always* has implications for your analysis. It’s not that you can’t do *any* kind of interpretation or analysis of the kind of squishy human-interpretation graph we put together—I did! I looked at it with my eyeballs and, ignoring things that I evaluated to be unimportant noise, determined that Katia and I drew similar graphs. But just like how you can’t meaningfully use methods that require a *lot* of text (like word vectors) on a single novel, you can’t apply quantitative methods to graphs like the ones we created.

Quinn's realization of failure in real time in the text of the book is significant. This moment of failure is equally a moment of revelation, the moment that the problem Quinn and Katia set out to address is resolved. Its resolution, of course, is failure. They did not find the sentiment analysis tool sufficiently useful for literary analysis. But through failing, we learn. Katia and Quinn came to the conclusion that it was not just that the tools gave insufficient nuance, but that the method itself was flawed for literary text analysis. Sentiment analysis might be a good method for doing work in some fields, but literature is not one of them. Katia summarized the takeaway from this experience: "Tagging individual words as positive or negative gets you into some subjective weeds, using those words as a model to tag other words gets trickier still, and then applying that model to something as complex as a novel, or even a story, becomes meaningless."

DSC #11 is a story of failure upon failure in pursuit of a method. In the end, Katia and Quinn learned that the method was not a good one for their needs. Often in DH, failure is a part of the process, but failure is not documented as often as success. Documenting failure, and incorporating it into our praxis as Data-Sitters, is important. *DSC #11* would not exist if we did not document our failure, and its loss would mean a lacuna in the scholarship around sentiment analysis, one that conceals our history of failure and, equally, our learning through failure. While reflecting on *DSC #11*, we made the welcome discovery that others have found value in our account of failure; the book has already garnered several citations from other scholars who use it in their work to support assertions about the inaccuracy of sentiment analysis for literary studies.

Final Thoughts

We offer the case study of the Data-Sitters Club in this article as a deep dive into the practices, processes, and messiness behind our collaborative work, with the hope that we can inspire others to invent their own creative approaches to collective knowledge creation. We cannot say that we have definitively cracked the code to collaborative praxis (if there even were such a code to begin with) or revolutionized scholarly communication. But we have created space for experimentation and play that gives greater meaning to our other scholarly endeavors and brings the kind of shared nerdy joy that made us want to become academics in the first place. Ironically, this nonstandard way of working has resulted in our group being invited to give talks, selected to appear on conference panels, accepted for publication in academic and nonacademic venues, and even recognized for our project with awards—unanticipated and, yes, CV-building outcomes that emerged from an effort initially understood as resisting the imperatives to reproduce the conditions of production within higher education. At least as gratifying

is the knowledge that our DSC books have been taught across many DH classrooms, introducing another generation of humanists to our collaborative, informal, and (we think!) fun way of working.

And yet—fittingly, given the spirit in which this project was conceived—the academic impact of the DSC is less immediately important to us than the connections it has forged between a group of friends who have seen one another through many personal and professional challenges over the years of this project. We recognize that our ability to come together in this way is facilitated at least in part by the various kinds of relative privilege that we have, foremost among which is stable employment within the academy. At the same time, we hope that our example will—both in spite and because of our own failures, false starts, and descents into rabbit holes—encourage other scholars to invent new approaches to research and, in doing so, create opportunities for collaboration with students and colleagues from an array of positions inside and outside of higher education.

Author Biographies

Lee Skallerup Bessette is Assistant Director for Digital Learning at Georgetown University and an affiliated faculty member in the Master's program in Learning, Design, and Technology. She has been teaching in higher education for over 15 years, primarily at regional, teaching-centered institutions that serve non-traditional students. Her current work in faculty development and technology focuses on the pedagogical and curricular implementation of digital fluency. She blogged at [InsideHigherEd.com](https://www.insidehighered.com) (College Ready Writing) and was also a regular contributor at ProffHacker. You can find out more about her at her website readywriting.org.

Katherine Bowers is an Associate Professor of Slavic Studies at the University of British Columbia. She holds a PhD in Slavic Languages and Literatures from Northwestern, and her research expertise is in Russian literature and culture, particularly that of the nineteenth century. Her research interests include genre, narrative, environmental humanities, imagined geography, and digital humanities. She is the author of *Writing Fear: Russian Realism and the Gothic* (2022) and has co-edited several volumes, most recently *The Oxford Handbook of Global Realisms* (2025).

Maria Sachiko Cecire is a program officer for Higher Learning at the Mellon Foundation and an associate professor of literature (on leave) at Bard College. She was the founding director of Bard's Center for Experimental Humanities, which focuses on how technologies mediate the human experience, and is author of *Re-Enchanted: The*

Rise of Children's Fantasy Literature in the Twentieth Century (2019). Maria has a BA in English language and literature from the University of Chicago and an MSt in English medieval studies and a DPhil in English from the University of Oxford, where she was a Rhodes Scholar.

Quinn Dombrowski is the Academic Technology Specialist in the Division of Literatures, Cultures, and Languages, and in the Library, at Stanford University. Quinn has a BA/MA in Slavic Linguistics from the University of Chicago, and an MLIS from the University of Illinois at Urbana-Champaign. Quinn is also the director of the Textile Makerspace, and advocates for better support for non-English DH projects.

Anouk Lang is Senior Lecturer at the University of Edinburgh, where she teaches twentieth- and twenty-first century literature and digital humanities. She is the editor of *From Codex to Hypertext* (2012), and co-editor of *Patrick White Beyond the Grave* (2015) and *Digital Futures of Graduate Study in the Humanities* (forthcoming in 2024).

Roopika Risam is Associate Professor of Digital Humanities and Social Engagement at Dartmouth. Her research focuses on data histories, ethics, and practices at intersections of postcolonial and African diaspora studies, digital humanities, and critical university studies. Risam is the author of *New Digital Worlds: Postcolonial Digital Humanities in Theory, Praxis, and Pedagogy*, and co-editor of multiple volumes, most recently *Anti-Racist Community Engagement* (2023) and *The Digital Black Atlantic* (2021).

References

- Bessette, Lee Skallerup, Katherine Bowers, Maria Sachiko Cecire, Quinn Dombrowski, Anouk Lang, and Roopika Risam. *DSC Super Special #1: The Data-Sitters Debate at Dartmouth*. The Data-Sitters Club, October 2, 2024. <https://datasittersclub.github.io/site/dscss1.html>.
- Bessette, Lee Skallerup, Quinn Dombrowski, and Isabelle Gribomont. *DSC #M4: Isabelle and the Missing Spaghetti-O's*. The Data-Sitters Club, July 22, 2021. <https://doi.org/10.25740/vz142ty4818>.
- Bowers, Katherine, and Quinn Dombrowski. *DSC #11: Katia and the Sentiment Snobs*. The Data-Sitters Club, October 25, 2021. <https://datasittersclub.github.io/site/dsc11.html>.
- Cordell, Cadence. *DSC #17: Cadence's Archives Mystery*. The Data-Sitters Club, September 21, 2022. <https://doi.org/10.25740/sd796vb8535>.
- Crawford, Marisa, and Megan Milks, eds. *We Are the Baby-Sitters Club: Essays and Artwork from Grown-Up Readers*. Chicago: Chicago Review Press, 2021.
- Dombrowski, Quinn. *DSC #1: Quinn's Great Idea*. The Data-Sitters Club, November 7, 2019. <https://doi.org/10.25740/jf827gc7731>.
- Fitzpatrick, Kathleen. *Generous Thinking: A Radical Approach to Saving the University*. Baltimore: Johns Hopkins University Press, 2019.

- Giroux, Henry A. "Critical Pedagogy in the Age of Fascist Politics." *Policy and Practice: A Development Education Review* 37 (2023): 159–75.
- Grande, Sandy. "Refusing the University." In *Toward What Justice? Describing Diverse Dreams of Justice in Education*, edited by Eve Tuck and K. Wayne Yang, 47–65. New York: Routledge, 2018.
- Knuth, Donald E. "Literate Programming." *Computer Journal* 27, no. 2 (1984): 97–111. <https://doi.org/10.1093/comjnl/27.2.97>.
- Kornbluh, Anna. *Immediacy, or The Style of Too Late Capitalism*. London: Verso, 2024.
- Mann, Rachel. "Paid to Do but Not to Think: Reevaluating the Role of Graduate Student Collaborators." In *Debates in the Digital Humanities 2019*, edited by Matthew K. Gold and Lauren F. Klein, 268–78. Minneapolis: University of Minnesota Press, 2019.
- Nowviskie, Bethany. "Evaluating Collaborative Digital Scholarship (or, Where Credit Is Due)." *Journal of Digital Humanities* 1, no. 4 (2012). <https://journalofdigitalhumanities.org/1-4/evaluating-collaborative-digital-scholarship-by-bethany-nowviskie/>.
- Radway, Janice. "Zines, Half-Lives, and Afterlives: On the Temporalities of Social and Political Change." *PMLA* 126, no. 1 (January 2011): 140–50. <https://doi.org/10.1632/pmla.2011.126.1.140>.

