Designing Digital Repositories: User Centered Design Thinking and Sustainable Professional Development

Hadi Riad Banat, Emily Palese, Hannah Morgan Gill, Shelley Staples, and Bradley Dilger

Learning to teach writing remains challenging for both student instructors and the experienced faculty, staff, and other professionals who mentor them. Many institutions have created local repositories of teaching-related materials such as syllabi and assignments, and some national efforts have also attempted to facilitate sharing resources to support instructor professional development. However, the labor-intensive nature of building digital resources has reduced the long-term effectiveness of these archives. Interdisciplinary approaches to building professional development resources can address these difficulties, creating more sustainable repositories and similar professional development resources. In this article, we describe the construction of the repository associated with the Corpus & Repository of Writing, a web-based archive of pedagogical materials from first year-writing courses at multiple institutions. We discuss the decisions we made regarding material and metadata collection, data processing, classification of materials, and interface design, and share insights from testing with administrators, teachers, and mentors keen on developing sustainable, equitable approaches to mentoring. By drawing from the methods of corpus linguistics and technical communication, and guiding our development processes with user-centered design, we offer a model for repository building that can shape mentoring, teaching, and research within and across diverse writing programs.

Recent scholarship and the continuous, collaborative efforts of professional organizations in rhetoric and composition have continued the call to prepare college writing teachers for the evolving demands of the profession. Because effective writing instruction can promote students' success in academic, professional, and public settings, "an investment in the training and professional development of writing instructors is an investment in student learning and success" (CCCC). Rhetorical knowledge, linguistic knowledge, instructional knowledge, knowledge of ethical research practices, and technical knowledge must be prioritized when professionalizing college writing teachers. While professional development can help writing teachers build knowledge, inequities in institutional resources can pose challenges towards

achieving sustainable mentorship. To address such barriers, Shelley E. Reid calls for creating online resource clearinghouses and pursuing research on writing pedagogy education, emphasizing that "few studies on writing pedagogy education are data driven, longitudinal, or inclusive of more than one program" (692). The lack of a systematically-produced body of knowledge on mentoring writing teachers creates an exigence for making the intensive intellectual labor of mentoring more visible. Enhancing the visibility of mentoring can take place through building accessible online resources and archives as well as producing practice-based scholarship. We take up Reid's charge by describing how we are creating a resource that moves beyond local situatedness and is documented in multiple venues. This resource, which supports the collection of datasets of pedagogical materials that reflect diverse writing cultures across institutions, will contribute to transforming the ethos of mentoring. Mentoring can then become more informed by diverse program representations, student and teacher populations, and various institutional realities. Thus, building and developing archival projects through cross-institutional collaboration can make mentoring more inclusive and sustainable.

To make the intensive work of building archives less burdensome, writing scholars keen on developing sustainable mentoring interventions should invest in cross-institutional collaboration. To make archives inclusive of diverse audiences, user experience design (UXD) can effectively inform building collaborative structures, processes, and teams. We have created a repository of pedagogical materials for first year writing (FYW) courses across two institutional sites, Purdue University and the University of Arizona, that responds to two specific needs: one, the need for building sustainable archives to facilitate ongoing professional development, and two, to provide resources for research on writing teacher preparedness. This archive of pedagogical materials is part of a larger project, the Corpus & Repository of Writing (CROW), which is building a corpus of FYW student texts alongside the teaching repository we discuss in this article (Staples and Dilger). In discussing the repository of teaching materials, we highlight our repository-building methods informed by corpus approaches and design thinking, as well as the approaches to collaboration we have adopted to build the repository and our team. We show how our interdisciplinary collaboration (corpus linguistics, technical/professional writing, and information sciences) with writing program administrators, writing teachers, and writing mentors running teacher practicums has helped us to make decisions pertaining to material and metadata collection, material processing and de-identification, archival classification systems, and interface design.

Addressing Sustainability Concerns in Building Archival Projects

User-centered design advances studying different populations of users to promote designing digital infrastructure that caters to the needs of various stakeholders in different contexts. As we began user experience research for the purpose of CROW development, we conducted environmental scans of existing projects and tools like Pedagogy Toolkit, the Corpus of Contemporary American English (COCA), Iowa State's Digital Repository for Academic Writing (DRAW), the Compleat Lexical Tutor (Lextutor), the Michigan Corpus of Upper-Level Student Papers (MICUSP), Sketch Engine, and Talk Bank. These environmental scans allowed us to trace the evolution of projects as well as examine hurdles that risk the sustainability of archival projects. Some notable challenges for sustainability included researchers changing jobs, limited numbers of researchers on teams, and the need for ongoing repository maintenance. For example, though Pedagogy Toolkit is still available, including its source code on GitHub, it has not been updated for eight years (Christie).

Even well-resourced institutions have encountered challenges related to sustainability. For example, the in-house repository for writing programs at the University of Arizona still exists, but because of the labor-intensive nature of keeping resources active and up to date, repository updates are typically sporadic, led by *ad hoc* committees of graduate students and faculty, and often limited in scope. Similarly, DRAW has been periodically active since its inception in 2012, but the recurrence of similar teaching materials and activities in their repository suggests their maintainers have been unable to sort and classify contributions. Local repositories have immediate value and benefit to the population of writing teachers at individual institutions, yet they are hard to maintain when the responsibility of keeping them alive is not formally assigned or when such labor is mostly done by students who eventually graduate and shift institutions.

The challenging conditions of maintaining in-house local repositories necessitate building more sustainable infrastructure in the form of cross-institutional archives. Team-based projects such as the Meaningful Writing Project, the Teaching for Transfer project, the Citation Project, and the International Corpus of Learner English (multilingual learner corpus research) suggest the merits of cross-institutional collaboration for building balanced and sustainable teams. Lessons from environmental scans helped us consider how to future-proof CROW's infrastructure as well as how to design a functional model for collaboration within our team: as a research and development site that includes faculty, graduate, and undergraduate students whose professional realities change with time.

The main outcome of environmental scans was educating ourselves about integrating user-centered design approaches across our interdisciplinary team. This can happen through leaning on a wider network of expertise, for example purposefully building teams that involve undergraduate and graduate students, writing teachers and mentors, and writing researchers and WPAs from multiple institutions. The collaboration model the CROW team has adopted aligns with Beth Brunk-Chavez, Stacey Pigg, Jessie Moore, Paula Rosinski, and Jeffrey T. Grabill's "hub-and-cluster network shape" that forms a network of interconnected nodes through the institutional partners invited to contribute to the repository (87). Our current repository includes two institutional hubs, Purdue University and University of Arizona, that are heavily involved in: 1) designing and maintaining the digital infrastructure, 2) planning strategic outreach that builds and connects networks of future partners, and 3) forming teams with a wide range and representation of interdisciplinary expertise. The team supporting the maintenance of CROW's repository has researchers at multiple institutions beyond the hub. Our team communication platform, Basecamp, and our strong documentation protocols facilitate collaboration and communication among researchers within our network. There are important benefits in such a model, like "pooling knowledge or expertise, bridging space, saving time, increasing innovation, and taking advantage of economies of scale" (Brunk-Chavez et al. 85). These benefits make such collaborative models economically viable and ensure long-term sustainability.

Building and operating such a team requires ongoing rhetorical negotiations because the process is not linear but iterative. Jason Swarts uses the term "network" to demonstrate how the interconnected nodes establish connections that generate "a feedback loop spurring the continued innovation, development, distribution, and integration of information generating, processing, and networking technologies" (120–21). The ongoing feedback we receive in CROW—through organized communication across sites, user testing, and outreach to new collaborators—allows continual improvements to the classification system of the archive, the development of automated data collection methods, and the overall interface design. These improvements, in turn, are tested as new institutions enter the network.

However, we recognize the difficulty of building and sustaining complex, diverse teams. Our "constructive distributed work" project is developing an integrated approach to team building that balances individual and team success (McMullin and Dilger). Michelle McMullin and Bradley Dilger explain: "Thinking three-dimensionally about all CROW activities—research design, software development, team building, infrastructure maintenance, and more—helps us to continually attend to sustainability. More important, it ensures our foundational commitment to ethical practice in both research

and professional development for faculty and students" (471). Understanding our collaboration as building a mentoring network facilitates collaboration logistics across institutions, exchange of expertise, coordination of efforts, and equitable distribution of intellectual labor among all team members: students, writing teachers, and administrators.

Teaching Material Archives at the Nexus of Praxis and Disciplinary Knowledge

This visionary and deliberate move is also informed by scholarship in writing program administration which encourages WPAs to adopt what Shirley Rose and Bud Weiser call a "researcher and archivist" mindset and to make sustainable long-term plans for creating, preserving, and studying teaching material archives (276). Such work can offer insight into how instructors design tasks and communicate expectations (Miller, Mitchell, and Pessoa; Palese). Additionally, knowledge gained from building and researching teaching archives can support the development of inclusive mentoring practices. Teaching material archives also promote the integration, customization, and accessibility of continuous professional development. Moving from strictly local institutional repositories by creating outreach networks for sharing these initiatives makes these efforts become better investments. Repository access and resource distribution across institutional sites become more visible and are promoted across multiple institutions, programs, and work sites. That is, building cross-institutional repositories like CROW's not only addresses local needs but creates widely available datasets for research on writing pedagogy and resources for writing teacher education.

Building pedagogical archives is a writing program administration commitment to advance and sustain professional development. This initiative is a strategic move to transform pedagogy into practice-based scholarship which further contributes to defining the field of writing studies. Rose and Weiser emphasize the commitment institutions have in "the stewardship of these digital materials, including long-term preservation where appropriate as well as organization and access or distribution" (329). Such work is not limited to archiving but employs strategies for creating visible and accessible infrastructure for promoting long term sustainability. Creating digital archives further facilitates the research dimension of WPA work to: 1) describe curriculum and classroom contexts with the use of data, 2) validate curriculum design and teaching practices by data driven assessment, and 3) produce new knowledge about mentoring and classroom learning.

In arguing that how digital archives shape the identity of a discipline for itself and for external audiences, James Purdy encourages us to "consider carefully what texts we save, how we organize them, and to whom we make them

available" (35). He invites us to reflect on the role writing studies teacher-scholars can play in building and using repositories for our classes, writing programs, and discipline, emphasizing that "our pedagogy, scholarship, and disciplinary identity are inextricably bound up in the digital archives we use today and design for the future" (27). Derek Mueller shows how collecting well-curated data presents "augmentative forms of evidence to cases grounded in local experiences" (159) and helps us build networks among "situated experiences and context-specific engagements [that] remain essential to disciplinary knowledge, action, and participation" (172). Similarly, we argue that building archives of teaching materials enhances the epistemological credibility of mentoring practices and writing teacher preparedness. Our argument illustrates how Purdy's three gifts of the digital archives—"integration, customization, and accessibility" (28)—apply to CROW because of our user-centered approach:

1. Integration: Teaching and mentoring are separate but interrelated activities, so digital repositories can promote the integration of these two scholarly activities. Pedagogical material archives have a great value for instructional mentorship of writing teachers. By using such archives, mentors and teachers can engage with authentic classroom materials.
2. Customization: Digital archives allow for customizing classification systems. CROW's classification practices are shaped by corpus-informed scholarship in applied linguistics and the available materials and metadata we gather from our sites (Kwon, Partridge, and Staples). By relying on CROW's interdisciplinary expertise in corpus linguistics, information science, and technical communication, we can tailor our classification systems to promote access and meet diverse user needs.
3. Accessibility: Projects like CROW reduce temporal and spatial obstacles to both research and professional development. The repository provides access to pedagogical materials collected from different institutional spaces and time periods in various writing programs. This archival characteristic allows for studying how writing programs are changing and how courses, curricula, and the genres of assignment design develop accordingly.

When the integration of teaching and mentoring takes place through digital infrastructures purposefully designed to promote equal and user-friendly access to all stakeholders, we can achieve sustainability in professionalizing writing teachers and build mentoring interventions more efficiently and effectively.

Changing Institutional Cultures Inspire New Models of Pedagogical Repositories

Both writing instructors and the wide variety of institutional cultures have inspired us to think of new models of pedagogical repositories that depend on collaboration and intentional outreach. The exigences and purposes for repository use vary widely and can change quickly, of course. Our user-centered approach and our goal of sharing repositories widely both help address needs in times of rapid institutional change (not only in special circumstances like the most recent pandemic). The changes we are witnessing at our institutions have led us to consider how we can continue to provide sustainable professional development in increasingly difficult conditions: increasing class sizes and teaching loads; reductions in the size of graduate programs in English; greater reliance on adjunct labor; limited support and resources for instructional mentorship, in part because of the decline in funding for the humanities; and observed shortcomings in our own local in-house repositories. We present these examples of changes in programmatic cultures because they are not unique to the institutions represented in CROW, but are taking place at many types of higher education institutions in the United States.

At many research-intensive doctorate-granting institutions like Purdue and Arizona, graduate teaching assistants (GTAs) in English departments teach FYW after enrolling in a teaching composition practicum that provides institution-specific mentoring with support from WPAs. At such institutions, GTAs often receive additional mentoring for specialized writing courses such as FYW for multilingual writers, technical and professional communication, and literature. However, the changes we describe above are putting pressure on this model, sharpening the need for new approaches to providing mentoring for writing instruction. And, of course, the vast majority of teaching writing takes place outside doctoral research institutions: at masters-granting research institutions, regional comprehensives, small liberal arts colleges, and community colleges. In these institutions, models for staffing vary, but often include fewer graduate students (for whom mentoring is part of a curriculum) and more lecturers and adjuncts (who do not enroll in mentoring practicums or participate in professional development interventions). The amount of programmatic support for such interventions also varies widely—WPAs may not be afforded sufficient course releases or financial resources to permit the dedication of time to the labor-intensive work of repository development. For all of these reasons, we see the need for more cross-institutional professional development resources that can be broadly used—more responsive mentoring interventions that facilitate professional development broadly.

The rise of teaching professionals who are not tenured or tenure-track but who have greater job security and compensation than adjuncts—professors of practice, teaching professors, clinical faculty, lecturers, etc.—is one specific change in hiring trends that merits discussion. Both Arizona and Purdue are hiring more of these types of positions, and a look at the job market shows this is common at other institutions, too. Some of these faculty are hired internally from the pool of available graduate students, meaning they possess familiarity with and knowledge of the local institutional context and may have training that is also well-adapted to the context. While we do not see the decline in tenure-track hiring as a positive, institutions that make strategic moves to reduce adjunct labor by embracing a workforce with more equitable and secure work conditions create the potential for investment in professional development and substantial contributions to digital archive projects that serve and sustain collaborative mentorship.

To summarize, the burdensome responsibility of mentoring writing instructors due to limited WPA capacity has become an equity concern for all stakeholders involved. Due to the different layers in teaching expertise, professional development should be intentionally informed by authentic data representing varied pedagogies and designed to help writing instructors serve diverse student populations. Thus, cross-institutional repositories like CROW can effectively respond by providing access to mentorship resources on teaching writing. By using such repositories, writing program stakeholders can save time on creating resources and instead dedicate more time to designing interventions that facilitate collaborative mentorship and potentially create communities of practice.

A User-Centered Approach to the Repository Construction Process

Our approach to repository construction and maintenance is inspired by software development (Thomas and Hunt) and user experience design (Garrett). We seek to understand what students, teachers, program administrators, and researchers need to do their work. Thinking like designers helps us ask questions about our stakeholders, their tasks, the contexts in which they operate, and how we can engage them in our project. A considerable portion of our outreach work is invested in finding actual and potential users of our CROW platform, training them to use it, and requesting feedback. The questions and commentary these users have shared with us have helped us to learn how students, instructors, and program administrators use and think of CROW. As illustrated in Figure 1, via direct observations, post-workshop surveys, and individual interviews, we have gathered data about how people use CROW. We have utilized this information to build personas, "realistic but fictional people that typify users likely for given contexts of use," derived from out-

reach efforts and thus representing students and faculty from diverse institutions (Banat et al. 3). Collating data we collected into coherent personas makes it easier for us to keep our different users in mind, which helps us to design for specific uses and well-defined targets. These personas are not fixed, but are altered and continue to develop through further outreach to new users and by the growth of our repository. We periodically update the personas as the CROW system changes, new materials are added, new users are identified, and user needs evolve. For example, initially we focused on personas representing students, teachers, researchers, and WPAs as users; however, we added developers later due to our evolving efforts in designing automated tools that help other institutions build their own corpora. In this manner, our approach to participatory design combines design and research (observations, interviews, artifact and protocol analysis, interviews, think aloud protocols) to "iteratively construct the emerging design" of the platform (Spinuzzi 164).

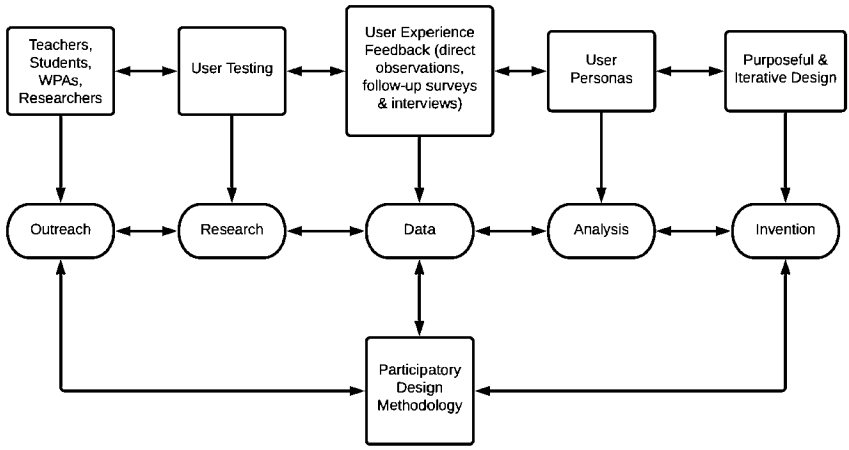

Fig. 1: Participatory design methodology for repository building.

The development of the repository is cyclical. Here we rely on design thinking as "an approach for creating solutions to difficult problems" that "encourages collaboration, creativity, and responsiveness" through "iterative cycles of design—which means designers work recursively to perfect a solution" (Tham and Thominet 3, 8). The approach is often represented by five stages: empathize with audiences, define the design problem, develop ideas to address it, build prototypes, and then test them (Dam and Siang). When doing this iterative work, Tim Brown recommends designers balance desirability, technical feasibility, and economic viability, i.e., what is feasible within the affordances and bounds of available resources and technologies (8). We have applied such approaches to emphasize that design thinking is human centered; it is facilitated

by the affordances of technology and should be informed by what our users need in the different institutional contexts they inhabit. When we were looking for potential users of the CROW platform, we considered various institutional contexts and writing program cultures. Even though the main CROW sites of data collection are doctoral research institutions, we do not aspire to design a resource that is strictly tailored to our contexts. On the contrary, we have designed CROW's repository by engaging users from multiple types of institutions: teaching-intensive universities, high schools, community colleges, regional comprehensive universities, and masters-granting institutions—both in the US and internationally. Instead of adopting a convergent approach with a predetermined set of options, we participate in a divergent approach which considers new options and possibilities outside the silos of our own institutional contexts and the local knowledge embedded in our writing programs.

By engaging these audiences, we periodically share our ideas for repository construction, which makes regular prototyping necessary. Sharing theoretical and conceptual ideas is not sufficient; we design prototypes, share them with users for testing, and make changes after user interaction and feedback. Bradley Dilger argues that the most neglected step in design processes is testing, although it "can be the engine of iteration that should be at the heart of design thinking" (41). For that reason, we invite real users to work with software, read a document, or try a process to validate how it works and to identify potential improvements. This process also promotes the premises of participatory design which, as Clay Spinuzzi argues, are "meant to improve quality of life," that is, to collaborate directly with people to improve the contexts we share, rather than imagining ourselves as experts saving users from bad design (169). We, the CROW team, have designed the repository to serve writing teachers, researchers, and administrators and help them utilize it for professional development, teaching, and research needs—at the same time we learn from these generous professionals as well.

Repository Interface Description and Features

As we have developed an understanding of how diverse audiences might use our repository, we have designed the interface in a manner that would allow quick, intuitive access to relevant materials. CROW's repository is classified by types and topics of pedagogical materials, in addition to assignment genres which represent the design of curricula at the institutional sites represented in this project. The metadata pertaining to these pedagogical materials also include institution, year, semester, course, mode of instruction, and course duration. Although the teaching materials are collected from our own institutions and writing programs, the categorization of materials by type, topic,

and assignment genre can make it applicable and transferable to other institutional and writing program contexts.

The repository's web interface is divided into two sections: left and right. The left side showcases a search box used to filter search results; thus, underneath it lies a list of pedagogical materials that live in the repository. The right side presents the filters that a user can apply to look for pedagogical materials based on assignment type, topic, genre, and other relevant metadata (see fig. 2).

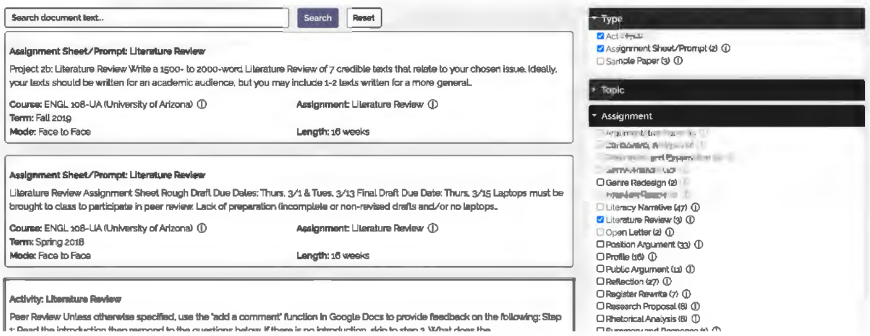

Fig. 2: Repository platform with filters representing material classification.

By selecting a pedagogical material in the search results, the interface presents access to the document itself in both PDF and plaintext formats, to cater to both teachers and researchers. Writing teachers can download PDFs of pedagogical materials to redesign and adapt to their own teaching contexts, while writing researchers can utilize the searchable text format for research purposes. For example, some writing researchers might want to examine how writing teachers familiarize multilingual students with institutional writing support beyond the classroom. In such scenarios, the writing researcher can type "writing center" in the search box and use the filter "assignment sheet." The search results will give them access to instances of the phrase "writing center" in assignment sheets, and then researchers can investigate the purposes behind these usages. The search engine highlights the word in yellow, which presents writing researchers with quick access to the portions of the assignment sheet that focus on how the writing center is presented to multilingual writers. In addition to providing access to the pedagogical materials, CROW's interface is designed to promote access to other instructors' materials related to this resource, related course materials from other instructors teaching the same course at the same institution, metadata about the material itself, and links to student drafts using the pedagogical material (see fig. 3). Related student texts are available in the corpus component of the interface. The links established between the repository of pedagogical materials and corpus of student texts

encourage users to explore and study the connections between teaching and learning in writing courses.

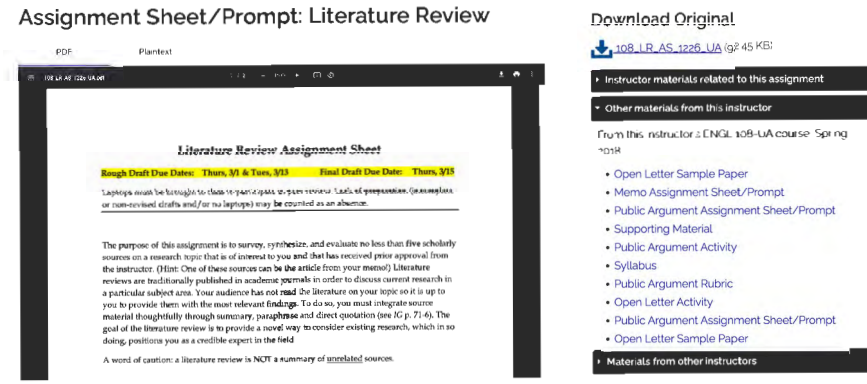

Fig. 3: Repository sample pedagogical material in PDF and plaintext formats.

Collecting and Classifying Repository Materials with User Experience in Mind

Our design processes, whether for the purposes of material collection, processing, or platform development, were always informed by UXD. For example, when building an intake form for collecting pedagogical materials, we piloted it with GTAs, lecturers, and writing program administrators. To gather feedback, we conducted think-aloud observations followed by a brief series of interview questions to discuss our users' experiences (Krug 63-64). In particular, we asked what they liked about the form, what questions they had after using it, what they wished the form would do in the future, and whether they had ideas for improvement. These UXD-informed updates helped us successfully launch the intake form and recruit contributors. Similarly, when we launched our repository online interface, we collected feedback on users' experiences via outreach workshops and demonstrations held at major conferences. The user experience feedback we collected also informed building our personas. We continue to implement suggested user experience improvements to ensure clear, intuitive navigation and use of the repository.

Our networked model of collaboration is based on maintaining visibility of material and digital infrastructures across our research team's work. We believe keeping infrastructure visible facilitates the participation of all team members irrespective of rank and time span in the project. It also promotes constructive negotiation when conflicts regarding design decisions occur. In the process of the repository construction, we developed documentation for recruitment, data collection, data processing, and data de-identification. To

facilitate such distributed work, we have Basecamp to involve team members, distribute roles, and coordinate teamwork. Basecamp integrates cloud-based calendars, schedules, to-do lists, and spaces for conversation, and is available on desktop and mobile platforms. The material affordances of Basecamp therefore allow different team members to participate in project management through creating projects, assigning themselves to projects and tasks of matching preference and skill, checking off to-do list items, adding comments, and reading over conversations. This non-hierarchical approach feeds sustainability and also contributes to our networked mentoring model which encourages peer-to-peer, student-to-faculty, and faculty-to-student mentoring.

Material Classification

With the user-centered design approach described above and a familiarity with CROW's corpus design, our team developed a shared vision of what materials we aimed to collect and a general understanding of how they might be classified, collected, and processed. When approaching how to classify materials, we also knew we would need to develop a clear protocol outlining our process to facilitate smoother collaboration among partners at various CROW sites. One of our first tasks in building the repository was designing a classification scheme for the pedagogical materials we were collecting, processing, and archiving. To do this, we reflected upon our existing personas (writing teacher, researcher, and administrator), reviewed the corpus classification system in CROW, and introduced new team members to the challenges faced with initial efforts to classify and organize the pedagogical materials. CROW colleagues had already begun the process of creating a classification system with materials from Purdue. These efforts helped us to anticipate challenges as we expanded the repository. Our bottom-up process entailed examining the materials that had initially been collected, deciding to classify pedagogical materials based on type in order to avoid vague classifications (e.g., "handouts") and to promote ease of access from a user-centered perspective. The type categories we ultimately chose were Activity, Assignment Sheet/Prompt, Lesson Plan, Rubric, Sample Paper (prior student model texts), Supporting Material (materials that facilitate classroom instruction and delivery, e.g. how-to-summarize tips), and Syllabus. These are differentiated for users on our web interface through embedded descriptions of the scope and function of each type of pedagogical material.

We also decided to identify key topics to allow CROW users to narrow searches. While not every material has a defined topic, we identified recurrent writing themes that instructors frequently addressed. For example, digital composition and language awareness were recurrent topics in lesson plans, activities, and assignment sheets, so we added them to our classification system.

In addition to the type and topic of pedagogical materials, we carried over several metadata categories already established for CROW's corpus: institution, course, semester, year, and assignment genre. We also added modality as we started to collect materials from online and hybrid courses.

To assess the efficacy of this preliminary classification scheme with the wide range of material types, assignments, and topics we encountered, we piloted the scheme on all previously collected materials. The repository team members examined each pedagogical material we had collected to validate the classification scheme and look for opportunities to expand or revise it. Challenges quickly arose—for example, presentation slides used in classes. Presentations had a wide range of content and functions, so it was difficult to identify one classification type that would encapsulate them all. We ultimately decided to classify presentation slides under Lesson Plans because they often functioned as class guides. This iterative refining process allowed us to more precisely determine what materials to collect, how to process and organize them, and how we might request them from contributors in subsequent semesters. The design thinking informing our interface development processes was inclusive of both potential contributors and future users of the CROW platform.

To envision the refined classification system as a functioning schema for processing new materials, we used Padlet to create a collaborative mind map (see fig. 4). The dynamic, collaborative nature of Padlet gave us flexibility to develop, modify, and reorganize classification levels and criteria. It also helped us visualize the processing routes of different materials, giving us a framework for our materials collection tool, as described in the next section. Once we devised a tentative classification scheme in Padlet, we invited other CROW colleagues—including our web developer—to share feedback on our scheme to ensure it was both comprehensive and technologically feasible.

Collecting Materials: Building and Piloting an Intake Form

Following CROW's process for corpus data collection, our process of collecting pedagogical materials shifted over time from in-person collection to a more efficient automated process. Initially, contributing instructors could either share their course site through the local LMS or come to a CROW lab and share their materials on a secure server. While these options aimed to limit labor for instructors, they limited instructors' agency in the classification process. Consequently, we decided to create an online intake form where instructors could independently submit and classify their materials.

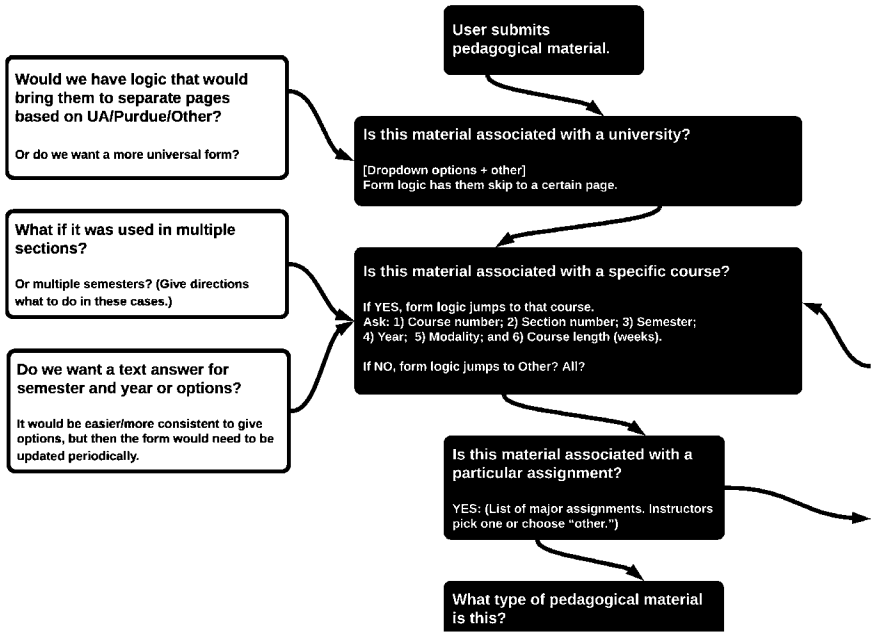

Fig. 4: Padlet mind map.

When choosing a platform for this online intake form, we identified four necessary criteria: 1) the platform must be secure, 2) it should have the affordance to collect and store files alongside corresponding metadata, 3) it must be accessible across institutions, and 4) it needed to have advanced internal functions that allowed users to seamlessly navigate to different submission pathways. For contributors, these features ensured their information would be protected, they could easily share their materials whenever it was convenient, and they would be able to access the correct consent forms and submission classifications corresponding to their respective institutions. For retrieval and processing, these criteria ensured that materials, metadata, and consent information were organized, accurate, accessible, and consistent across institutions.

Several collection platforms were considered, including Qualtrics, Google Forms, and JotForm. Ultimately, we chose JotForm as the material intake platform for five reasons: 1) it allows question pathways that facilitate collection from multiple courses and sites, 2) the material download is smoother, 3) metadata are stored in a spreadsheet with corresponding materials at the backend, 4) it is a platform that could be shared across multiple institutions, and 5) we could embed a consent form.

After selecting our material collection platform, the next challenge involved deciding the metadata to request from contributors alongside their material

submissions. Metadata helps contextualize pedagogical materials, facilitating subsequent material processing and classification. To determine the type and quantity of metadata to request, we aimed to balance our classification needs with contributors' time and labor. As such, we only selected readily available metadata: the course context (e. g. institution, year, semester, course number, mode) and the material itself (e.g., type and assignment genre). In addition to helping our team classify these materials, this metadata helped link pedagogical materials in the repository with student texts in the corpus. On the user end, these connections allow seamless navigation across CROW's corpus and repository interface.

Material Processing

Our pedagogical material processing and archiving processes are closely aligned with CROW's corpus linguistics expertise, for example, in the CIA-BATTA corpus building project (Staples et al.), though with key differences important for audiences unfamiliar with corpus use. Figure 5 maps out the material processing workflow laid out in detail below.

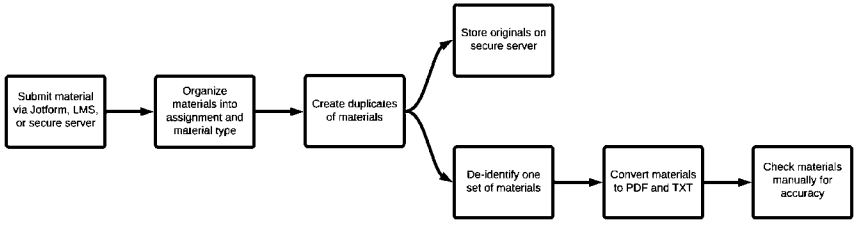

Fig. 5: Material processing workflow.

After we receive pedagogical materials from contributors, our team downloads them to a secure server. We then organize materials following a nested folder structure we have developed and documented based on repository intake data. This structure is designed to attach metadata to individual repository files by course, semester, and then by instructor code. Materials are further organized by major course assignments and material type. The options for material type are: syllabus, lesson plan, assignment sheet, rubric, activity, supporting materials, and sample paper. Finally, materials are grouped into sub-folders according to topic: digital composition, language awareness, corpus based, and peer review.

The nested folder structure has several benefits. First, it supports project sustainability and consistency for organizing and retrieving materials: team members follow a processing guide we created with step-by-step directions. Second, it allows our team to use two Python scripts in later stages of reposi-

tory development: one that creates a filename for each pedagogical material based on the folder structure, and another that adds this same information along with additional metadata in a header within the material. These automated scripts dramatically speed up processing time while also reducing the possibility of human error.

We then move to manual de-identification: carefully reviewing each document for identifying information such as names, section numbers, contact information, room locations, and personal photos. When removing this information, we replace the original text or image with angled brackets (< >) enclosing a standardized description of the content, e.g., <Professor Name> instead of Prof. Martinez. This helps preserve the composition of the original document while removing personal details.

Finally, we convert materials into PDFs so they can be displayed on our web interface with their original formatting. To ensure PDFs no longer retain identifying data, we use our Corpus Text Processor (github.com/writecrow/corpus_text_processor/releases) as part of the CIABATTA toolkit to strip authorship metadata (Fullmer and Picoral). Finally, the PDFs are additionally converted to plain text files that are included in the repository to maximize searchability.

Repository Growth and Sustainability

Repository growth depends on our writing programs' instructors who contribute to CROW by informed consent. We promote CROW's repository at different times of the year, at different institutional sites, by visiting classrooms, sending emails, and promoting CROW at professional development events as both a resource instructors can access for their own needs and an archive they can contribute to. We have been able to encourage contributions from instructors teaching students from diverse backgrounds through support from the American Council of Learned Societies (ACLS). Instructors teaching students who identified as Spanish heritage speakers contributed to the repository; these pedagogical materials will be added thus expanding the repertoire of classrooms serving diverse multilingual writers.

Our approach to sustainability helps us to avoid the uncertain futures that research projects can encounter when "faculty roles, research interests, or home institutions change" (McMullin and Dilger 470). As mentioned earlier, our strategy in response to such risks has centered on building a distributed network of scholars across institutions and disciplines. Whenever possible, we automate processes for accuracy and efficiency. Additionally, the varied interdisciplinary expertise the team intentionally attracts makes the construction of such tools and resources more technically feasible and economically viable—promoting research opportunities across multiple sites.

Repository development is also driven by carefully tested and highly accessible documentation that we repeatedly "socialize" across our team (Mulder and Yaar). Visible, clear, and detailed documentation helps our team members learn how to adopt and comment on processes and protocols we follow. The team members involved in repository development engage in data collection, data processing, process automation, and software development. We invest in members with specialized skills to supervise work processes and build detailed documentation to support our partner institutions without as much expertise on site. This vision of sustainability is inspired by Susan Leigh Star's recommendations for reliable characteristics of infrastructure: networked through multiple situations and sites, built over time, and shaped by the conventions, expectations, and institutions in which it is formed. As McMullin and Dilger point out, the application of Star's three-dimensional integration in CROW keeps networks and methods visible and thus open to observation, reflection, negotiation, and improvement, making iteration inevitable in the process development (22).

Interinstitutional Digital Repositories: Multiple Benefits and Users

CROW's repository of pedagogical materials has helped us integrate teaching writing and mentoring teachers of writing, customize our classification systems to benefit multiple institutions and users, and promote outreach and professional development broadly. Digital repositories like CROW's reduce temporal and spatial obstacles and afford inclusion from all of the institutions engaged in writing instruction. CROW not only encourages professional development but further promotes access to datasets that include authentic teaching materials and student texts. CROW's most recently awarded ACLS grant has funded outreach work to early career scholars at community colleges and Hispanic Serving Institutions (HSIs), which is expanding our network model. Providing diverse teacher and student populations access to the CROW system helps them develop awareness about language use and varieties through direct engagement with writing. With the mentoring we provide, both writing instructors and students can take up corpus informed methods for language instruction.

Our goal is to add pedagogical materials developed by our partners at community colleges and HSIs to the repository, thus expanding the diversity of material types and instructor profiles available on the CROW platform. Such professional development opportunities are part of our outreach work, including our CROW Fellows initiative, which mentors instructors interested in designing teaching materials informed by corpus-based methods, authentic multilingual student writing in our corpus, and the diverse teaching materials in our repository. Our outreach work has also included offering professional

development workshops to professionalize instructors at high schools, community colleges, and HSIs. The affordances of an interface like CROW allows us to promote sustainable professionalization of instructors beyond our own institutions.

Our collaborative engagement in infrastructure development serves to improve research and teaching practices, with an emphasis on challenging the deficit model especially harmful to multilingual students. As Shelley Staples argued in her plenary at Teaching and Language Corpora 2022, focusing on asset-oriented models of language learning and writing instruction promotes multilingual learner texts as models for other multilingual writers (mentor texts) and as sites for discussion of functional language use and choices in contexts relevant to student writers. Thus, we believe that certain pedagogical materials might reinforce monolingual bias for standard English ideologies—sometimes marked as ineffective and unjust to multilingual student populations. In other words, even if some repository materials may harbor such troublesome design decisions, they spur discussion and are useful for teacher training. Thus, our repository attempts to demonstrate how instructors can bring linguistics research to bear on assignment design and model an inclusive approach to linguistic diversity.

We continue development of the repository because we are convinced that working in the silos of our own institutions, disciplines, and programs does not advance professional development in ways to meet the evolving needs of our students and our teaching professions. Heejung Kwon, Scott Partridge, and Shelley Staples argue that "relying on existing [localized] systems and infrastructures do not help us critically reflect on our teaching practices, research methods, and program administration" (125). What's more, the changing dynamics of our classrooms, teaching, mentoring, and research contexts due to the most recent pandemic show that data-driven projects like CROW can help writing teachers, researchers, mentors, and program administrators sustain their own professionalization agendas at times when conducting research and holding face-to-face professional development trainings are challenging. CROW offers material commodities that respond to dynamic exigencies in research and pedagogy.

Different users can benefit from CROW's repository: 1) researchers keen on studying how assignment and pedagogical material design affect student learning, 2) researchers interested in exploring diverse design of pedagogical materials and writing assignment genres in different writing programs, 3) teachers who can use these artifacts as templates and examples to learn from and design their own, 4) mentors who look for teaching artifacts for writing teacher training, and 5) writing and ESL program administrators who adopt a researcher and mentor mindset to administration. We invite writing researchers

and teacher-scholars to visit writecrow.org and explore ways they could collaborate or benefit from our interinstitutional, interdisciplinary, data-driven project.

Works Cited

Banat et al. "Initiating and Sustaining Student Professionalization through Grant Writing." *Proceedings of the 38th ACM International Conference on Design of Communication (ACM SIGDOC).* Association of Computing Machinery, Article No. 2, 2020, pp. 1–6.

Brown, Tim. "Design Thinking." *Harvard Business Review,* vol. 86, no. 2, 2008, pp. 1–10.

Brunk-Chavez et al. "Designing, Building, and Connecting Networks to Support Distributed Collaborative Empirical Writing Research." *Composition Studies*, vol. 46, no. 1, 2018, pp. 81–101.

Christie, Alex. *Pedagogy Toolkit.* github.com/axchristie/toolkit. Accessed 20 February 2023.

Conference on College Composition and Communication (CCCC). "CCCC Statement on Preparing Teachers of College Writing." National Council of Teachers of English (NCTE), 2015. cccc.ncte.org/cccc/resources/positions/statementonprep. Accessed date.

Dilger, Bradley. "Testing." *Keywords in Design Thinking: A Lexical Primer for Technical Communication Today*, edited by Jason C. K. Tham, WAC Clearinghouse, 2022, pp. 41–44.

Dam, Rikke Friis, and Teo Yu Siang. "Ten Insightful Design Thinking Frameworks: A Quick Overview." *Interaction Design Foundation,* www.interaction-design.org/literature/article/design-thinking-a-quick-overview. Accessed 20 February 2023.

Fullmer, Mark, and Adriana Picoral. "Corpus Text Processor." *Writecrow Github*. github.com/writecrow/corpus_text_processor.

Garrett, Jesse James. *The Elements of User Experience: User-Centered Design for the Web and Beyond*. 2nd ed., New Riders, 2010.

Krug, Steve. *Rocket Surgery Made Easy: The Do-It-Yourself Guide to Finding and Fixing Usability Problems.* New Riders, 2010.

Kwon, Monica H., R. Scott Partridge, and Shelley Staples. "Building a Local Learner Corpus: Construction of a First-Year ESL Writing Corpus for Research, Teaching, Mentoring, and Collaboration." *International Journal of Learner Corpus Research,* vol. 4, no. 1, 2018, pp. 112–27.

McMullin, Michelle and Bradley Dilger. "Constructive Distributed Work: An Integrated Approach to Sustainable Collaboration and Research for Distributed Teams." *Journal of Business and Technical Communi*cation*,* vol. 35, no. 4, 2021, pp. 469–95.

Miller, Ryan, Thomas Mitchell, and Silvia Pessoa. "Impact of Source Texts and Prompts on Students' Genre Uptake." *Journal of Second Language Writing,* vol. 31, 2016, pp. 11–24.

Mulder, Steve, and Ziv Yaar. *The User Is Always Right: A Practical Guide to Creating and Using Personas for the Web.* New Riders, 2007.

Mueller, Derek N. *Network Sense: Methods for Visualizing a Discipline*. The WAC Clearinghouse; UP of Colorado, 2017.

Palese, Emily Anne. *Prompting Students to Write: Designing and Using Second Language Writing Assignment Prompts*. 2021. U of Arizona, PhD dissertation.

Purdy, James P. "Three Gifts of Digital Archives." *Journal of Literacy and Technology*, vol. 12, no. 3, 2011, pp. 24–49.

Reid, Shelley E. "Preparing Writing Teachers: A Case Study in Constructing a More Connected Future for CCCC and NCTE." *College Composition and Communication*, vol. 62, no. 4, 2011, pp. 687–703.

Rose, Shirley K, and Irwin Weiser, editors. *The Writing Program Administrator as Theorist: Making Knowledge Work*. Heinemann, 2002.

Spinuzzi, Clay. "The Methodology of Participatory Design." *Technical Communication*, vol. 52, no. 2, 2005, pp. 163–74.

Staples, Shelley, et al. "Corpus in a Box: Automated Tools, Tutorials, & Advising." *CIABATTA*, 2021, writecrow.org/CIABATTA/

Staples, Shelley, and Bradley Dilger. *The Corpus and Repository of Writing*. 2023. crow.corporaproject.org.

Staples, Shelley. "Learner Corpora and Data-Driven Learning: Moving Toward an Additive Approach." Fifteenth Teaching and Language Corpora Conference (TaLC), U of Limerick, Ireland, 16 July 2022. Plenary address.

Star, Susan L. "The Ethnography of Infrastructure." *American Behavioral Scientist*, vol. 43, no. 3, 1999, pp. 377–91.

Swarts, Jason. "Network." *Keywords in Writing Studies*, edited by Paul Heilker and Peter Vandenberg, UP of Colorado, 2015, pp. 120–24.

Tham, Jason, and Luke Thominet. "Introduction to Design Thinking and *Keywords." Keywords in Design Thinking: A Lexical Primer for Technical Communication Today,* edited by Jason C. K. Tham, The WAC Clearinghouse; UP of Colorado, 2022, pp. 3–17.

Thomas, David, and Andrew Hunt. *The Pragmatic Programmer: Your Journey To Mastery.* 20th Anniversary Edition (2nd ed.), Addison-Wesley, 2019.