Collaborative Research: Modeling Genre Ecologies, Mapping Communication Events: Re-imagining What Workflow Representations Are and What They Do for Knowledge Workers

William Hart-Davidson, Rensselaer Polytechnic Institute, PI
Clay Spinuzzi, University of Texas at Austin, PI
Mark Zachry, Utah State University, PI

Linked proposals submitted to NSF Program #03-611, Digital Society and Technology
January 8th, 2004

Project Summary

Intellectual merit. This project aims to discover previously hidden patterns of work organization (workflows) by taking a novel approach to the creation of work process visualizations of knowledge-intensive (KI) work. In order to answer compelling questions about the ways mature KI organizations can be both successful over time and innovative – that is, both stable and agile in terms of carrying out KI work – we adopt a new analytic approach that combines qualitative data-gathering methods with both qualitative and quantitative analysis to produce dynamic, sortable work process models at two levels of granularity:

Genre ecology models show how different types of information objects are called upon to coordinate and carry out knowledge work across an organization, illuminating the power of genres and combinations of forms to mediate key processes,

Communication event diagrams zoom in to show projects as more-or-less routinized patterns of communication events, revealing process and content-oriented logics of practice responsible for making individuals and teams effective, well-coordinated and motivated.

Both modeling techniques rely on representations of communicative events – both written and spoken – to build a representation of work rather than relying exclusively on more conventional units found in workflow systems: tasks, decisions, and milestones. We believe the resulting models are much more thoroughly appropriate for representing KI work. Rather than prescribing normative practices, our modeling technique provides a self-mediational tool that allows KI workers to become more aware of stabilizations and innovations in their work processes. Traditional models that do not depict communication events tend to lack representations of the tacit knowledge that provides the “connective tissue” for formalized work practices.

In this three-year project, we will develop our modeling techniques through a series of field research studies of KI organizations with the aim of discovering, in phase one of the project, what patterns of work characterize success, team satisfaction, stability, and innovation in mature KI organizations. In phase 2, we will turn our attention to start-up KI organizations to see, first, if similar patterns of work are characteristic of success, satisfaction, innovation, etc. In phase 3, we will actively introduce strategies from the mature organizations made explicit in phase 1 to teams in the start-up organizations in order to determine if these same work patterns can be leveraged to achieve similar results. We hope, in the process, to demystify the process of depicting knowledge work, rendering key KI strategies intelligible, interpretable, and transferable both within and across organizations.

Broader Impacts. To accomplish this ambitious goal, we will be developing a software tool to implement and extend our modeling technique. Our intention in developing the tool is to allow a wide range of users, including researchers and KI workers themselves, to enter data, to construct and manipulate visualizations of both ongoing and past work processes, and to interpret these data with varying degrees of rigor appropriate to the types of reasoning they may be doing about work. The tool will provide workers with ways to actively mediate and self-regulate knowledge work. At the same time it will create a platform for analyzing knowledge work that will allow researchers to address some of the most vexing questions about the nature and quality of communication in knowledge work. In the final year of our project, we hope to devote a substantial amount of time to testing the modeling tool with workers in both mature and start-up
KI organizations, refining the ways projects can be tracked and visualized, and increasing the capacity of the tool to handle large amounts of project data in order to allow for more and more rigorous modeling.

Understanding the Complexities of Knowledge Work: A Challenge for Mature and Startup Organizations in an Information Economy

According to reports by the U.S. Department of Labor Bureau of Labor Statistics, trends in occupational growth through the year 2010 favor what has traditionally been called “knowledge work” – business and service-sector jobs requiring education and on the job training at or beyond the post-secondary level (Hecker 2001). The fastest growing category is projected to be “professional and related occupations,” a group which is expected to comprise twenty percent of the total workforce by the year 2010 (Hecker 2001, p. 2). Within this category, specific occupations on the rise include business communication and business office jobs, both of which are growing as a result of “the corporate world wiring itself to incorporate advances in communications” (U.S. BLS 2001, p.1). These trends point to an increasing awareness on the part of the corporate world to expand capacity to perform knowledge work, and to leverage their existing human and technological assets in order to create sustainable, yet flexible practices that can be employed across the enterprise.

Firms of all sizes are beginning to invest in systems designed to help them manage the critical information that their business depends upon, and there is evidence that investment in content management systems will significantly outpace the software industry as a whole in the foreseeable future (Julian, 2002). But there is also evidence that much work remains in understanding how to best use systems designed to support knowledge work. According to one IDC/Xerox report, for example, knowledge workers spend 15-30% of their time at work conducting searches for information, but up to 50% of these searches are unsuccessful. This sort of failure to understand and leverage regularities in knowledge work can add up to U.S. $2.5-3.5 million annually per 1000 knowledge workers in lost productivity, according to the report. And while we may have reason to suspect that these dramatic numbers have everything to do with IDC/Xerox having a stake in convincing businesses that a crisis is at hand, the proposed “solutions” to the problem remain challenging in their own right: “reengineer business process and document workflow,” “develop customized solutions…starting with an assessment” (Boyd, 2003).

The clear message, apart from the marketing speak in industry whitepapers and the broad summaries of occupational trends in government reports, is that our ability to capture and manage information is growing, but so is our need to understand how knowledge work organizations rely on information to achieve success.

Our Approach to Modeling Knowledge Work

This project proposes a new way to study this problem that promises to change the way we see knowledge work – literally and figuratively – by introducing a modeling method and software for visualizing, analyzing, and enacting knowledge work. Our approach draws upon a tradition of research on written and technologically-mediated communication in workplace settings, incorporating theoretical perspectives and research techniques from distributed cognition (Cole & Engestrom, 1993; Hutchins, 1995; Winsor, 2001) and Activity Theory (Bazerman, 1997; Engestrom, 1999; Nardi, 1996; Spinuzzi, 2003).

Our technique views knowledge work as chains of coordinated communication events, organized and understood as genres. These events become the primary unit for creating representations of work in our approach, supplanting, but not totally replacing other important units such as tasks, decision points, users or actors,

Emerging Consensus on Workflow

Work process representations, as plans which precede any given activity related to the work process, are neither fully adequate prescriptions for future work, nor exhaustive descriptions of past work (Suchman, 1987; Dourish, 2001).

Work representations are most valuable to users when they are event-based, as this allows users to understand work in terms of motivated, explicitly goal-oriented activity (Gerson & Star, 1986; Anderson et al., 2003).

Work process representations are mediational resources for workers, helping them to reason and account for specific actions (Dourish, 2001).

Work process descriptions mediate reasoning about work processes by helping workers do “articulation work,” visualizing and making connections among important events, actors or roles, artifacts, constraints, and outcomes (Simone, Divinity, & Schmidt, 1995; Star & Strauss, 1999; Hayes, 2000).
documents, or combinations of these such as “active documents” (Dourish, et. al., 2000), or process/constraint representations (Bernstein, 2000). We have learned much from others’ approaches to modeling work practices, paying close attention to the ways various types of annotated process models such as Dourish et. al.’s “Freeflow” system (1996) have been developed and subsequently analyzed in workplace settings (Dourish, et. al., 1999a; Dourish, et. al., 1999b; Dourish, et. al., 2000). Dourish et. al.’s overall approach to workflow moves toward the creation of open systems which guide and support, but don’t entirely anticipate and control work processes.

We share a similar viewpoint, which has arisen from our own studies of workplace writing (Spinuzzi, 2003a, 2003b; Spinuzzi & Zachry, 2000; Zachry, 2001) and our attempts to develop modeling methods which represent writing practices (Hart-Davidson, 2002a; 2002b; Spinuzzi, 2002; Hart-Davidson, 2003). In these efforts, we have come to understand communication genres and the various instances of these genres which circulate in organizations as useful tools for representing the nature of work. Workers, themselves, rely on genres to understand their own tasks and to coordinate tasks with others (Bazerman, 1988; Smart, 2003; Mirel, 2003). Researchers of workplace writing have shown that it is not only the skillful employment of genres that leads to success (e.g. Spilka, 1993), but it is also the ability to understand knowledge work as successful communication which is important (Freedman & Smart, 1997; Geisler, 2001; Smart, 2000). The “window on work” that genres provide, combined with the specific task and context related content that resides in specific genre instances make for a powerful set of interpretive resources for knowledge workers (Dias, et.al., 1999).

So while we have examined approaches to workflow system design and work process visualization that incorporate specific types of genres or instances of genres as an approach to what is known as “exception handling,” (e.g. see Casati, et. al., 2000), we are making a fundamental break from this approach in proposing communication events as the primary unit for modeling. To put it simply, we do not believe that communication events are “exceptions” in the context of knowledge work. They are the rule. And while communication events happen as a way for people to handle unexpected circumstances, whether they be “planned” or “unplanned” exceptions from the perspective of a managed workflow system, they are not ancillary to doing tasks or making decisions. They are, rather, constituents of these!

In reviewing the literature on workflow, we believe we can make a contribution to both the theory and practice of work process descriptions by introducing key concepts from the study of workplace writing, specifically from the study of “genre,” seen as a crucial facilitator of coordination and action in organizational settings. In so doing, we expose patterns of knowledge work previously invisible in most workflow systems.

**Research Questions**

Our modeling approach allows us to address compelling questions fundamental to carrying out knowledge work. Our project will focus on three questions:

1. How do successful knowledge work (KW) organizations balance work process stability with agility in order to sustain themselves over time and continuously innovate?
   
   a. What patterns of communication events correlate with project success metrics? (e.g. quality of key deliverables? levels of client/customer satisfaction, success or adoption rate of project proposals)
   
   b. What patterns of communication events correlate with team satisfaction? (as measured by satisfaction scales and exhibited in observations and interviews)

   c. What patterns of communication events correlate with stability, over time, of work practices? (e.g. as measured by repetition, explicitness of key patterns, adherence to common plans or managed workflows)

   d. What patterns of communication events correlate with innovation? (as measured by substantial and favorable gains in success or satisfaction or by substantial favorable changes to work patterns [e.g. shortening of overall time-to-completion] without sacrificing success or satisfaction).
2. Do start-up organizations exhibit patterns associated with project success, team satisfaction, stability, and innovation in mature KW organizations?

3. Can patterns shown to be positively correlated with success, satisfaction, stability, and innovation in mature organizations be transferred to start-up organizations with similar results?

We will address these questions in a three-phase study that will 1) model knowledge work in mature, successful knowledge work organizations in phase one, 2) compare these models to those derived from the study of start-up KI organizations in phase two, and, 3) introduce successful patterns of work derived from the study of mature organizations in phase one in order to determine if they are similarly valuable for start-up organizations who employ them.

In short, we want to see if new KI organizations can learn from mature ones the strategies needed to be successful, stable over time, and adaptable to new challenges as they arise.

We will pursue this goal by focusing on one particular activity that is notoriously ad-hoc and perhaps necessarily so: grant proposal writing. Success over time at this activity requires an organization to leverage past experience but to be flexible as well, willing to alter both the means by which it works as well as the ends it aims to accomplish. And while we understand that there may be some strange reverberations here, proposing a study of proposal writing to a panel that is steeped in this genre and its associated activities, we have chosen the activity of proposal writing primarily for its ability to put mature organizations and start-up organizations on relatively equal footing so that we can better account for factors leading to success, stability, team satisfaction, and innovation. We do not assume that patterns of activity coming from mature organizations will automatically be “better” than those emerging from start-ups, in part due to the rather volatile nature of the activity. We do assume, though, that mature organizations have developed patterns that allow

---

**Contributions Our Approach Makes to Research on Knowledge Work Process Representation**

The minimal meaningful unit of a knowledge work process representation is a communication event. Some other units, such as tasks, decision points, and constraints – indeed workflows themselves – can be understood as assemblages of communication events. Actors, artifacts, and deliverables, moreover, participate in communication events, and so they too are understood to be constituents of these events in models of knowledge work.

Communication events in knowledge work representations may be modeled as objects events with value-pair attributes that function as metadata for a communicative artifact – e.g. a document, a chat session transcript, etc. – in order to preserve both the goal-oriented nature of the event at the action level, and the motivations and traditions associated with the event-as-genre at the sociocultural level (Hart-Davidson, 2001; Spinuzzi & Zachry, 2000; Hart-Davidson, 2002a; Spinuzzi, 2003b).

Genres are functional for writers in organizations in two complementary ways that lend themselves to constructing useful pictures of knowledge work: 1) as relatively stable representations of strategies that have been successful in the past, genres can stand in for rather complex arrays of social interaction, helping writers to leverage the skill and knowledge of the past; 2) at the same time, genres permit writers to handle novel situations because they are malleable and readily used in combination with other genres (Spinuzzi, 2003b).

Workflows, to the degree that they represent relatively stable representations of work processes made up of assemblages of communication events, are best seen as interconnected networks of genres or genre ecologies, “interrelated group[s] of genres (artifact types and the interpretive habits that have developed around them) used to jointly mediate the activities that allow people to accomplish complex objectives” (Spinuzzi & Zachry, p. 172).

Work process representations built from communication events can serve as mediational resources for writers within organizations, helping them to reason and account for specific actions; they may also help researchers answer important questions and/or triangulate findings from controlled studies with respect to real-world project data (Hart-Davidson, 2003).

Work process descriptions mediate reasoning about knowledge work by helping both researchers and workers themselves understand connections among genres and goal-oriented actions as well as opportunities for substituting and/or modifying genres to achieve specific aims (Spinuzzi, 2002; 2003b).
them to persist and, in the cases of the organizations we have chosen as research sites in phase 1, to be successful, stable, and innovative. Our qualitative research methods, in conjunction with the formal modeling, will help us to more fully explore the possibilities and pitfalls of transferring work strategies from mature to start-up organizations.

We also believe that proposal writing is an area that the organizations, themselves, will want to have a more explicit grasp of, as it can be a high-stakes and high-risk proposition. Therefore, sharing work process representations in the “modeling interpretation” sessions described below will be potentially valuable for the participants rather than a drain on their time. Following trends in user-centered and participatory design approaches to the development of information systems, we will use the data from these interpretation sessions not only to triangulate research results, but to improve the way the modeling software works as well.

Answering the Questions: Using Concepts of Genre and Writing Activity to Model Knowledge Work

Our modeling approach relies on two key concepts which serve as units of analysis for the construction of meaningful visualizations of knowledge work:

1) chains of coordinated communication events understood as writing activities or projects, and
2) genres, understood as typified responses to recurrent social situations, and recognizable as regularities in the formal features of discourse

Together, these concepts allow us to create pictures of an activity that is notoriously difficult to visualize due to its distributed nature. Projects are visualized in Communication Event Models (CEM), built from a record of all the communication events that members of a given project team participate in. Each event in a CEM is represented as a database record with multiple attributes that allow for visualizing and sorting the CEM in order to interpret the dynamics of single project in the past or one which is ongoing. As CEMs for individual projects accumulate, we can begin to see patterns of use that cut across projects, which we capture in another visual format called a Genre Ecology Model (GEM). These models highlight the ways specific discursive forms coordinate with recurring organizational goals, and they also show the ways genres are typically employed in conjunction with one another. This sort of view, we argue, affords forward-looking and, over time and with enough project data, perhaps an inferentially valid means of deciding on best practices for employing genres to address specific organizational goals.

The image below offers a visual summary of our modeling approach. As described above, data is gathered for the models along two dimensions. We adopt the linguistic terms syntagmatic and paradigmatic to describe the basic analytic stance for each dimension.

Syntagmatic analysis asks “what elements may follow other elements and still yield a coherent, meaningful linguistic unit (known as a “syntagm”)? At the level of syntax, a linguist conducting syntagmatic analysis might ask what words can be added to a clause that would extend it without making it ungrammatical. At the level of writing activity, we are asking about much larger units of discourse, but our analytic stance is similar. We ask “what communicative events can come next in this project?” This allows us to gather information about how individual events are associated with one another.

Paradigmatic analysis asks “what elements may be substituted for other elements in a given syntagm while preserving the meaning and coherence of the unit?” When used together, the two dimensions allow for the rules of language systems to be inferred, tested, and compared with those of other language systems. In this project, we are interested in the qualities mentioned in the research questions above: what sorts of patterns can be linked to the success of a given project? To the satisfaction of the team? To the stability of work processes during and across projects? And to the fostering of innovation across projects of similar types, with similar participants, etc.?
In the two sections that follow, we describe the modeling approach in greater detail along each of the two dimensions depicted above. We have, thus far, constructed CEMs and GEMs only by hand, which has been valuable for determining the necessary data and how best to obtain and represent it. The proposed project would allow us to automate each of these steps to some degree by building software to facilitate collection, analysis, and representation of data. This would, in turn, allow us to refine and extend the modeling techniques, making them more rigorous and attuned to the criteria we have targeted in the research questions above. We also believe that the models will become more useful to knowledge workers who may use them to monitor ongoing work, review past work, and plan for future work.

**Seeing Stability and Agility with Communication Event and Genre Ecology Models**

*Communication Event Models (CEMs)*

We choose to depict writing activities or projects as strings of coordinated communication events because this approach makes the distributed, collaborative nature of such projects visible and intelligible. Taken together, the collection of events that make up a project are examples of motivated, explicit, goal-oriented knowledge work. And while, for the purpose of modeling, project “syntagms” can be of \( N \) length, in practice, they have beginnings and endings, and names that correspond to familiar workplace goals and actions. Projects are meaningful units for knowledge workers because, unlike some units found in work process representations (i.e. workflows), they correspond exactly with what workers perceive themselves doing in real life. The also correspond meaningfully with workers’ goals and motivations, which makes them more robust than “tasks” or “decision points.”

The CEM to the right depicts a routine writing activity in an organization: the development of a quarterly human-resource allocation report. Adapting a convention proposed by...
Gunnarson (1997), each icon in the model represents a specific communication event. Elliptical icons denote an oral genre such as a phone call or a face-to-face meeting. Rectangular icons denote written genres such as e-mail or printed documents. Events are ordered chronologically according to the date on which they occurred. In this view, the events have been sorted to emphasize how efficient the project was perceived to be according to one of the participants’ ratings of “on-task” communication events vs. those having other purposes (socializing, coordinating, troubleshooting, etc.). Each event rated as “on-task” by the participant starts a new horizontal line, allowing efficiency to be understood as “working vertically.” This particular project proceeded in a fairly efficient manner, with a relatively high percentage of on-task events relative to the total number. We might expect this level of efficiency in a process that recurs quarterly, as this one does. But this model allows us to see precisely how this efficiency is achieved and, more importantly, how the strategies used here might be leveraged in other projects.

What makes the CEMs especially valuable for this sort of reasoning is the ability to sort project data in various ways to emphasize different aspects of the project. This is possible because projects and communication events are modeled as data objects with attribute-value pairs. Below is a listing of the data categories for the communication event object developed in our proof-of-concept work on CEMs (Hart-Davidson, 2002; 2003). These represent canonical categories and data types for the first iteration of the modeling software, currently in progress, as indicated by the user-interface mock-up for “adding a new event to an existing project.” We anticipate adding to, testing and refining these categories in order to increase the reasoning power of the models for researchers and knowledge workers alike.

Hart-Davidson (2003) reports on one such iteration of the CEM approach which involved sorting to emphasize the overall level of coordination exhibited by teams with two different sets of collaborative constraints. For this study, student teams were chosen in order to ensure that both groups would have an identical task – a report writing assignment, in this case – with identical beginning and end dates. One team, the co-located team, was composed entirely of on-campus students, while the other was non-co-located, consisting of students separated from the campus and from each other. While the two teams began the project the same way, the distance team got off to a bad start soon after by working in an uncoordinated way. Analysis of the communication event types in the early going showed that many of the opportunities for planning the scope of the project used by the local team were either not available to the distance team (f2f meetings) or were carried out with at least one team member missing. As a result, the distance team “struggled to get on the same page throughout the whole project,” according to one of the members.

From a modeling perspective, the interesting thing to point out is that this struggle was quite noticeable from the CEMs, both to the researchers and to participants who saw them in model interpretation sessions. Perhaps most promising about these pictures is the fact that we can see the trouble for the distance team brewing early, before the project is even half completed. This suggests that we might not only learn from this example, but that in future situations, we can monitor projects in real time, paying attention to functions such as “coordination” and “progress” in order to address any problems that arise before they can derail a project.

**Genre Ecology Models (GEMs)**

Whereas CEMs provide a syntagmic view of how events are assembled chronologically and enacted through media – that is to say, associations – GEMs provide a paradigmatic view of the divergent information resources that are used during these events – that is to say, substitutions among these resources.
For instance, one worker might track her work by stacking forms, pages, and notes in the order that she expects to consult them; another worker might leave these different genres in different places but use a checklist to order and track her tasks. In this case, the checklist and the stack substitute for each other. Indeed, workers often innovate at the individual and group level by introducing new genres into a given ecology (i.e., assemblage) of genres, using new genres to replace or supplement existing ones. As one of us has argued elsewhere, popular modeling techniques tend to ignore or dismiss this innovative substitution work (Spinuzzi, 2003a). This tendency to innovatively substitute and develop genres tends to lead to stability over time, as genres and their relationships become established within organizations, while further innovations provide the flexibility necessary to adapt to changes in the work and the project environment.

Genre ecologies are a way of talking about the material set of interpreted informational resources on which workers draw as they do their work. Genre ecologies tend to expand quickly because workers draw on genres they have learned elsewhere and often experimentally substitute genres during their work, particularly (but not simply) when they are dissatisfied with how work is performed. Rather than discrete tools, genres are used in assemblages to jointly mediate the work activity. Workplace researchers have deployed various terms and frameworks to explore this joint mediation, including datacloud (Johnson-Eilola, 2001), information ecologies (Nardi & O'Day, 1999), tool ecologies (Hutchins, 1995), genre sets (Devitt, 1991), genre systems (Bazerman 1994; Yates & Orlikowski 2002), and genre repertoires (Orlikowski & Yates, 1994). Without getting into the differences, we recognize a rich set of theoretical frameworks and empirical work that inform the genre ecology framework.

We have modeled genre ecologies by hand to represent the genres on which groups of workers drew during various field studies (Spinuzzi, 2002; 2003b). For instance, Figure 4 shows an ecology of genres on which a group of workers drew during observations. Figure 5 shows a more detailed genre ecology in which regular patterns of mediation are described. In this study, we plan to develop these models in ways that more clearly show how individuals draw on their genre ecologies, the frequency of genre use, and substitutions, both across participants' work and within each participant's work.

![Figure 4. An informal genre ecology diagram (Spinuzzi 2002, p.201)](image)

Figure 4. An informal genre ecology diagram (Spinuzzi 2002, p.201)
Research Methods
As discussed above, our technique seeks to make communication events the primary unit for modeling, allowing us to build pictures of regularized and divergent communicative practices in workplaces as well as the assemblage of resources used to accomplish those practices. Furthermore, our technique provides a way to graphically share insights with those whose work is represented, allowing them to benefit from these insights and us to gain more insights through our conversations with them.

To develop the technique – including the models, the software used to generate them, and the mode of interaction with workers – we will study the activity of grant proposal writing. This activity is a strong fit for developing the technique because it is an activity with a definable start and end; it exhibits regularity across sites and regions; and it is explicitly collaborative knowledge work in which writing is a central activity. Furthermore, many guidelines on the general process exist to give us background, yet few formal studies have attempted to model the activity as we plan to do (c.f. Van Nostrand, 1997).

In accordance with the research questions, our methods will allow us to explore how successful knowledge work organizations balance work process stability with agility in order to sustain themselves over time and continuously innovate; whether start-up organizations exhibit patterns similar to those in established organizations; and whether patterns that have been successful in established organizations can be successfully transferred to start-up organizations.

The studies will occur at three sites simultaneously. In Year 1, we will study three mature organizations: The Academy of Electronic Media (Troy, New York), Space Dynamics Laboratory (Logan, Utah), and Knowbility (Austin, Texas).

- The Academy of Electronic Media, a research and multimedia development center on the Rensselaer campus, conducts research and develops new learning technologies for academic, non-profit and corporate partner organizations across a wide variety of content areas. Current activity areas include interactive design tools and tutorial units, interactive simulation modules, multimedia case studies, virtual reality environments, and modular multimedia curricula in content areas such as language learning, health care, engineering, and computer science.
Space Dynamics Laboratory (SDL), a unit of the Utah State University Research Foundation, is a not-for-profit research corporation. SDL is responsible for the design, fabrication, and operation of payloads, ranging from aircraft and rocket-borne experiments to space shuttle experiments, small satellites, and satellite-based sensor systems. As one of ten University Affiliated Research Centers (UARCs) in the nation, SDL conceives and develops state-of-the-art sensor and satellite systems; performs space, air, and ground-based experiments; conducts rapid, experimental development of prototype sensor hardware and associated software; performs concept validation studies and demonstrations; and develops data-fusing technology for passive and active sensors.

Knowbility is an Austin-based nonprofit organization that works to provide accessibility awareness and training services to government agencies, industry, the educational community, and people of all ages with various disabilities. It partners with various entities, such as the Austin Independent School District, the University of Texas at Austin, and the Texas Commission for the Blind, to promote Internet accessibility through awareness building, education, and employment programs; regional Accessibility Internet Rallies in Texas, Colorado, California, Georgia, and Washington, DC; and accessible web page design curriculums.

Conducting studies at mature organizations will allow us to examine relatively stable work activity as opposed to the rapidly changing, emergent activity typical at less mature organizations. At the same time, since the three sites are in different regions and different sectors, we will be able to identify commonalities and differences among the sites. The result, we believe, will be a more robust modeling technique that can then be applied to a wider variety of organizations doing similar work.

In Year 2 and 3, we will study three emerging organizations in matching industries, drawn from the technology incubators at our respective universities. Conducting these studies will allow us to compare practices and resources between mature and emerging organizations in the same industry as well as across industries. Studies in all three years will involve the following:

**Data Collection**

Data collection involves diverse methods for exploring the communicative patterns and resources that organizations follow to sustain work activity while remaining flexible, and how these patterns relate to project success and team satisfaction. As the same data are collected across established and start-up organizations, it will become possible to compare differences in these organizations’ work.

- **Diary studies:** Using preprinted forms, workers will keep diaries of their communicative events and the artifacts they use during these events. Researchers will periodically remind workers to keep diaries, collect diaries as the project progresses, and enter data into a database accessible by all researchers.

- **Observations:** For a subset of the groups studied, researchers will observe key events in proposal writing and take detailed field notes. These field notes will be entered into the research database.

- **Post-observational interviewing:** Researchers will conduct semi-structured interviews with observed workers. These interviews will be recorded digitally and stored in the research database, along with the researcher’s notes.

- **Questionnaires:** At the end of each project, all participants will fill out confidential questionnaires in which they rate their satisfaction with key deliverables, their work resources, and their collaborative work with their teams. These questionnaires will include Likert scales and open-ended questions.

- **Artifact collection:** Researchers will periodically visit the work site and collect artifacts such as drafts, email, and memory aids. Artifact collection will be guided by diary studies, observations,
and post-observational interviews. Artifacts will be digitized and stored in the research database along with the researcher's notes, and linked to other relevant data in the research database.

**Data Analysis**

Data analysis will focus on the tension between stability and innovation and how patterns lead to (or away from) project success and satisfaction. In years 2 and 3, the modeling interpretation sessions will provide a vehicle for transferring successful patterns from emerging to start-up organizations.

**Coding diaries.** As diaries are entered into the research database, researchers will code them. Coding will be based on the following:

- **Communicative events.** Communicative events will be defined in advance through a flexible coding system built into the diary forms.

- **Genres.** Genres, or artifact types, will be inferred from the artifacts workers list as being used during communicative events. Researchers will check their inferences periodically through direct inquiries or inquiries during interviews.

- **Predetermined categories.** Researchers will agree on categories before data collection begins. A table of such categories will be available through the research database, and will include events and resources.

- **Emergent categories.** Other categories will emerge as researchers examine the data. This coding is vital since the other two types of categories are not flexible enough to catch unexpected behavior and innovations. Researchers will develop emergent categories based on their own sites' data, then discuss and synchronize those categories periodically. Emergent categories will be stored in the research database, associated by researcher, explicited with examples, and available to all three researchers. Emergent categories will include, among other things, genres (artifact types) used at the site.

  - Coding will be tested for inter-rater reliability through Cohen's Kappa.

Coding the diaries will provide a basis for visualizing communicative patterns and resources and serve to explore the relative stability of those patterns and resources.

**Coding field notes.** As field notes are entered into the research database, researchers will code them according to communicative events, genres, predetermined categories, and emergent categories. Like diaries, field notes will provide a basis for visualizing communicative patterns and resources and serve to generate more insight into the context in which these patterns and resources are used.

**Coding interviews.** As interviews are entered into the research database, researchers will code them according to communicative events, genres, predetermined categories, and emergent categories. Interviews will provide insight into the reasons workers use particular communicative patterns and resources, their perceived success, and how satisfied workers are with them.

**Analyzing artifacts.** Researchers will conduct a genre analysis of selected artifacts. Artifact analysis will be particularly important for examining the stability of relationships among genres and the innovative introductions of new genres. Artifact analysis will also be triangulated with questionnaires to provide measures of the quality of deliverables. The genre analysis will include

- **Coherence.** Examining the artifact as it is used across events and participants, particularly in terms of convergent and divergent understandings across data sources (see Mol 2002). Examining coherence will involve triangulating the artifact with the diaries and field notes (See Spinuzzi, 2003b; Yates & Orlikowski, 1992, 2000).
- History. Examining the artifact as it relates to previous versions of the artifact in the organization's archives. For instance, meeting minutes for an observed meeting might be compared to those from earlier years to detect changes and reorientations. (See Bazerman, 1988; Spinuzzi 2003b; Zachry, 2001.)

- Comparisons. Examining the artifact as it relates to other artifacts used during the same communicative event. For instance, meeting minutes can be compared to other documents used during the meeting, the proposal draft discussed during the meeting, and emails related to the meeting. Comparisons across artifacts can provide insight into how information gets circulated and transformed across the artifacts. (See Bazerman, 1994; Berkenkotter & Huckin, 1995; Orlikowski & Yates, 1994; Spinuzzi, 2003b; Yates & Orlikowski, 1992, 2002.)

**Triangulating.** Based on the above work, researchers will triangulate the results of the analyses.

- Individual sequences. Selected communicative event sequences from individuals' diaries and field notes will be examined, drawing on interviews and genre analyses to better understand the described interactions as enacted by the individual workers.

- Comparative sequences. Selected communicative event sequences will be compared across participants in each team. Doing so should give the researchers insight into how different team members interpret the same events and how they rate the success and satisfaction of these events.

Triangulating the analyses will provide a more unified picture of stability, innovation, perceived success, and satisfaction. In years 2 and 3, these can be compared with success metrics for the proposals written during the previous years.

**Modeling.** Based on the coding work, researchers will build models based on the above.

- Communication event models. This modeling technique is described in (Hart-Davidson, 2002a; 2003). Based on the coding in the diary studies, field notes, and interviews, researchers will generate models of communication events through proof-of-concept software that has been generated by the lead PI, itself based on a non-automated modeling technique. (This work will also serve to guide further development of the modeling software, resulting in alpha-quality software by the end of the engagement.)

- Genre ecology models. This modeling technique is based on the genre ecology work of Spinuzzi and Zachry (2000), particularly a strand of modeling work (Spinuzzi, 2002; 2003b). Based on the emergent coding in the diary studies, field notes, and interviews, as well as the genre analyses, researchers will generate models describing the assemblage of genres (artifact types) supporting each communicative event. In particular, the genre ecology models should allow researchers to examine how workers (a) use clusters of genres as a unit for accomplishing certain actions and (b) substitute genres in these clusters from event to event, worker to worker, and site to site.

**Model interpretation sessions.** During this final stage, workers and researchers will examine models during model interpretation sessions. Based on scenario-based participatory design work (e.g., Bodker and Grønbaek, 1996; Ehn, 1989; Muller 1999), these model sessions will be run as semistructured interviews focused on the models as representations of work; however, since the focus is on the co-analysis of the models, we describe this stage as analysis rather than data collection. Model interpretation sessions will be videotaped. Since the models will be developed as a way to support workers' self-examination and self-reflection, these sessions will be crucial for fine-tuning models for the workers' use as well as discovering ways in which the models could better serve the researchers' work. In years 2 and 3, model interpretation
sessions will provide a mechanism for transferring successful patterns from mature organizations to start-up organizations.

- **Reactions.** Workers examine the models, cointerpret them with the researchers, and comment on how well or poorly they reflect how the workers perceive their own work.

- **Reflections.** Workers suggest changes to their own work based on the models.

- **Suggestions.** Workers suggest changes to the models that would make them easier to use, more useful, or more representative.

- **Changes.** Based on reactions across the sessions, researchers will modify the models appropriately.

**Success metrics.** In years 2 and 3, researchers will count the number of accepted proposals from the previous year. This count will provide one important measure of project success.

**Schedule of Research Activities**

<table>
<thead>
<tr>
<th>Year 1</th>
<th>Year 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.15.2004</td>
<td>Initiate diary study 1</td>
</tr>
<tr>
<td>12.15.2004</td>
<td>Progress report 1</td>
</tr>
<tr>
<td>3.15.2005</td>
<td>Progress report 2</td>
</tr>
<tr>
<td>3.15.2005</td>
<td>Complete field work</td>
</tr>
<tr>
<td>4.1.2005</td>
<td>Complete coding &amp; analysis of diaries</td>
</tr>
<tr>
<td>4.15.2005</td>
<td>Complete coding &amp; analysis of field notes &amp; interviews</td>
</tr>
<tr>
<td>4.15.2005</td>
<td>Progress report 3</td>
</tr>
<tr>
<td>5.15.2005</td>
<td>Complete genre analysis</td>
</tr>
<tr>
<td>6.15.2005</td>
<td>Complete model building</td>
</tr>
<tr>
<td>6.15.2005</td>
<td>Complete software development</td>
</tr>
<tr>
<td>7.15.2005</td>
<td>Complete model interpretation sessions.</td>
</tr>
<tr>
<td>7.15.2005</td>
<td>Progress report 4</td>
</tr>
<tr>
<td>8.1.2005</td>
<td>Model &amp; software refinement</td>
</tr>
<tr>
<td>8.30.2005</td>
<td>Year 1 report</td>
</tr>
<tr>
<td>8.30.2005</td>
<td>Release alpha version of software</td>
</tr>
<tr>
<td>9.15.2005</td>
<td>Initiate diary study 2</td>
</tr>
<tr>
<td>10.2005</td>
<td>Present panel at SIGDOC</td>
</tr>
<tr>
<td>12.15.2005</td>
<td>Progress report 5</td>
</tr>
<tr>
<td>11.30.2006</td>
<td>Complete course materials &amp; website</td>
</tr>
<tr>
<td>12.15.2005</td>
<td>Progress report 6</td>
</tr>
<tr>
<td>3.15.2006</td>
<td>Complete field work</td>
</tr>
<tr>
<td>4.1.2006</td>
<td>Complete coding &amp; analysis of diaries</td>
</tr>
<tr>
<td>4.15.2006</td>
<td>Complete coding &amp; analysis of field notes &amp; interviews</td>
</tr>
<tr>
<td>4.15.2006</td>
<td>Progress report 7</td>
</tr>
<tr>
<td>5.15.2006</td>
<td>Complete genre analysis</td>
</tr>
<tr>
<td>5.15.2006</td>
<td>Complete genre analysis</td>
</tr>
<tr>
<td>6.15.2006</td>
<td>Complete model building</td>
</tr>
<tr>
<td>6.15.2006</td>
<td>Complete model building</td>
</tr>
<tr>
<td>7.15.2006</td>
<td>Complete model interpretation sessions.</td>
</tr>
<tr>
<td>7.15.2006</td>
<td>Progress report 8</td>
</tr>
<tr>
<td>8.1.2006</td>
<td>Compare models with Year 1 models.</td>
</tr>
<tr>
<td>8.15.2006</td>
<td>Year 2 report</td>
</tr>
<tr>
<td>8.30.2006</td>
<td>Check on success of Year 1 proposals.</td>
</tr>
<tr>
<td>8.30.2006</td>
<td>Release beta version of software</td>
</tr>
</tbody>
</table>

**Year 3**

<table>
<thead>
<tr>
<th>Year 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.15.2006</td>
</tr>
<tr>
<td>10.2006</td>
</tr>
<tr>
<td>12.15.2006</td>
</tr>
<tr>
<td>3.15.2007</td>
</tr>
<tr>
<td>3.15.2007</td>
</tr>
<tr>
<td>4.1.2007</td>
</tr>
<tr>
<td>4.15.2007</td>
</tr>
<tr>
<td>4.15.2007</td>
</tr>
<tr>
<td>4.15.2007</td>
</tr>
<tr>
<td>6.15.2007</td>
</tr>
<tr>
<td>7.15.2007</td>
</tr>
<tr>
<td>7.15.2007</td>
</tr>
</tbody>
</table>

**Year 3, cont.**
8.1.2007  Compare models with Year 1 and 2 models.
8.1.2007  Check on success of Year 2 proposals.
9.30.2007  Year 3 report
9.30.2007  Release the open-source software application
10.30.2007  Submit book manuscript
10.2007  Present full-day tutorial at SIGDOC

Deliverables
The proposed project will generate a series of deliverables, which are designed to build upon on another before culminating in three products to be publicly distributed at completion of the project: 1) an open-source software application, 2) a tutorial to be offered at an international conference, and 3) a textbook manuscript to be submitted for publication.

The researchers’ work on these deliverables will be documented in a series of quarterly progress reports, which will be published on an intranet supporting the collaborative work of the research team.

Year 1
Deliverable 1: Analysis of data gathered from three mature KI organizations
The data gathered concurrently from the three research sites will be analyzed to create a richly detailed picture of the work processes of knowledge workers creating proposals. These analyses will be shared with representatives at the research sites and then used as a resource for development of an alpha version of the software as described below.

Deliverable 2: A panel session at the annual meeting of ACM SIGDOC
In October 2004, the research team will present the first in its planned series of three annual sessions at the international ACM SIGDOC conference. This first planned session is a panel presentation and discussion of the project, including its theoretical assumptions, its innovations in the area of work models, and its execution. This presentation will be published in the annual proceedings of the conference and will be available online through the ACM Portal.

Deliverable 3: An alpha version of the open-source software application
At the end of Year 1 of the project, the research team will release an alpha version of the open-source software application being developed to support the analysis of knowledge work in organizations. The software will be released on an open developer site such as sourceforge.net. Upon release, notification of the availability of the software and descriptions of its use will be published through relevant online channels where representatives of the interested interdisciplinary community are most likely to encounter the information (e.g., open-source repositories such as SourceForge www.sourceforge.org or Freshmeat www.freshmeat.org, as well as sites maintained by the PI’s labs). We will also announce these releases on relevant disciplinary mailing lists and web forums. Feedback from the open source development community will be used to refine the software application in subsequent years of the project.

Year 2
Deliverable 1: Analysis of data gathered while using models to consult at three start-up KI organizations
As described earlier in this proposal, the research team will focus on three start-up organizations during Year 2 of the project. They will be gathering and analyzing data about how their findings from Year 1 are useful for inventing work process in new organizations. Specifically, during this phase of the project, the research team will be collecting and analyzing information about how its software is used to envision alternative work processes by organization’s that have yet to realize stable patterns in their KI work activities. The data gathered during this year will be reported in the Year 2 ACM session (described below) and will be used in the Year 2 course materials/delivery (described below).

Deliverable 2: Materials for delivery of a course related to the modeling procedure/software being developed and tested
To disseminate knowledge about the work process modeling practices that underpin this research project, instructional materials will be developed to support course offerings at the three institutions involved. Course materials will include (a) a bibliography of sources informing the modeling process being developed, (b) instructional units with related exercises to educate students in procedures for modeling knowledge work, (c) semester projects with related write-ups that demonstrate and analyze activities related to the modeling work, and (d) development of internship opportunities for interested students to further extend their learning in non-class contexts.

In addition, a jointly developed, educational web forum will be created to distribute these materials and to foster the exchange of ideas and practices between students at the three institutions. The three investigators will be responsible for maintenance of the forum, ensuring the quality of material available there for future classes.

Deliverable 3: Offerings of this course at three institutions
Instruction using these materials will take place during Year 2 in seminar and special topics courses at the three institutions. In these courses, which are already routinely offered at the institutions, students expect to be exposed to leading edge theory and practice in the design of user-centered systems. At the three institutions, each the investigators serves as the director of a lab space within which many of the learning activities will take place.

The instructional experiences at these three institutions will be analyzed by the research team and used to refine similar instructional offerings in Year 3 and to develop the textbook being prepared in Year 3 (see below).

Deliverable 4: A presentation at the annual meeting of ACM SIGDOC
In October 2005, the research team will offer a panel discussion of their findings-to-date in the project, focusing on their efforts to migrate lessons learned from their studies of KI work in three mature organizations to work being conducting in three start-up organizations. As in Year 1, the paper generated for this panel presentation will be published in the annual proceedings of the conference and will be available online through the ACM Portal.

Deliverable 5: A beta version of the open-source software application for supporting the analysis of knowledge work in organizations
At the completion of Year 2 of the project, the research team will release a beta version of its software application. The team will solicit and use feedback from the development community, refining the application for its eventual full release at the end of Year 3.

Year 3
Deliverable 1: Delivery of a textbook manuscript for publication
Refining their instructional materials based on Year 2 experiences, the research team will prepare a textbook manuscript for publication by a publisher with related work (e.g., Wiley Computer Publishing, or the Morgan Kaufmann series in interactive technologies). The textbook will include access to and instruction in the released version of the software (see below).

Deliverable 2: A tutorial session at the annual meeting of ACM SIGDOC
Building on their panel presentations at the previous two annual meetings of ACM SIGDOC, the research team will in 2006 offer a day-long tutorial session regarding their modeling software and related processes. This tutorial session will include a refined and abbreviated version of the instructional materials the team began developing in Year 2 and included in the textbook manuscript being produced in Year 3.

Deliverable 3: Release of the open-source software application for supporting the analysis of knowledge work in organizations
At the completion of the Year 3 of the project, the research team will release an updated version of its open-source application as in each previous year of the project. However, this release will be
complemented by extensive set of instructional materials included in a manuscript to be submitted to a publisher as described above.

**Future Plans**

If this project is successful, an additional proposal will be submitted to NSF to expand the scope of this study beyond the modeling of KI work related to proposal development. In the future, the research team is interested in researching patterns of activity and processes of modeling in a broader domain of KI work. In addition, the research team plans to offer an international symposium in modeling the complexities of knowledge work in the emerging Information Economy. With the credentials developed through the first three-year phase of this project, the researchers believe they would be able to attract participation in such a symposium, perhaps as early as 2007.

**Impact**

The software developed as part of this project will advance knowledge about the centrality of communicative events in the knowledge work that underpins our emerging digital society. It will, in fact, provide a new way of understanding how work is accomplished in digitally-mediated, dynamic organizations where traditional hierarchical structures and pre-planned task flows no longer provide adequate explanations of how large-scale projects are accomplished. The software is specifically being envisioned as a tool for understanding work practices and for anticipating what choices knowledge workers might make to refine the communicative activities in which they engage to realize generalizable goals in recurrent situations.

Further, the project will integrate technical and social theories/methods into the software, advancing an understanding of how technical and social knowledge can work together to complement the objectives of human participants. Specifically, the underlying logic of the software assumes that while humans always retain agency in how they decide to employ communicative events in their work activities, they often choose to structure such events in a typified fashion. By allowing participants to see communicative events at both the macro and micro levels and to view such patterns across many different historical instances, the technology allows for alternative scenarios to be visualized and factored into emerging plans.

Finally, the project helps to create and foster a coherent interdisciplinary community of practice out of different groups doing research on the digital society. By bringing together researchers from three institutions, representing three research perspectives (i.e., human-computer interaction, computers and writing, and technical communication) the project draws on new combinations of technical and social knowledge. The dissemination of knowledge during the project will also foster interdisciplinary connections as students from each of the programs participate and collaborate in the planned classes and then as each of the investigators represents the project and its findings in discipline-specific forums.

Beyond that, however, is the promise that this project will garner attention from several disciplines as the software becomes available to anyone interested in instruments for studying work processes.
Works Cited


