Modeling Knowledge Work

Clay Spinuzzi
University of Texas at Austin

William Hart-Davidson
Rensselaer Polytechnic Institute

Mark Zachry
Utah State University

5 May 2004

Keywords: genre sets, genre systems, genre repertoires, genre ecologies
Abstract: Many models have been proposed for examining work. However, work models have typically assumed a strongly determined work structure and have often filed idiosyncracies and innovations under "exception handling." In this white paper, we propose two models, communication event models and genre ecology models, that can be paired to productively model and study knowledge work without assuming a strongly determined work structure. These models are intended not just for researchers but also for workers: ideally, they should help workers to examine and mediate their own knowledge work.

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.[1] 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.[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."[3] 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.

[1]Hecker, D. E. (2001)

[2]Hecker, page 2

[3]United States Department of Labor Statistics.

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.[4] 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."[5]

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

We propose 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 and activity theory.[6]

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, documents, or combinations of these such as "active documents," or process/constraint representations.[7] 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 have been developed and subsequently analyzed in workplace settings.[8] Dourish et. al.'s overall

[4]Julian, H. (2002)

[5]Boyd, A. (2003): http://whitepaper s.informationweek.com/detail/RES/ 1066151826_999.html

[6] For distributed cognition, see Cole & Engestrom, 1993; Hutchins, 1995; Winsor, 2001. For activity theory, see Bazerman, 1997; Engestrom, 1999; Nardi, 1996; Spinuzzi, 2003b

[7]For "active documents," Dourish, et al (2000). For process constraint representations, see Bernstein, A. (2000). ACM.

[8] See studies by Dourish, et. al., 1996, 1999a, 1999b, 2000.

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 and our attempts to develop modeling methods which represent writing practices.[9] 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.[10] Researchers of workplace writing have shown that it is not only the skillful employment of genres that leads to success (e.g. Spilka), but it is also the ability to understand knowledge work as successful communication which is important.[11] 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.[12]

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," 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.

Using 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

[9]For our studies of workplace writing, see Spinuzzi, 2003a, 2003b; Spinuzzi & Zachry, 2000; Zachry, 2001. For representing writing practices, see Hart-Davidson, 2002a; 2002b; Spinuzzi, 2002; Hart-Davidson, 2003
[10]See Bazerman, 1988; Smart, 2003; Mirel, 2003

[11] See Freedman & Smart, 1997; Geisler, 2001; Smart, 2000

[12]Dias, P., Freedman, A., Medway, P., and Paré, A. (1999).

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.

Figure 1 (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.?
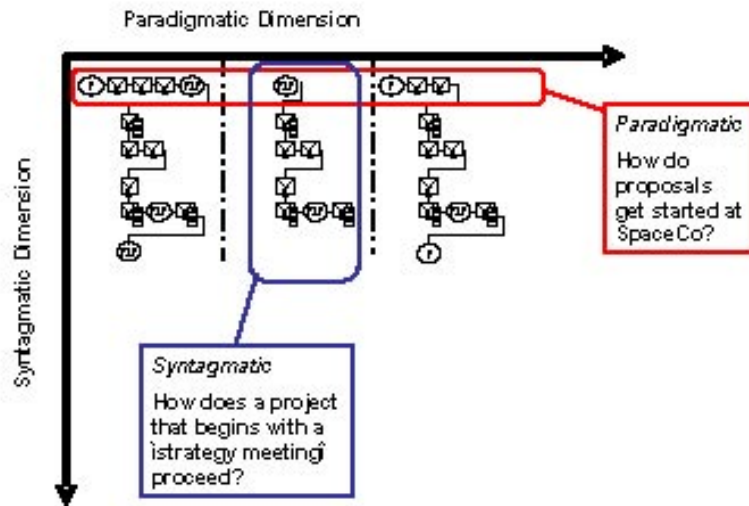
Figure 1. Writing Activity in Two Dimensions: Conditions for Sequencing and Substituting Genres

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. In the future, we plan 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.

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 in Figure 2 depicts a routine writing activity in an organization: the development of a quarterly human-resource allocation report. Adapting a convention proposed by Gunnarsson, 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.

Figure 2. A Progress Sort of the Communication Events in a Routine Report Task (Hart-Davidson 2002b, p. 463)

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. Figure 3 shows a listing of the data categories for the communication event object developed in our proof-of-concept work on CEMs.[13] 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.

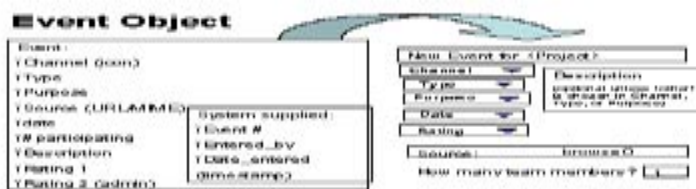[13]Hart-Davidson 2002a, 2002b, 2003

Figure 3 Attributes and UI Mock-Up for Communication Event Object

Hart-Davidson 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.[14] 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.

[14]Spinuzzi, 2003b

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, information ecologies, tool ecologies, genre sets, genre systems, and genre repertoires.[15] Without getting into the differences, we recognize a rich set of theoretical frameworks and empirical work that inform the genre ecology framework.

[15]For datacloud, see Johnson-Eilola, 2001; for information ecologies, Nardi & O'Day, 1999; for tool ecologies, Hutchins, 1995; genre sets, Devitt, 1991; genre systems, Bazerman 1994 and Yates & Orlikowski 2002; and genre repertoires Orlikowski & Yates, 1994.

We have modeled genre ecologies by hand to represent the genres on which groups of workers drew during various field studies.[16] 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

[16] Spinuzzi, 2002; 2003a, b

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)

Figure 5. A formal genre ecology model showing specific, directional mediational
relationships (Spinuzzi 2002, p. 205)

Conclusion

We believe that by pairing these two models, we can model knowledge work in ways that can help researchers and workers alike to visualize work in productive ways, without falling into the trap of overrationalizing or overdetermining work (as has often been the case with modeling techniques). Our goal is to use these models sensitively, to examine knowledge work with its contingencies and idiosyncracies, and to look for patterns without mistaking those patterns for the whole of the work. In the future, we plan to develop these models through coordinated research and application development.

Works Cited

Bazerman, C. (1988). *Shaping written knowledge: The genre and activity of the experimental article in science.* University of Wisconsin Press, Madison, WI.

Bazerman, C. (1994). Systems of genre and the enactment of social intentions. In Freedman, A. and Medway, P., editors, *Genre and the new rhetoric* (pp. 79-99). Taylor & Francis, London; Bristol, PA.

Bazerman, C. (1997). Discursively structured activities. *Mind, Culture, and Activity*, 4(4):296-308.

Bernstein, A. (2000). How can cooperative work tools support dynamic group processes? Bridging the specificity frontier. *Proceedings of Computer Supported Cooperative Work '00* (pp. 279-288). Philadelphia, PA: ACM.

Boyd, A. (2003). Improving office workflow to reduce costs and increase productivity. InformationWeek. Multimedia webcast. http://whitepapers.informationweek.com/detail/RES/1066151826_999.html

Casati, F., Ceri, S., Paraboschi, S., and Pozzi, G. (2000). Specification and implementation of exceptions in workflow systems. *ACM Transactions on Database Systems* 24(3):405–451.

Cole, M. and Engeström, Y. (1993). A cultural-historical approach to distributed cognition. In Salomon, G., editor, *Distributed cognitions: Psychological and educational considerations* (pp. 1-46). Cambridge University Press, Cambridge.

Devitt, A. J. (1991). Intertextuality in tax accounting: Generic, referential, and functional. In Bazerman, C. and Paradis, J. G., editors, *Textual dynamics of the professions: Historical and contemporary studies of writing in professional communities* (pp. 336-357). University of Wisconsin Press, Madison, WI.

Dias, P., Freedman, A., Medway, P., and Paré, A. (1999). *Worlds Apart: Acting and Writing in Academic and Workplace Contexts.* Mahwah, NJ: Lawrence Erlbaum.

Dourish, P., Bentley, R., Jones, R., and Maclean, A. (1999a). Getting some perspective: Using process descriptions to index document history. *Proceedings of GROUP'99* (pp. 375-384). Phoenix, AZ: ACM.

Dourish, P., Edwards, W. K., LaMarca, A., Lamping, J., Petersen, K., Salisbury, M., Terry, D. B., & Thornton, J. (2000). Extending document

management systems with user-specific active properties. *ACM Transactions on Information Systems* 18(2):40–170.

Dourish, P., Edwards, W. K., LaMarca, A., and Salisbury, M. (1999b). Using properties for uniform interaction in the presto document system. *CHI Letters 1*(1):55-64.

Dourish, P., Holmes, J., MacLean, A., Marqvardsen, P., and Zbyslaw, A. (1996). Freeflow: Mediating between representation and action in workflow systems. *Proceedings of Computer Supported Cooperative Work 96* (pp. 190-198). Cambridge, MA: ACM.

Engeström, Y. (1999). Expansive visibilization of work: An activity-theoretical perspective. *Computer Supported Cooperative Work*, 8:63-93.

Freedman, A. and Smart, G. (1997). Navigating the current of economic policy: Written genres and the distribution of cognitive work at a financial institution. *Mind, Culture, and Activity*, 4(4):238-255.

Geisler, C. (2001). Textual objects: Accounting for the role of texts in the everyday life of complex organizations. *Written Communication*, 18(3): 296-325.

Gunnarsson, B. L. (1997). The writing process from a sociolinguistic viewpoint. *Written Communication* 14(2): 139-187.

Hart-Davidson, W. (2002a). Turning reflections into technology: Leveragingtheory and research in the design of communication software. *Proceedings of the International Professional Communication Conference* (pp. 455-467). Portland, OR: IEEE.

Hart-Davidson, W. (2002b) Modeling document-mediated interaction. *Proceedings of the ACM SigDoc 2002 Conference* (pp. 60-71). Toronto: Association of Computing Machinery.

Hart-Davidson, W. (2003). Seeing the project: Mapping patterns of intra-team communication events. In *ACM SIGDOC 2003 Conference Proceedings* (pp. 28-34). ACM, Inc., New York.

Hecker, D. E. (2001). Employment outlook 2000-10: Occupational employment projections to 2010. *Monthly Labor Review* November 2001. United States Department of Labor: Office of Occupational Statistics and Employment Projections, Bureau of Labor Statistics.

Hutchins, E. (1995). *Cognition in the wild*. MIT Press, Cambridge, MA.

Johnson-Eilola, J. (2001). Datacloud: Expanding the roles and locations of information. In *SIGDOC 2001 Conference Proceedings* (pp. 47-54). ACM, Inc., New York.

Julian, H. (2002). Strategies and capabilities of professional services firms that offer content management services. *IDC Opinion Volume 1*. IDC Corporation. IDC #28377.

Mirel, B. (2003). *Interaction design for complex problem solving: Getting the work right*. San Francisco: Morgan Kaufman.

Nardi, B. A. and O'Day, V. L. (1999). *Information ecologies: Using technology with heart*. MIT Press, Cambridge, MA.

Nardi, B. A., editor (1996). *Context and consciousness: Activity theory and human-computer interaction*. MIT Press, Cambridge, MA.

Orlikowski, W. J. and Yates, J. (1994). Genre repertoire: The structuring of communicative practices in organizations. *Administrative Science Quarterly*, 39:541-574.

Smart, G. (2000). Reinventing expertise: Experienced writers in the workplace encounter a new genre. In Dias, P. and Paré, A., editors, *Transitions: Writing in academic and workplace settings* (pp. 223-252). Hampton Press, Cresskill, NJ.

Smart, G. (2003). A central bank's combined use of written/oral genres and technology to orchestrate its communications strategy in the arena of public policy. In Bazerman, C. and Russell, D. R., editors, *Writing selves, writing society*. http://wac.colostate.edu/books/selves_societies/.

Spilka, R. (1993). *Writing in the workplace: New research perspectives*. Southern Illinois University Press, Carbondale.

Spinuzzi, C. (2002). Modeling genre ecologies. In *ACM SIGDOC 2002 Conference Proceedings* (pp. 200-207). ACM, New York.

Spinuzzi, C. (2003a). Knowledge circulation in a telecommunications company: A preliminary survey. In *ACM SIGDOC 2003 Conference Proceedings* (pp. 178-183). ACM, New York.

Spinuzzi, C. (2003b). *Tracing genres through organizations: A sociocultural approach to information design*. MIT Press, Cambridge, MA.

Spinuzzi, C. and Zachry, M. (2000). Genre ecologies: An open-system approach to understanding and constructing documentation. *Journal of Computer Documentation*, 24(3):169-181.

United States Department of Labor Bureau of Labor Statistics. (2001). New and emerging occupations. *Issues in labor statistics*. Summary 01-06, December 2001.

Winsor, D. A. (2001). Learning to do knowledge work in systems of distributed cognition. *Journal of Business and Technical Communication,* 15(1):5-28.

Yates, J. and Orlikowski, W. (2002). Genre systems: Structuring interaction through communicative norms. *Journal of Business Communication*, 39(1): 13-35.

Zachry, M. (2001). Constructing usable documentation: A study of communicative practices and the early uses of mainframe computing in industry. *Journal of Technical Writing and Communication*, 31(1):61-76.