# Coping with Errors: the Importance of Process Data in Robust Sociotechnical Systems

**Michael B. Twidale**
Graduate School of Library and Information Science
University of Illinois at Urbana-Champaign
Champaign, IL 61820 USA
+1 217 265 0510
twidale@uiuc.edu

**Paul F. Marty**
Graduate School of Library and Information Science
University of Illinois at Urbana-Champaign
Champaign, IL 61820 USA
+1 217 265 0475
p-marty@uiuc.edu

**ABSTRACT**

This paper presents an analysis of written and electronic records that document the collaborative process of packing museum artifacts in preparation for a move. The majority of data recorded detailed the *process* of packing, while only a small amount of the data concerned which artifacts were packed in which boxes. Museum staff members were able to use these process data to solve the numerous errors that occurred during packing. We explore the design implications for collaborative systems which focus on supporting error recovery rather than error prevention.

**Keywords**

Workflow, error analysis, workplace study, problem solving, process data

## INTRODUCTION

People make mistakes, and computer systems need to be designed to help people cope with the results of errors. A robust system needs to be able to handle many different kinds of exceptions. Errors, especially those that are avoidable, form a particularly interesting category of exceptions. However, the examination of errors in work practices has negative connotations and has therefore not received the attention that it merits. Despite this, system designers must embrace the inevitability of errors: a system that assumes the infallibility of its operators will clearly not be sufficiently robust.

This paper examines the capture and use of information management data in a complex sociotechnical system designed to govern workflow. It presents an analysis of errors within this system and demonstrates how the management of those individual errors led to system-wide improvements in detecting and remediating errors as well as supporting future error prevention.


*CSCW'00*, December 2-6, 2000, Philadelphia, PA.


For the purposes of this study, we had access to a museum in the midst of a major transition in which staff members were moving 30,000 museum artifacts from old to new museum facilities. This museum provides an excellent example of an organization which is not only managing large quantities of information but also collecting extensive amounts of data about information management. This organization was faced with continually evolving processes, with many museum employees collaborating in the manipulation of data including museum artifacts as well as paper and electronic records. For the past four years, we have been studying the activities of the museum staff as they prepared, inventoried, packed, and moved the museum's collections; we are currently continuing our observations as the museum staff unpacks and prepares to open the new facility. During this time, we gathered longitudinal data on the museum's evolving information systems, including all paper records and electronic database versions, providing us with a continuous picture of the changing state of the museum over time.

We were interested in how the museum's information systems (computerized and paper-based) help manage the museum's evolving work practices, allow for continual policy improvement, and coordinate information sharing among the museum's employees. Using data that documented the day-to-day tasks of the museum's staff, we analyzed changes that occurred in their work practices over time. One issue that emerged from the study was how an explicit acknowledgement of the inevitability of errors leads to a set of activities that can deal with them robustly. The museum under study represents an environment extremely prone to human error; moreover, all such errors are potentially very serious for the museum and must be immediately resolved for the security of the museum's collections. Rather than attempting to foresee and eliminate all errors from the workflow in the museum, elaborate mechanisms, including the careful recording of process information to support the diagnosis and remediation of errors, evolved to cope with problems.

The initial results of our study at this museum prompt us to ask the following question: in an environment where errors are both likely to occur and costly to the organization, what kinds of data should be gathered about information management and workflow so that policies can be improved and mistakes easily corrected after errors are detected? This question is relevant not only to museums but to all organizations which manipulate data that may contain hidden errors. We explore the implications of acknowledging the need to cope with errors when designing robust collaborative systems.

## SOCIOTECHNICAL SYSTEMS AT THE SPURLOCK MUSEUM

The Spurlock Museum at the University of Illinois in Urbana-Champaign is a typical university museum: it is understaffed, underfunded, and possesses an eclectic collection of 30,000 cultural heritage artifacts from around the world and throughout history. In 1996, the University Board of Trustees approved the necessary funds to build a new museum facility. The new building was completed in April 2000 and is expected to open to the public in early 2002.

The construction of a new home for the museum's collections represents a major change for the staff at the Spurlock Museum. They face the daunting task of moving the museum across campus from its current location in a 15,000 square foot attic in a one-hundred year-old building to a 53,000 square foot modern museum facility. Moreover, although funds were made available to build the facility, few additional funds were allocated to support the task of moving the collections. The museum staff had to perform three major tasks to prepare for the move. First, they had to conduct an inventory of the museum's entire collections, the first complete inventory in the museum's history (the museum conducted a partial inventory in 1972). Second, they had to design exhibits for the museum's new and much larger galleries. Finally, they had to pack the museum's 30,000 artifacts to be transported to their new home across campus. Most museums take far more time to accomplish far fewer tasks. The fact that these tasks had to be performed simultaneously over a period of three years with few additional personnel or resources necessitated a strong collaborative effort.

In order to manage these complex and interrelated tasks, staff members at the Spurlock developed a new information system designed to support knowledge sharing among the employees at the museum [26]. There were many different potential users of this system, from undergraduate students employed by the museum to university professors assisting in exhibit design. The information system had to exist as an integral part of the larger system of policies and procedures that governed the process of moving the museum itself. Moreover, the system had to allow for the constant interaction between people, artifacts, and data in an environment that was recreating its goals, procedures, and policies on an almost daily basis. In sum, the situation represented a constantly evolving and extremely complicated sociotechnical system, with museum staff members, paper and electronic records, and museum artifacts all acting as integral components.

The fact that many different people were performing many different tasks while collaborating with many other individuals in a constantly changing complex process with tremendous amounts of data of various types had the potential for very serious consequences [27]. Spurlock staff members had no prior experience in dealing with projects anywhere near this size or of this complexity. The large number of individuals collaborating on this project with various levels and areas of expertise meant that there was a great potential for errors to occur. Because of this concern and because the system was constantly evolving as the project proceeded, museum staff members decided early on in the project to track as much data as possible about the process of moving the museum, especially the process of packing the museum's collections.

## TRACKING WORKFLOW PROCESSES: GATHERING DATA WHILE PACKING ARTIFACTS

While preparing the museum for the move to the new facility, staff members gathered much data about the artifacts in their collections. They also recorded extensive information about what they were doing to the artifacts while they were preparing the artifacts for the move; we call this information "process data." This kind of data is difficult for many organizations to record. It is expensive to document and ephemeral in nature. The information that Spurlock staff members recorded about the process of preparing their museum to move is the sort of information that normally gets lost in an organization. Few individuals see the need to save such data, given that it is usually of no value once the task at hand is completed. The actions of the Spurlock staff stand in contrast to usual procedures. For the purposes of this study, we have chosen to focus on just one small section of the available process data: data about the process of packing the collections.

The physical packing of the museum's collections began officially in November 1998 and concluded in April 2000. During this time period approximately 2000 boxes were packed; these boxes were custom created and individually designed for each artifact in the museum's collections. Extensive documentation, both paper and electronic records, was recorded for each box. We examined these records to learn more about the process of managing information concerning the packing of the museum's collections.

As might be expected, the packing records document such basic facts as "artifact x is packed in box y." Surprisingly, however, they also document approximately seven to ten times as much data about the *process* of packing artifact x into box y. The primary paper form (the Packing Data Sheet) completed by the museum staff while packing each box clearly shows this disparity. There are 25 fields on this form that need to be filled out while a box is being packed. Of these fields, only three (the identification number of the box, the contents of the box, and the location of the box) deal with factual data about the box and its artifacts. The other 22 fields record such information as the manufacturing methods used in creating this box, the packing methods used while packing the artifacts, and the types of problems encountered while packing the box. Additionally, these process data fields are used to track information about "who did what, when" to the box: who packed it, who supervised the packing process, who moved it into storage, who entered this information into the computer, who double-checked the data entry, etc.

When we examined the electronic records, we found an even greater ratio between factual and process data. The digital records about each box were initially created from the paper forms completed during the packing process and featured similar data fields. However, these records then grew in size over the months as museum staff members manipulated boxes in storage. The electronic records tracked each time a box was moved within the museum, each time an individual performed a contents inventory for a box, and each time a box was re-examined because of a suspected error in the packing process. This additional data documented the process by which museum staff members interacted with the packed boxes while preparing them for the move. An electronic modification history associated with each box would record such events as "problem alerts," "contents inventories," or "problem resolutions," noting details such as who first noticed a problem, who followed up on it, and how errors were corrected. The recording of this information dramatically increased the amount of process data maintained on each box in the museum's database systems. As problems were identified and resolved, even more data was generated, from another entry in a database to another slip of paper affixed to a box.

In retrospect, these records represent an amazing amount of work; museum staff members expended a great deal of effort gathering data that would be of little use to them after the move is completed. Once boxes are opened, artifacts removed, and boxes destroyed or recycled, the fact that a given box was once moved from storage zone nine to storage zone fifteen on the fifth of December 1999 by a certain student employee will be essentially obsolete data. Given this situation, why did the Spurlock Museum staff members bother to record so much data about the packing process?

To answer this question, let us imagine that the only data available after an artifact is packed is the basic fact that "artifact x is packed in box y." Then let us further imagine that, unnoticed at the time, something went wrong while "artifact x" was being packed: a number was misread, a contents inventory was inaccurate, or the artifact was incorrectly identified. And finally, let us imagine that one day "box y" is unpacked and it is discovered that "artifact x" is not there. Where is it? How can it be found? If the only fact available to the Spurlock Museum staff was that "artifact x is packed in box y," and that fact was suddenly discovered to be wrong, then the staff would have had no other data is available that could help them track down where the error occurred or provide clues as to where the lost item actually is.

How serious a problem would this have been for the Spurlock staff members? The museum has over 30,000 items packed in over 2000 boxes. Out of those 2000 boxes, twenty percent have had to be re-examined or re-opened for one reason or another. Twenty percent of all boxes packed had some kind of problem, ranging from misread accession numbers to inconsistent counts on inventory sheets. All of these problems were potentially serious ones that could have resulted in lost artifacts. How could the museum have possibly coped with such a high error rate without having collected the ephemeral process data?

## COPING WITH ERRORS: ERROR RECOVERY VS. ERROR PREVENTION

Information about the process of packing proved to be of vital importance to the museum staff members. Given the magnitude of the task at hand, they knew that they faced the potential for an extremely high number of errors. In almost any information system, errors are inevitable; moreover, even data once thought to be correct can become erroneous over time [46]. New errors arise, processes are redesigned, and information systems must evolve through these changes if they are going to successfully support communication and collaboration within an organization.

When faced with such a high probability that errors will occur, institutions normally focus on *preventing* errors by designing information systems that organize workflow into highly structured collaborative processes [1]. Such systems indicate to individuals working within the collaborative group when and what kinds of data to share with whom. This is especially true in situations where many different individuals must collaborate to achieve an overall goal while each performing smaller, individual tasks. Thus, the desire to prevent errors tends to result in systems for knowledge sharing that attempt to organize and even control collaborative processes [41].

At the Spurlock Museum, however, this could not be done; there was no way to provide *a priori* an organizational structure that could be embedded in an information system [40]. The museum professionals at the Spurlock Museum were, in essence, making the process up as they went along. They had no prior experience in moving a museum, they had no time for "test runs," and they had no money to hire consultants. Even if they could have hired experts, each museum's situation is unique and Spurlock staff members would not have been able to rely on specific instructions or guidance from their peers. For most tasks, they had to rely on undergraduate employees who, though diligent, hardworking, and reliable, could hardly be considered experts in inventorying artifacts, packing collections, or moving museums.

Moreover, even if Spurlock staff members had attempted to design in advance a rigid system that covered the possible conditions and circumstances that might be encountered during the packing process, any one artifact could have disrupted it. Every artifact that moved through the system was unique, every case was equally important, and every situation a possible exception to a rule. For any rule that might have created in an attempt to prevent errors by rigidly regulating the packing process, any one artifact could have broken it at any time.

The Spurlock staff had to be able to change their operating procedures at the drop of a hat. In addition, the museum's information system had be able to evolve quickly enough to keep up with the museum's constantly changing environment. The resulting system was fluid, constantly evolving, and highly unstructured; but it was also extremely prone to error.

One might think that advanced technology could help minimize errors in the packing process while still allowing the overall moving procedures to remain flexible, fluid, and adaptable. After all, the majority of the packing errors encountered by the Spurlock staff were attributable to transmission errors, misread numbers, or inaccurate data entry into a computer. Ignoring problems of expense, could not identification technologies such as bar coding or radio frequency tags attached to the artifacts have helped prevent such errors? Perhaps, but museum staff members would still have faced the initial error-prone task of coding each artifact's identification number into the system and onto the piece itself. Moreover, once the system was in place, there would simply have been new types of transmission errors to deal with. Even with high tech solutions [48], valuable as they are, museum staff members would still have faced errors.

Therefore, given the fact that initial error prevention through organizing workflow was impossible, a complex socio-technical system evolved that allowed the Spurlock staff to perform error recovery by focusing on tracking process data through the Spurlock's information management systems. It should be noted that the Spurlock staff members did not set out to design such a system. In fact, where they attempted to design systems to cope with errors directly, they actually tried to design systems to prevent errors from occurring. Not only did these attempts not work, but the elements of the system originally intended for error prevention evolved along with the rest of the environment until they themselves were used for error recovery.

For example, soon after they began packing the museum, staff members initiated a new system explicitly designed to reduce the number of packing errors. A system of "prepack checks" was implemented in an attempt to ensure that certain tasks which absolutely had to be completed before an artifact could be packed were in fact completed at the proper time. This system had both paper and electronic components: a series of check boxes was added to the main artifacts database system and a paper packing slip was created that followed each artifact from station to station. In theory, no artifact was to be packed unless all of these prepack checks were completed and the artifact was allocated a "ready to pack" status. However, given the rapid pace of change in the overall moving procedures in place at the museum, these prepack checks quickly became a burden rather than an aid. The attempt to enforce rigid rules about the packing process could not keep up with the rapidly evolving needs of the staff members actually involved in packing the collections. One by one, each rigid "prepack check" was skipped, made an exception, or flat out dropped as special cases arose that required special treatment or workarounds [15]. Eventually, the prepack checks stopped being used for their intended purpose; instead the data transformed into a way of tracking where an artifact was in the packing process rather than a way of assuring that each artifact be packed in the same rigid method, step by step. In this way, the prepack checks, both on the paper slips and online, were transformed into a valuable source of process data for the museum staff members. They documented in a clear fashion what steps in the packing process each artifact had actually completed and in what order these steps were taken. From their intended purpose of preventing errors, the prepack checks had evolved into valuable clues that could be read and interpreted by museum staff members working to recover from and correct packing errors.

We turn now to a couple of examples that demonstrate how process data was used by the museum staff members to solve packing problems in a collaborative fashion.

## COLLABORATIVE PROBLEM SOLVING WITH PROCESS DATA

As discussed above, the many steps in the packing process, combined with the many different individuals working on packing the collection, meant that there was a great deal of opportunity for error. Indeed, approximately twenty percent of boxes packed at the museum had to be re-examined and double-checked for some sort of error. Errors were especially dangerous when they occurred in the early stages of the packing process. Errors could propagate throughout the entire system, and the fact that many different people handled each individual artifact meant that tracking the original source of the error could be extremely difficult.

However, despite the high error rate, not a single artifact was permanently misplaced. In each instance, museum staff members were able to use data gathered during the packing process itself to backtrack the steps in packing any given artifact, determine the source of the packing error, and correct the problem. Because of this process data, no error came back through the system that could not be identified, tracked down, and corrected by the individuals working as a team to solve packing problems. The following examples are representative of the types of problems that occurred and the methods used in solving them.

One of the most common types of problems found in the records documenting packing problems is referred to as a "no such number" problem. For example, a student worker is packing a given artifact, accession number 22.4.11, into box 831. However, the student, reading the accession number off of the object, misreads a '1' as a '7' and records the artifact's identification number on the Packing Data Sheet as 22.4.71.

This form is then delivered to a student employee working in the Registration department to be entered into the database system. When this student enters 22.4.71 into the database he learns that no artifact with that accession number exists. The student writes "no such number" on the Packing Data Sheet, circling the offending number, and returns the sheet to the Assistant Collections Manager (who is in charge of solving packing problems). There are several possible explanations for this problem. For instance, no record could exist for this artifact because it was accidentally skipped during the inventory procedure. The student working on the computer in Registration could have mistyped the value into the system or misinterpreted the result. The artifact could even be mismarked with the wrong identification number. However, when the Assistant Collections Manager opens box 831, she discovers that the student packer had simply misread the inventory number and written the wrong number on the Packing Data Sheet. She corrects the number, changing it from 22.4.71 to 22.4.11, writes "number corrected" on the sheet, signs and dates her contribution, and returns it to the Registration department. There, another student employee will pick up the form, enter the new value into the system, and complete the problem resolution. Note that all of these actions are tracked and recorded as process data in the packing records (written and electronic) for this box.

Another very common type of packing problem is referred to as an "already packed" problem. For example, a student entering a Packing Data Sheet for box 1131 into the database system tries to record that artifact number 77.12.2 is packed in this box. However, the computer system informs her that 77.12.2 is apparently already packed in box 689. This is a more insidious type of problem, for the possible explanations have repercussions which could easily affect more than one box. In this case, the Assistant Collections Manager first checks the contents of box 1131, where she discovers that artifact 77.12.2 is indeed packed in box 1131. She records this fact on the Packing Data Sheet for box 1131 and then proceeds to open box 689. According to the inventory for this box, there should be ten artifacts packed in the box; a quick count confirms this fact. However, artifact 77.12.2 is not one of those ten artifacts. To discover the source of the problem, she physically compares each artifact in box 689 with the inventory for that box, checking off inventory numbers as she removes each artifact from the box. In doing so, she discovers the source of the error: a different artifact, accession number 77.1.22, was packed in box 689 but accidentally recorded on the Packing Data Sheet as artifact number 77.12.2. She records this information on the Packing Data Sheets and sends the results of her investigation to the Registration department, where the correct box locations for each artifact will be updated and a complete description of the Assistant Collections Manager's interactions with each box will be recorded.

Faced with problems such as these on a daily basis, museum staff members found the process data to be of immense value. It was considered so valuable that even when budget crises dictated continual cutbacks in the amount of data that could be gathered for each artifact, data about the packing process was never cut. The inventory process, for example, was eventually reduced to recording the bare minimum of data needed for artifact identification. In contrast, museum staff members continually added to the amount of process data gathered while artifacts were packed.

Why was this process data considered so valuable that it was kept when all other data was being cut? The simple answer is that process data did more than help solve packing problems. By recording this data, museum staff members were constantly documenting details of the packing process itself which, over time, helped implement changes that improved the moving procedures in the museum. In particular, having access to this process data actually improved the museum's problem solving procedures.

## EVOLVING WORK PRACTICES: CONTINUAL PROCESS IMPROVEMENT

Throughout the course of the packing process, many steps were taken to reduce the number of errors that occurred. Some modifications represented explicit changes made to the system by the museum professionals in an attempt to minimize errors; others represented changes that evolved in the system over time, inadvertently resulting in process improvements. We will now present two examples of process improvements that helped solve packing problems.

The first example documents a physical change to the packing process that also resulted in the improved accuracy in the transmission of artifact identification numbers. Analysis of the process data gathered while packing revealed that a major source of packing errors could be attributed to difficulties in reading, recording, and transmitting artifact identification or accession numbers. Many museum staff members (especially undergraduate student employees) had difficulty accurately reading the small accession numbers written on each of the museum's artifacts. These numbers, the only certain means of identifying some objects, were often difficult to read, vague, unclear, or ambiguous. Student packers would frequently record 18's as 81's or make other transposition errors. Errors made at this stage in the packing process would propagate throughout the system, possibly ruining the contents inventory for an entire box. A process innovation designed to solve a completely unrelated problem helped reduce inaccuracies caused by misread numbers. This other problem was that student workers assigned to pack the collections would often put off packing artifacts they perceived as being more difficult to pack or more fragile to handle. Artifacts believed to be more of a challenge would sit on the packing tables for weeks while other artifacts were packed almost immediately. To solve this problem, the Collections Manager instituted a new procedure in September 1999 about halfway through the packing of the museum's collections. In this new "Magic Fairy" procedure, she and the Assistant Collections Manager, the two museum employees most skilled in handling artifacts, would stay late at night and "prewrap" a day's worth of artifacts to be packed (the first step in packing any object). When the students returned to work the next morning, they found that a "magic fairy" had visited the museum during the night and left prewrapped artifacts for them to pack. These prewrapped objects were much less frightening for the students to pack; fragile-looking museum artifacts were transformed into easy-to-handle misshapen blobs of ethafoam. However, so that the student packers would know the accession number for each object, the Collections Manager and her assistant, the two museum employees most skilled in reading artifact accession numbers, would write the artifact's identification number on the outside of the prewrapped piece (the second step in packing any object). This they did clearly and carefully in black ink on yellow tape on the white ethafoam. Thus, with one innovation, they not only made the task of selecting artifacts to pack easier on the students but they also made the task of reading accession numbers much simpler by introducing expertise into a crucial phase of the process.

The second example illustrates how a minor change to the Packing Data Sheet resulted in more efficient packing problem solving. Over the lifecycle of a packed box, the Packing Data Sheet ends up in the hands of many different employees; each time a box is opened, a label examined, or an artifact checked, the person responsible for this act was expected to take the Packing Data Sheet from the box, note their action, and sign their name. Thus, like the electronic modification history associated with each digital record, these signatures formed a kind of history for the Packing Data Sheet [5, 20]. This record of discrete actions by different individuals with the same box over many months was of great value to the employees working on solving packing problems. It provided clues as to who had handled the box when and for what reason. But there were problems with this ad hoc process. Sometimes people would forget to provide a date. Sometimes, because of the number of signatures on the form, it would be difficult to determine whose name was associated with which action. One action/signature pair of great significance in the problem solving process was the "packing label check," which indicated whether a museum staff member had cross-checked the handwritten contents inventory list contained on the Packing Data Sheet with the computer-generated "Packing Label" to be affixed to the outside of the box and also listing the box's contents. A discrepancy between these two lists would likely indicate a data input error. Such a check was time-consuming to perform, and it was useful to know when and if this check had been performed in the past (especially when two or more packing labels had been generated for any one box because of previous errors with that same box). However, the large number of signatures on a given Packing Data Sheet combined with the lack of a formalized process for recording these signatures meant that there was a good chance that a packing label check might have been previously performed but go unnoticed by the problem solving team. To clear up potential confusion, in the Fall of 1999, the Assistant Collections Manager added a new field to the Packing Data Sheet, providing a clear location for individuals performing packing label checks to note their actions. This act of formalizing an already occurring informal activity helped employees interpret Packing Data Sheet notations more accurately and more quickly.

Whether such changes to the system evolved by accident or were instituted on purpose, the museum staff members were able to identify what was helping and what was not simply because they were tracking as much data as possible about the packing process. These examples demonstrate the value of tracking process data for improving work practices. What then are the implications this raises for CSCW researchers interested in designing collaborative systems for error detection and correction?

## RESEARCH QUESTIONS AND DESIGN IMPLICATIONS

The initial research question that must be addressed when designing a system for an error susceptible environment is whether it is better to focus one's efforts on error prevention or error recovery. Our experiences with the Spurlock Museum have led us to believe that procedures to allow for error recovery are at least as important in system design as methods for attempting error prevention. We see parallels between our study in a museum and various other studies of very different kinds of work.

Hughes et al. [21] note how a robust, reliable and trustable system for air traffic control arose from the activities of fallible individuals and technologies. The successful operation of this system depended on the seemingly inefficient cross-checking activities of air traffic control employees in much the same way that the seemingly inefficient recording of process data in the museum assured the robustness of their information systems. Flight strips played an important role in this robust system, affording the recording (and subsequent easy checking) of what has been done and by whom [17].

In a Finnish paper mill [4, 24], an electronic diary provided a means of communicating problems and potential resolutions between people who worked on a task at different times; like the Packing Data Sheets, this allowed error diagnoses and repairs to occur as part of the ongoing workflow. The diary was often used to facilitate awareness of the existence of problems without the need to explicitly direct the problem report to the proper individual. A relatively simple technology was easily incorporated into the activities of reporting errors and process improvement.

As in a study of workflow on a printing shopfloor [11], the workflow observed at the Spurlock has activity components with physical items (the artifacts) that are external to the computer systems used. However, unlike that study, the museum's databases merely required that users record the results of their activity. Compared with the print shop's workflow system, they were even less constraining in the ordering of the work done and therefore did not interfere with the smooth operation of day to day tasks. Furthermore, the overhead of use at the museum was relatively low.

Barthelmess and Wainer [6] explore exceptions and propose viewing workflows as exception handling tools. This involves providing support for different kinds of exception handling actions, including mechanisms for recording who did what. This approach often leads to solutions that require data that were incorrect or missing to be re-entered. However, with errors in the operation of the workflow system itself, consequent actions based on erroneous data can cause problems, so that the undoing of the actions is more complex and a simple redo is insufficient. Hence our concern in this paper was to concentrate on errors.

Saastamoinen [37] examined exceptions and noted errors as an important sub-category, amounting in some circumstances for 82% of all exceptions. He particularly noted problems caused by errors made by people outside the department under study. Such errors are harder to resolve because of the additional complexities of interacting with people from another department, even if they are in the same building. We would expect that a substantial number of errors also occur within any department (it seems unlikely that errors only ever occur somewhere else), but because these are easier to resolve, they are not regarded as being so problematic and thus are less visible.

Hartswood & Proctor's study [18] of radiography notes how breakdowns (including errors) lead to repair activity. They explore the consequences of cases when the people identifying problems are different from those who can fix them. Again, the preservation of contextual information was found to be important to effect repair. They note that: "…breast screening staff (say) may find it difficult to recognise the work that they do when presented in this way, precisely because the skills and resourcefulness that people learn on the job enable them to treat many breakdowns as routine." This emphasizes that an organization can robustly deal with certain types of errors, not by eliminating them but by having mechanisms for detection and remediation. Furthermore, it shows how those processes can be so much a part of normal activities that they are considered unremarkable.

Similarly, white collar work may contain explicit checks of work by senior staff (e.g. [8], [49]). Such organizations note the inevitability of error and so add checking routines into work practice. Wenger [49] notes how errors can be used in organizations to blame workers, which can turn an exploration of error patterns into a very controversial area of study.

There are parallels with work on the causes of accidents in safety-critical organizations [31, 38]. It important to try to understand the causes of accidents in order to prevent future ones. However, blame and particularly the politics of blame can get in the way. It is tempting to trace all causal chains of failures to human error. Instead, we need to consider a wider picture of the causes of such errors, such as deficiencies in training, poor interface design, and badly designed procedures and work arrangements. We must look for the organizational context of error which can lead to the errors of individuals having systemic effects [28]. Within the field of human reliability assessment there is an emerging acceptance that errors are unavoidable events. This leads to regarding people in the system not just as the cause of problems, but as contributing to overall system robustness. Hence cooperative working contributes to the overall reliability of sociotechnical systems [36]. The errors we have studied are not life threatening, nor do they lead to catastrophic system failures. Nevertheless, a concentration on kinds of errors and how poor organizational arrangements can increase error rates can be useful in white collar work.

Workflows can serve as useful organizing structures around which exceptions adapt. The workflow may embody what ought to happen (and what ought to have happened), but in cases of problems we also need information on what actually happened [7, 28], and perhaps on what the participants intended to happen. This understanding (which otherwise may soon be lost) can be valuable not only in reconstructing and recovering from exceptions, but also in redesigning workflows to minimize or eliminate error patterns. Recording large amounts of process information is costly [7], but as we found, can be justifiable. Thus we would advocate recording as much history as possible, not just that of unstructured exceptional activities.

Other than these examples (but see also [41, 42, 44, 45]), there seems to be relatively little attention paid in the CSCW literature to the role of errors as either a source of insight into current problems or an inevitable aspect of human performance that should be considered during systems design. The CSCW literature does focus, however, on such topics as exceptions to general rules caused by the impossibility of specifying all cases, the need for negotiated rules, the difficulty of articulating tacit knowledge, the inherent complexity of life, and the changing nature of work [40]. It would thus seem useful to add the study of errors into CSCW research, much in the way that error analysis and preventative design have been incorporated into research in Human Computer Interaction, and particularly Human Factors [22, 33, 34].

If we accept that error recovery is as important as error prevention, then several research questions should be considered. How does one go about collecting process data that can be successfully used for error recovery and problem solving? How much process data is enough to perform these tasks? How do individuals in the organization contribute their share of process data? How can knowledge about each step in a process be recorded so that human experts skilled in problem resolution can correct errors? How expensive is process data recording? Is having access to this data worth the time it takes to gather it?

The answers to many of these questions depend on the particulars of individual circumstances. However, any organization that depends on a fluid and continually evolving information system for its successful operation could benefit from the recording and tracking of process data. The more adaptable an information system has to be, the more likely it is that information managers will require new approaches, such as the acquisition of process data, in order to support the needs of their users.

Each of the research questions raised above represents social or organizational questions that given institutions must attempt to answer individually. However, each issue has associated technical questions that could lead to the development of general applications designed to support the gathering and interpretation of process data. As noted earlier, ubiquitous computing offers great potential for tracking the movement and use of physical elements of the workflow [48]. Possible technological innovations directly related to the recording of process data could include the use of logins to automatically track who changed what, when, at the field level in the system; one could incorporate a digital history of data modifications that tracks not only all previous values but also why each value was changed, when, by whom, and with what authority [19, 25]. Design rationale [29], which serves a similar role in software development, should be used as inspiration in developing systems with lightweight features to record this additional contextual information.

Such history-enriched digital data could allow advanced systems to provide visualizations that track processes and errors, displaying on a screen many of the connections and relationships usually held only in the minds of the systems' expert users [32, 43]. Using explicit representations of work processes to help in diagnosing errors has parallels with software debugging [7], and so we might consider the utility of analogous functions from a sophisticated debugging environment. In software debugging, the results of errors may only manifest later in the chain of execution, corrupting data along the way. Diagnosis involves stepping through the code and comparing what should have happened with what actually occurred (particularly as values in data structures are changed). We can envisage analogous functionalities for workflows, noting how the Spurlock's detailed recording of process information serves as a valuable trace of actual execution steps.

The importance of ensuring that workflow systems can handle large numbers of exceptional cases has been noted in prior studies [2, 13, 39] and has inspired work on developing highly flexible systems [1, 12, 16]. Some common exceptions can be predicted and mechanisms included in the workflow to cope with them. Adaptive workflow systems acknowledge that it is not possible to include within the original workflow model pre-defined conditional branches for all possible contingencies. Instead, there also needs to be support for modifying a process model once it has started executing [10, 23]. One important component of such systems is the ability to undo parts of the process and recover by redoing erroneous or anomalous steps [3, 14]. It is interesting to note that CSCW researchers are willing to acknowledge the possibility of errors occurring in the *specification* of workflows, and so design systems to support detection and remediation at that level [9]. As developers, we acknowledge our own fallibility [47], but perhaps because of understandable concerns for the misuse of error analysis of the behavior of less empowered workers, we are reluctant to consider that they too can and will make mistakes.

What was especially notable to us in the museum study was how frequently exceptions arose and how often processes were changed in the light of evolving practice. Thus not only do such systems have to be flexible, but they also have to be easily adaptable and provide clear indications to all users of the current evolutionary stage of the work process. We believe that certain features of the tools used in the workflow (paper forms, conventional databases and email) allowed people to understand the process enough to adapt the workflow in the light of evolving understanding of the causes of errors and the desire to improve the quality of work. These tools allow users to understand the effects of certain changes in use (e.g. what the result will be of adapting the use and meaning of a database field) even if they don't necessarily understand all the underlying mechanisms (how the database actually works) [30]. Thus they serve as Robinson's "common artefacts", affording easy evolution of use [35].

## CONCLUSION

We should operate under the assumption that no matter how hard we try to prevent errors, errors will always occur in collaborative work processes. Therefore, we should design information systems that incorporate mechanisms specifically intended to cope with errors. Our analysis has led us to conclude that successful error correction is not so much about increasing automation in computer-supported sociotechnical systems as it is about improving collaboration between the human and computer elements of these systems. Systems which only regulate user actions to prevent errors from occurring cannot easily support error recovery. Thus, it is more important to build systems which record data about processes so that process data is available for future error recovery than it is to build systems that dictate successful policies in an attempt to prevent errors from occurring at all.

Our findings have revealed that the effective use of process data requires the assistance of human experts skilled in the interpretation of such data. Although CSCW technologies can help gather data about a given process, they cannot interpret it; technologies, however, can support the analysis of this data. Extensive communication between members of the organization is necessary for solving complex problems; the more difficult the problem, the more important it is that mechanisms exist to record and communicate problems and changes in workflow over time.

Information systems must be able to support collaborative problem solving directly, not merely enforce rules designed to minimize errors. We cannot rely on systems that work only when rules are obeyed and standard procedures followed. Such systems can cause organizations to lose valuable data when the never before considered rare exception occurs. As CSCW researchers, we must not only organize workflow systems, we must track them, and then make the resulting data available to experts who can use them to solve other problems. From such examples, experts will be able to analyze the ways in which errors develop and then recommend modifications to existing sociotechnical systems that may minimize future errors.

## REFERENCES


1. Abbott, K.R. and Sarin, S.K. Experiences with workflow management: issues for the next generation. In *Proceedings, CSCW'94*, (Chapel Hill NC, 1994), ACM Press, 113-120.
2. Agostini, A., De Michaelis, G., Grasso, M.A. and Patriarca, S. Reengineering a business process with an innovative workflow management system: a case study. In *Proceedings, COOCS'93. Conference on Organizational Computing Systems*, (Milpitas, CA, 1993), ACM Press, 154-165.
3. Agostini, A. and De Michelis, G.A. Light Workflow Management System Using Simple Process Models. *Computer Supported Cooperative Work: The Journal of Collaborative Computing*, *9* (3-4). 335-363.
4. Auramäki, E., Robinson, M., Aaltonen, A., Kovalainen, M., Liinamaa, A. and Tuuna-Väiskä, T. Paperwork at 78kph. In *Proceedings, CSCW'96*, (Boston, MA, 1996), ACM Press, 370-379.
5. Baker, N. Discards. *The New Yorker*, *70* (7). 64-86.
6. Barthelmess, P. and Wainer, J. Workflow Systems: a few definitions and a few suggestions. In *Proceedings, COOCS'95. Conference on Organizational Computing Systems*, (Milpitas, CA, 1995), ACM Press, 138-147.
7. Blumenthal, R. and Nutt, G.J. Supporting unstructured workflow activities in the Bramble ICN system. In *Proceedings, COOCS'95. Conference on Organizational Computing Systems*, (Milpitas, CA, 1995), ACM Press, 130-137.
8. Blythin, S., Rouncefield, M. and Hughes, J.A. "Never mind the ethno stuff, what does all this mean and what do we do now?" Ethnography in the commercial world. *interactions*, *4* (3). 38-47.
9. Bogia, D.P. and Kaplan, S. Flexibility and control for dynamic workflows in the wOrlds environment. In *Proceedings, COOCS'95. Conference on Organizational Computing Systems*, (Milpitas, CA, 1995), ACM Press, 148-159.
10. Borgida, A. and Murata, Y. Tolerating exceptions in workflows: a unified framework for data and processes. In *International Joint Conference on Work Activities, Coordination and Collaboration (WACC'99)*, (San Francisco, CA, 1999), ACM Press, 59-68.
11. Bowers, J., Button, G. and Sharrock, W. Workflow from Within and Without. In *Proceedings, European Conference on Computer-Supported Cooperative Work, ECSCW'95*, (Stockholm, Sweden, 1995), Kluwer, 51-66.
12. Dourish, P., Holmes, J., MacLean, A., Marqvarsdsen, P. and Zbyslaw, A. Freeflow: mediating between representation and action in workflow systems. In *Proceedings, CSCW'96*, (Boston, MA, 1996), ACM, 190-208.
13. Ellis, C.A., Gibbs, S.J. and Rein, G.L. Groupware: Some issues and experiences. *Communications of the ACM*, *34* (1). 38-58.
14. Faustman, G. Configuration for adaptation - a human centered approach to flexible workflow enactment. *Computer Supported Cooperative Work: The Journal of Collaborative Computing*, *9* (3-4). 413-434.
15. Gasser, L. The integration of computing and routine work. *ACM Transactions on Information Systems*, *4* (3). 205-225.
16. Glance, N.S., Pagani, D. and Pareschi, R. Generalized Process Structure Grammars GPSG for Flexible Representations of Work. In *Proceedings, CSCW'96*, (Boston, MA, 1996), ACM Press, 180-189.
17. Harper, R.H.R. and Hughes, J.A. 'What a f-ing system! send 'em all to the same place and then expect us to stop 'em hitting'. Making technology work in air traffic control. In Button, G. ed. *Technology in working order: studies of work, interaction, and technology*, Routledge, London, 1993, 127-148.
18. Hartswood, M. and Proctor, R. Design guidelines for dealing with breakdowns and repairs in collaborative work settings. *International Journal of Human-Computer Studies*, *53* (1). 91-120.
19. Hill, W.C. and Hollan, J.D. History-enriched digital objects: prototypes and policy issues. *The Information Society*, *10* (2). 139-145.
20. Hughes, J.A., King, V., Mariani, J.A., Rodden, T. and Twidale, M.B. Paperwork and its lessons for database systems: an Initial Assessment. In Shapiro, D., Tauber, M. and Traunmueller, R. eds. *The Design of Computer Supported Cooperative Work and Groupware Systems*, North Holland, Amsterdam, The Netherlands, 1996, 43 - 66.
21. Hughes, J.A., Randall, D. and Shapiro, D. Faltering from ethnography to design. In *Proceedings, CSCW'92*, (Toronto, 1992), ACM Press, 115-122.
22. Jones, P.M. Human error and its amelioration. In Sage, A.P. and Rouse, W.B. eds. *Handbook of Systems Engineering and Management*, John Wiley, 1999, 687-702.

23. Klein, M., Dellarcos, C. and Bernstein, A. Introduction to the special issue on adaptive workflow systems. *Computer Supported Cooperative Work: The Journal of Collaborative Computing*, *9* (3-4). 265-267.

24. Kovalainen, M., Robinson, M. and Auramäki, E. Diaries at Work. In *Proceedings, CSCW'98*, (Seattle, WA, 1998), ACM Press, 49-58.

25. Mariani, J.A. and Rodden, T. The Impact of CSCW on Database Technology. In *Proceedings, IFIP International Workshop on CSCW*, (Berlin, Germany, 1991), 146-161.

26. Marty, P.F. Museum Informatics and Collaborative Technologies. *Journal of the American Society for Information Science*, *50* (12). 1083-1091.

27. Marty, P.F. Museum Informatics: Sociotechnical Infrastructures in Museums. *Bulletin of the American Society for Information Science*, *26* (3). 22-24.

28. McCarthy, J.C., Healey, P.G.T., Wright, P.C. and Harrison, M.D. Accountability of work activity in high-consequence work systems: Human error in context. *International Journal of Human-Computer Studies*, *47* (6). 735-766.

29. Moran, T.P. and Carroll, J.M. (eds.). *Design Rationale: Concepts, Techniques, and Use*. Lawrence Erlbaum Associates, Mahwah, NJ, 1996.

30. Nardi, B.A. and Miller, J.R. Twinkling lights and nested loops: distributed problem solving and spreadsheet development. *International Journal of Man-Machine Studies*, *34* (2). 161-184.

31. Perrow, C. *Normal accidents: Living with high-risk technologies*. Basic Books, New York, NY, 1984.

32. Pycock, J., Palfreyman, K., Allanson, J. and Button, G. Representing Fieldwork and Articulating Requirements Through VR. In *Proceedings, CSCW'98*, (Seattle, WA, 1998), ACM Press, 383-392.

33. Rasmussen, J. *Information processing and human-machine interaction: an approach to cognitive engineering*. Elsevier Sciences, Amsterdam, 1986.

34. Reason, J. *Human Error*. Cambridge University Press, Cambridge, 1990.

35. Robinson, M. Design for Unanticipated Use..... In *Third European Conference on Computer-Supported Cooperative Work - ECSCW'93*, (Milan, Italy, 1993), 187-202.

36. Rognin, L., Salembier, P. and Zouinar, M. Cooperation, reliability of socio-technical systems and allocation of function. *International Journal of Human-Computer Studies*, *52* (2). 357-379.

37. Saastamoinen, H.T. Case study on Exceptions. *Information Technology and People*, *8* (4). 48-78.

38. Sagan, S.D. *The Limits of Safety*. Princeton University Press, Princeton NJ, 1993.

39. Schael, T. Workflow management systems for financial services. In *Proceedings, COOCS'93. Conference on Organizational Computing Systems*, (Milpitas, CA, 1993), ACM Press, 142-153.

40. Shipman, F.M. and Marshall, C.C. Formality considered Harmful: experiences, emerging themes and directions on the use of formal representations in interactive systems. *Computer Supported Cooperative Work: The Journal of Collaborative Computing*, *8* (4). 333-352.

41. Strong, D.M. and Miller, S.M. Exceptions and exception handling in computerized information processes. *ACM Transaction on Information Systems*, *13* (2). 206-233.

42. Suchman, L. Office procedures as practical action: models of work and system design. *ACM transactions on Office Information Systems*, *1* (4). 320-328.

43. Swenson, K.D. Visual support for reengineering work processes. In *Proceedings, COOCS'93. Conference on Organizational Computing Systems*, (Milpitas, CA, 1993), ACM Press, 130-141.

44. Symon, G., Long, K. and Ellis, J. The coordination of work activities: cooperation and conflict in a hospital context. *Computer Supported Cooperative Work: The Journal of Collaborative Computing*, *5* (1). 1-31.

45. Trepess, D. and Stockman, A. A classification and analysis of erroneous actions in computer supported co-operative work environment. *Interacting with Computers*, *11* (6). 611-622.

46. Twidale, M.B. and Marty, P.F. An Investigation of Data Quality and Collaboration. Technical Report, GSLIS, University of Illinois at Urbana-Champaign, Champaign, 1999. Available at http://www.lis.uiuc.edu/~twidale/pubs/dq.html

47. Viller, S., Bowers, J. and Rodden, T. Human factors in requirements engineering: A survey of human sciences literature relevant to the improvement of dependable systems development processes. *Interacting with Computers*, *11* (6). 665-698.

48. Want, R., Weiser, M. and Mynatt, B. Activating Everyday Objects. In *1998 DARPA/NIST Smart Spaces Workshop*, (Gaithersburg, Maryland, 1998), 7-140 to 143.

49. Wenger, E. *Communities of practice: learning, meaning, and identity*. Cambridge University Press, Cambridge, UK, 1998.