Complexities of Multi-organisational Error Management

John Dobson*, Simon Lock, David Martin
Department of Computing, InfoLab21,University of Lancaster, Lancaster LA1 4WA

Abstract: In this paper we shall look at some of the problems in designing an information and communication (ICT) system for an organisation located in a complex multi-organisational setting. We shall look in particular at the handling of errors both within the ICT itself and in the complex multi-organisational activities which the ICT is designed to support. We shall discuss the role of ethnography in uncovering some of the complexities and possible sources of error. We shall conclude by offering some advice to system designers which should prevent them from repeating mistakes which have been made all too often before.

Keywords: responsibility modelling, organisational boundaries

Introduction

Learning from system failures is a difficult process. Analysing major failures after the fact is a crucial resource for developing more dependable systems in the future, as examples such as London Ambulance, Ariane, Therac, Ladbroke Grove, Space Shuttle attest. However, retrospective analysis often reveals that the precise set of events that caused a failure consisted of an unlikely or highly unusual convergence of factors. This means that while the analysis of single system failures at a fine level of detail and particularity is crucial to exposing their root causes there may be problems with how much we can extrapolate to aid dependability in other systems. Furthermore, of greater concern, as Reason (1997) states, focusing dependability measures on eliminating the possibility of previous failures can actually make a system more susceptible to other, different, future failures. Reason's approach, and that of others, to address this problem, is to focus not so much the particular failures in themselves but instead to look at the organisational structures and procedures that give rise to them. This is predicated on the idea that problems in organisational structures and procedures can lead to systemic weaknesses that may lead to failures in a variety of different specific operational areas. Indeed, such a view is entirely consonant with the current legal position on catastrophic failures, as elucidated by Reason:

"If these were individual accidents, the discovery of unsafe acts immediately prior to the bad outcome would probably be the end of the story. Indeed, it is only within the last 20 years or so that the identification of proximal active failures would not have closed the book on the investigation of a major accident. Limiting responsibility to erring front-line individuals suited both the investigators and the individuals concerned – to say nothing of the lawyers who continue to have problems with establishing causal links between top-level decisions and specific events. Today, neither investigators nor responsible organizations are likely to end their search for the causes of organizational accidents with the mere identification of 'sharp-end' human failures. Such unsafe acts are now seen more as consequences than as principal causes." (Reason 1997, page 10)

Following in this tradition, in this paper we outline an approach that aims to identify systemic weaknesses in the structure and procedures of complex systems. Our analysis, although in some ways inspired by and resonant with some of Reason's work, takes a different approach to identifying weaknesses. We outline a position whereby ethnographic field studies of complex organisations can be employed to gather data from which, firstly, we can identify the procedures, the personnel and the technologies responsible for error handling. The second stage is then to model these structures of responsibility and communication. Through this process we can identify biases in the manner in which errors are handled, and ways in which the search for latent faults may be curtailed due to incoherence in responsibility structures. By incoherence we mean situations, for instance, where it is unclear whose responsibility a problem is, or where there is poor communication between parties jointly implicated in dealing with dependable operation, or where contractual relations between organisations do not stipulate a means of enforcement.

* Address for correspondence: J.E.Dobson, 31 Wentworth Park, Allendale Northumberland NE47 9DR, U.K. email <J.E.Dobson@ncl.ac.uk>

We have developed our approach through applying responsibility modelling (Dobson and Strens 1994) and the fault-error-failure model of dependability (Randall 2000) to details of the Ladbroke Grove Railway disaster as elucidated in the transcripts and other materials of the Public Inquiry. The Public Inquiry provides us with a level of access, albeit retrospectively, to the particulars of inter-organisational operations (as well as, clearly the crash details) that we would not normally be able to access. It is for this reason that we suggest that we would want to employ ethnographic studies – studies that can explicate operational procedures, practices and communications in detail – when collecting the materials to apply our approach to a novel setting. Plans are underway to do carry do this, but the main focus of this paper is about outlining the approach.

One clarification is necessary at the start. When we use the term "information and communication system" we are not assuming anything about the extent to which it has been computerised. Information systems are taken to include not only paper records but also individual and organisational memory. Similarly, communication systems are taken to include teleconferencing and face-to-face meetings. To stress this point, we have deliberately chosen to illustrate our points by reference to an example in which serious failings in the information and communication systems were uncovered, though no computer systems were implicated. More will be said about this later, when the example is introduced. However, our recommendations and conclusions are intended to be applied to systems built using information and communication technology; we are hoping to show how the kinds of problems that such systems have to deal with requires a certain reconceptualisation of information and approach to design when complex shared responsibilities are to be supported.

Many ICT systems are designed for a context which is restricted to the organisation that deploys them. This is often an oversimplification since organisations often do not work as a closed system with relationships confined to defined interfaces. Standard system design paradigms are not well adapted to designing for systems to be deployed in complex multi-organisational settings, often because the procurement process is specifically designed to exclude this. Procurement is thought of as a single contract between a single purchasing organisation and a single supplier (though the supplier may be a consortium of organisations). This model works well when the nature of the goods to be supplied is well-understood ("Please supply half a ton of broken biscuits"), but fails when the relationship between the parties is more complex than a consumer-supplier one, or when it involves something more complex than goods, or when recovery from failure is problematical and involves society as a whole, not just the interested parties.

To make this clear, here are three examples of organisational relationships that are well-understood and standard system design paradigms can be made to work quite well: support for consumer-supplier relationships; implementation of straightforward financial transactions; licence-handling applications. Here, by contrast, are three examples of more complex multi-organisational systems where standard system design paradigms have been found not to work too well: systems to support healthcare for the citizen; integrated transport systems; military command and control systems. These systems are all complex because they all include patterns of shared responsibilities which are implicit, negotiated and dynamic; and, as we shall see, it is often not until a failure occurs that the full complexity of these shared responsibilities is appreciated, and the simplified assumptions about them that are implicit in the information and communication systems that support the joint enterprise are exposed and break down.

This make such systems hard to design because more attention has to be paid to what happens when things go wrong. It is in the presence of failure that responsibilities are assumed, rearticulated and renegotiated; this requires flexibility of role definitions and organisational boundaries. Rigidity of boundaries and interface definitions so often serve to prevent recoverability. Many system design methods start from the assumption that the functionality of the system can be defined in terms of activities that are expressed in terms of their behaviour when accessed through defined interfaces. Although this is a simplified model of the way real things work in the real world, it works well enough as a representation when things are working correctly in a stable well-understood environment. But in the kinds of complex multi-organisational systems we are considering, when things do not work correctly, human ingenuity is often able to conceive some kind of workaround which may well require some extension or renegotiation of responsibilities, and this in turn may require some adaption of the information and communication system representation of the real world entities.

An illustrative example

One of the best simple examples of multi-organisational complexity is the relationship between Railtrack and the train operating companies as illustrated by the Ladbroke Grove disaster.[1]

On the fifth of October 1999, at Ladbroke Grove in London, a catastrophic crash occurred between two trains resulting in the deaths of both drivers and 29 passengers and the injury of approximately 414 persons. The trains involved in the crash were owned by two private companies - Thames Trains and Great Western Trains. The Thames Train was driving away from Paddington, while the Great Western was travelling towards Paddington. The Thames train passed a signal that was reading 'stop' (signal passed at danger (SPAD)), and indeed continued accelerating into the path of the Great Western. It is this action that has been described in the Inquiry into the tragedy as the *"immediate cause of the accident*". The Ladbroke Grove Rail Inquiry, which provides our material, took place in the months of May and December the following year. As an inquiry rather than a criminal case, the court proceedings had a wide ranging remit, provoking a detailed examination of organisational circumstances at the time, which is felicitous for our purposes. This remit was stated by Mr Hendy (acting solicitor on behalf of the families of victims):

"The evidence will show that a multitude of failings together brought about this crash. Probably every one of them was forseeable and avoidable. Our clients want each of them exposed and remedied for the future. In the course of the enquiry failings will be demonstrated which it will be argued were not causative of the collision. Our clients are naturally concerned to find out which factors were causative and which were not. But they are much more concerned that every factor exposed in this Inquiry which might lead to an accident in the future will be remedied than in any abtruse debate about whether particular factor was or was not a causative of this crash." (2,5,23)

Following the privatisation in 1996 British Rail was restructured into more than 100 separate businesses. What had previously been an integrated network bound together in a hierarchical bureaucratic structure with a few modern concessions to divisional decentralization now became an interconnected array of contracts linking companies accountable to their own shareholders and to regulatory bodies (Martin 2002). This radical re-organisation meant that new inter-operational relationships between the new companies with different responsibilities needed to be forged. Perhaps unsurprisingly, establishing effective new working relationships between separate companies, when previously such matters were handled internally within a single organisation, created particular problems (Martin, 2002). An important element in these difficulties can be seen to be a lack of clear definition of the nature of, and lines of control in the inter-organisational relationships and poor inter-organisational communication. In this paper we focus on the inter-operational relationships relating to dependability that The Inquiry identified as implicated in the Ladbroke Grove disaster.

The Scrutinised Parties: In The Inquiry the scrutiny most directly falls on two companies – Thames Trains, the owners of the train whose driver (Driver Hodder) committed the SPAD and Railtrack, the company responsible for the maintenance and running of the rail infrastructure, including signals, tracks and so forth. Thames Trains are implicated and scrutinised as the employers of Driver Hodder and as owners of the train. As employers, their procedures of selection, training and support are questioned. As owners of the train their level of technical safety measures is questioned. Railtrack is questioned on the design and maintenance of the infrastructure, particularly signal design, placing, maintenance and evaluation and the attendant procedures for dealing with these. Interestingly, both companies are also scrutinised on what may be thought of as organisational ethos — whether their priorities were for profit or safety, how were these attended to, was the balance right? Other companies such as Jarvis (the maintenance company) might also

[1] [For readers not familiar with the UK railway system it will help to know that at the time of the disaster, railway responsibility was divided between a single organisation (Railtrack) responsible for the entire infrastructure of track, signalling, bridges and property, and a number of train operating companies (of whom Thames Trains was one) responsible for operating and maintaining rolling stock and the conveyance of passengers or goods. A regulator was appointed to oversee that the whole system worked in the public interest, to award franchises to the operating companies, and to enforce public policy in the areas of competition and cooperation.]

be partially implicated as the ramp on the track set to trigger the advanced warning system (AWS) –an in cab alarm– if the driver committed a SPAD may have been faulty. However, we will focus on the main two parties as this serves as an apt illustration of our approach.[2]

As we shall show, the information and communication systems in these organisations, though partly manual, were deficient. There is no reason to believe that fully automated systems would have been any better, given the system design paradigms for computer-based systems prevalent at the time.

But before we can discuss the dependability implications, we need to establish a clear vocabulary. The next section introduces some basic terms which we shall use when our analysis resumes in the subsequent section.

Dependability Basics

In this section we introduce, with examples, some basic vocabulary for talking about dependability. The terms used are standard in the domain of dependability of computer-based systems, but we will explain them in the context of any socio-technical system, including those in which the technology is not computer-based (or can be so regarded: computers are in fact used in signalling systems, but in the Ladbroke Grove case this was completely irrelevant).

FAULT: a fault is something that is not as it should be; a state. It may be a state of a human or of a machine. It may be latent (not visible) and it may be benign (does not cause an error).
ERROR: an error is the manifestation of a fault, and is a behaviour of something. Often an error is a manifestation of an interaction between two or more faults.
CONSEQUENCE: a consequence is the observable effect or outcome of an error.
FAILURE: a failure has occurred when a undesirable consequence is experienced. It is a judgement about erroneous behaviour, based on its consequences.

We shall endeavour to be quite consistent in our usage, which is based on strong typing: a fault is a state, an error is a behaviour, a consequence is a result (a causal concept), a failure is a judgement.

We shall also introduce some terms associated with the achievement of dependability. Faults can be avoided during the creation or implementation of system components. Faults can be removed from components after they have been created or implemented. Faults can also be tolerated, which means that if an error occurs and is detected as such, some recovery or workaround is initiated which prevents the error from causing a consequence judged to be a failure. This implies the need for monitoring and exception handling. The risk of failure can also be accepted, with the cost of failure (if it occurs) being met, for example through insurance or compensation or writing off. Acceptance of risk can be transferred to users or operators through the use of disclaimers and warning signs.

Example: There are many possible accounts of an incident which leads to an adjudged failure, taken from different viewpoints. Indeed, though not here, some accounts may lead to the view that a consequence was not, in fact, a failure — since a different judge is making the judgement. One possible account is that which places the responsibility with the train operator (there are others, equally valid):

FAULT a poorly trained driver
"A Mr Holmes of the Railway Inspectorate told a Mr Franks of Thames Trains that he was "very concerned about driver training"
ERROR a signal passed at danger (called a SPAD)
FAILURE a crash

Another possible account is that which places the responsibility with the infrastructure provider:

FAULT a badly designed and/or positioned operational signal

[2] Our approach to modelling is based on the premise that it is important to extend the analysis of error handling and the responsibility modelling further and wider in an attempt to discover latent conditions –systemic problems– (Reason, 1997) implicated in the failure but at first not apparent. However, the judgement of how extensive any modelling should be – i.e. across how many processes and organisations – will depend on the purposes of the project.

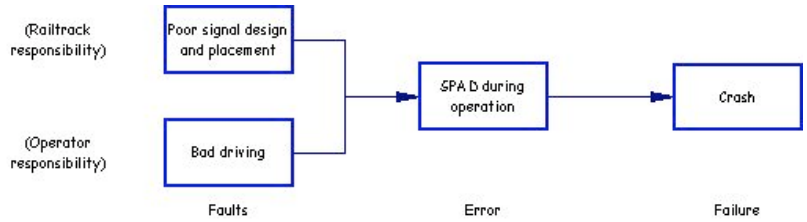
"there would appear to be three principal factors – the siting of signals in terms of design, local orientation and conspicuity – visibility – and human factors which may have contributed to the error."

FAULT *Inadequate monitoring and countermeasure guidelines and practice*

ERROR *SN109 not identified as dangerous due to poor monitoring and countermeasure processes*

CONSEQUENCE *the continued use in operation of SN109*

We can summarise the story so far in the following picture:



We now briefly look at the mechanisms in place that were intended to achieve dependability of the system.

Operational faults

Removal	It was assumed that (re)training and information would remove driver errors due to faults in insufficient skill and knowledge *"But…. Mr Adams, who supervised Driver Hodder's practical training……was not aware that SN109 was a multi-SPAD signal."*
Tolerance	It was assumed that an automatic warning system in the driver's cab and a 700-yard run-on (between the signal and the points it controlled) would be sufficient to allow error recovery and 700 yard run on would allow error recovery to avoid failure

Signal design and Placement Faults

Removal	Procedures were in place to identify and rectify problematic signals but a solution had not been found/agreed upon for SN109 *"What is unquestionably the case is that the bodies that I have identified generated a considerable quantity of paper. What is less clear is how effective they were at identifying problems and rectifying them."*
Tolerance	In addition to the assumed tolerance of driver error, procedures were in place in the signal control room to detect SPADs and to take appropriate action in signals on the other line(s)

So although fault removal and tolerance procedures were in place, they were not perfect. What we see here is another example of "This system is safe provided that system is reliable". It is this chain of faults errors and failures that crosses both system boundaries and organisational boundaries that leads to its complexity. Thus identifying and remedying faults involves tracing back from manifest failures to identify causally implicated faults. Other features of the model are that is draws attention to the fact that fault-error-failure chains may be nested within one another recursively, and may also exist alongside one another, interacting on occasion to produce failures. For example, in a technical system a fault may be due to a failure in the development process, traceable back to a faulty design that leads to errors in programming. Or the interaction between faults in technical and human systems may be more likely to compound each other and cause failures. Although Randall focuses on technical systems, the fact that he mentions the design process, and also how failures in technical systems may set off fault-error-failure chains in the social systems of people using technologies provides us with ready justification to extend his model to look at inter-organisational socio-technical systems.[3]

[3] Subsequently Randall outlines an approach to fault tolerance whereby latent faults may be remedied through technical instantiations of redundancy and diversity. We believe that these ideas may also be important for dependable fault tolerance in socio-technical systems and will discuss them in our conclusions.

A single failure may be the consequence of multiple faults, all acting together. The removal or tolerance (recovery) from a single fault may prevent a subsequent failure occurring. The danger is (especially with multi-organisational systems) that the faults which are not removed of protected against will remain latent and may later become reactivated by changing conditions or the injection of further faults. For example, if the infrastructure provider (namely Railtrack) did all that they could to remove faults from the system, this would at best improve the positioning of the signal, removing only one fault from the system. The adequacy of driver training would not be affected; indeed, the deficiencies in training might go unnoticed as the improved signal positioning would be likely to reduce or prevent failures.

We would also lie to draw attention to the fact that any of the above identified faults can be seen to arise from systemic failures in the systems sitting behind the operational system. For example, systematic failures in the design and placement of signals may arise from a faulty design premises leading to errors in their design process and protocols for placement. A similar situation may arise out of systematic problems in driver training and selection or processes through which monitoring and error handling guidelines, processes and protocols are defined and implemented.

Indeed The Inquiry judged that these were all the case, and that senior management in Railtrack and Thames trains were therefore implicated in the operational failures down the line. Our position, and this is backed up by the inquiry, is that the investigations following SPADs were limited, primarily considering them to be driver error, as most drivers admitted some liability following a SPAD. Efforts to look for other antecedent causes were limited – the positioning and design of SN109 was considered but difficulties in advancing a solution, agreeing upon it and implementing it meant the problem was not solved in a timely manner. The fact that there might have been a systemic problem in the systems and procedures to monitor SPADs and deal with problem signals with did not seem to have been considered. On committing SPADs drivers were retrained, and there was some information available on problem signals, but the basic procedures for driver selection and training were never reviewed for their own faults.

It should now be possible to see that by applying the fault-error-failure model of dependability we are necessarily inclined to scrutinise whether existing error handling procedures dangerously curtail investigations into non catastrophic failures (as SPADs were treated at the time). Our position being that efforts in this situation were not made to trace back to more distal antecedent causes, and that this is what was required. And that this should be the approach when investigating error handling in other situations. Our analysis also uncovers another problem – that potential faults were not adequately dealt in unison and cooperatively by the two organisations involved. Thames trains were meant to be dealing with the potential causes of driver error, Railtrack with signal problems. Uncertainty reigned as to where the problem actually lay when a coordinated holistic approach would have treated the problem as due to both and potentially the antecedent causes of both.

This brings us to the issue which is at the heart of this paper: the complexity arising from multiple faults situated in different organisations. Examples of this complexity are:

* With different organisations, how do different possible faults interact? Whose responsibility is it to work this out? Who is responsible for the interaction?
* What is the model of the relationship between companies? What is the nature of the contract between them?
* Is the peer relationship of very loose cooperation adequate for creating a safety structure?
* How do faults, error and failure in the system that creates a given system undermine the effectiveness of fault avoidance strategies? In a similar way, how are fault-error-failure chains associated with the other failure management schemes (fault removal, fault tolerance, failure acceptance)?

---

[3] Subsequently Randall outlines an approach to fault tolerance whereby latent faults may be remedied through technical instantiations of redundancy and diversity. We believe that these ideas may also be important for dependable fault tolerance in socio-technical systems and will discuss them in our conclusions.

Responsibilities for Handling Errors

In this section, we shall expand a theme mentioned in the introduction: responsibilities for handling errors. We maintain that in complex multi-organisational settings, failures often occur because mechanisms for sharing responsibilities are inadequate or absent; and this is particularly true for responsibilities for preventing or managing failure. In order to further develop this analysis, in the following section, we apply responsibility modelling (Dobson and Strens 1994) to the inter-organisational relationships that existed between Railtrack and Thames trains at the time.

Causal and consequential responsibility: There are many meanings of the word 'responsibility', which we will not discuss here. (Good books to read on the nature and importance of responsibility are Lucas (1995) and Jonas (1984), respectively.) However, for our present purposes it is useful to distinguish between *causal responsibility*, when an agent has an obligation to make something happen or prevent it from happening or to maintain a state, from *consequential responsibility*, when an agent is answerable when something happens or does not happen or a state is not maintained. These different responsibilities do not always rest on the same agent (the doctrine of 'ministerial responsibility') and consequential responsibility may be held to rest with an organisation as a whole whereas causal responsibility most usually can be traced to an individual or the fact that no particular individual at the time held the responsibility. causal responsibility may sometimes be delegated, though some responsibility remains with the delegating agent (i.e. the responsibility for having chosen to delegate), whereas consequential responsibility is not normally capable of delegation, though it may sometimes be transferred. We shall refer to these different responsibilities in our discussion of Ladbroke Grove, and deal with the complexities they pose for system design in a later section.

Lifecycle and responsibilities: In preparation for mapping out the responsibilities implicated in a failure, it is useful to start by looking at the major life-cycle phases of an operational system as a way of distinguishing different responsibilities. There are four major phases (defined by processes) in the life cycle of an operational system: procurement; operation; maintenance; decommissioning (in the case of Ladbroke Grove, decommissioning was not an issue). It is easier to deal with particular faults in particular ways at particular points in the life-cycle:

Procurement includes making assessments of the risks and consequences of operational failures.

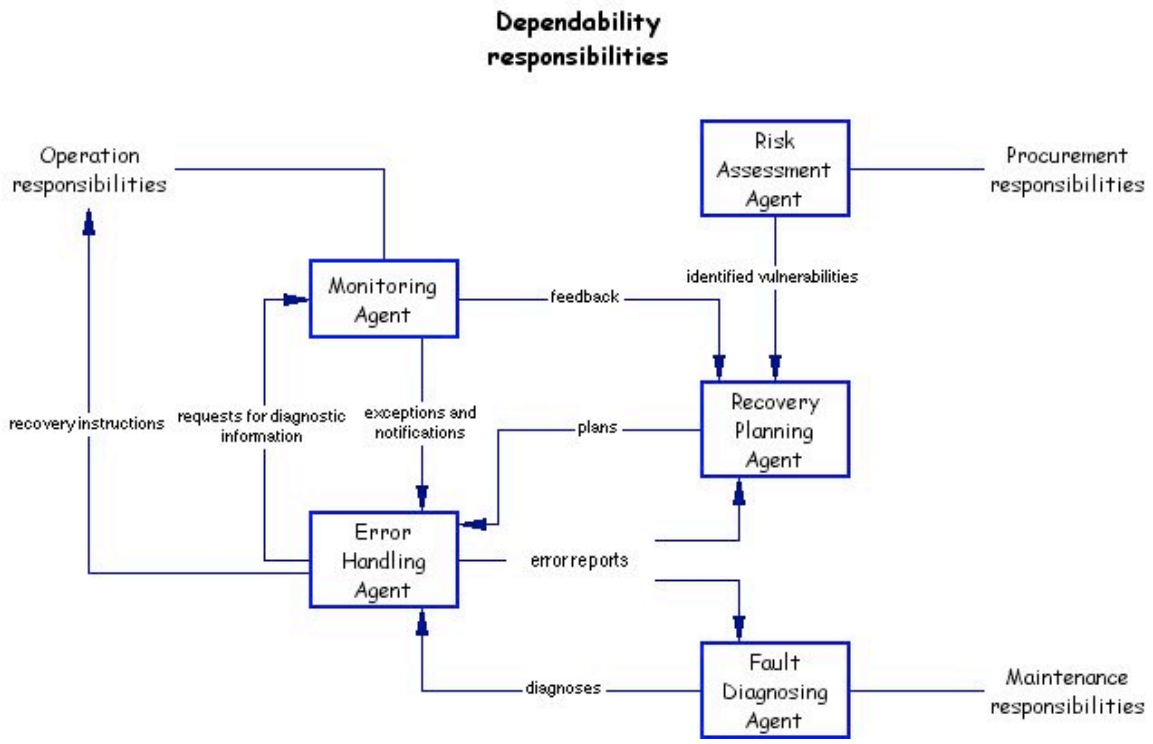
Operation includes monitoring errors and following plans for recovering from the errors so as to prevent them from giving rise to failures.

Maintenance includes taking retrospective action to prevent subsequent occurrences of F-E-F chains.

Decommissioning includes ensuring that documentation concerning the (in)accuracy of the failure mode assumptions and (un)successful ways discovered of managing failures is preserved for posterity.

There are a lot of responsibilities implicated here. Where they can all be contained in a single organisation, as was the case with the former nationalised unitary organisation British Rail, then consequential responsibility for a failure clearly rests with that organisation. But where, as in Ladbroke Grove, the responsibilities are distributed over a number of organisations, it may well be the case that attribution of consequential responsibility is not at all clear cut and a detailed judicial investigation is needed. Such an investigation should examine not only the causal errors, including errors in the preventin and tolerance processes and mechanisms, but the risk assessment, monitoring and recovery procedures; and, more importantly, it should look at where the various responsibilities for those processes and procedures lay and the extent to which they were satisfactorily discharged.

The following diagram shows the main responsibilities for managing errors and how those responsibilities relate to those for procurement, operation, and maintenance:



The use of the word 'agent' here indicates a responsibility for doing something or seeing that it gets done – the actual execution could be performed by a machine or delegated to humans. An agent is always a person or group of people sharing a responsibility.

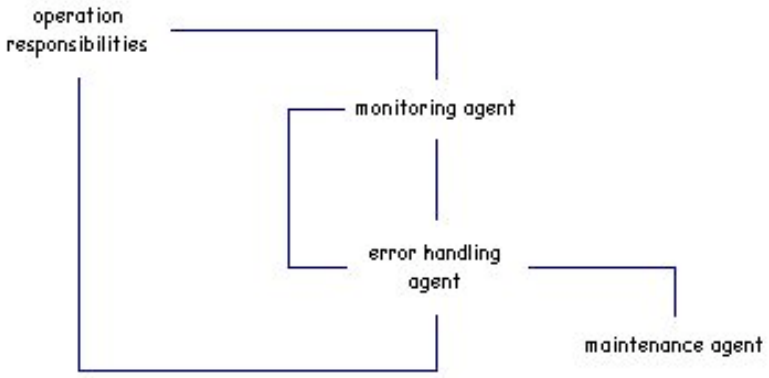
The lines in the diagram represent not just information flows but conversations. A conversation is a possibly extended series of exchanges, distributed in time and space, between two agents. The information exchanged can also be seen as a partial state of that conversation as it exists at any instant. More details about the modelling here presented will appear in a forthcoming book (Clarke and Hardstone 2005).

The picture is intended to be normative. Its use is in performing a comparison with a description of the responsibilities as they are articulated in the actual setting, in order to identify such things as missing or ill-defined responsibilities, or shared responsibilities that cross inter- or intra-organisational boundaries, as it is these that so often give rise to failures, and in particular in failures in failure prevention or management. This comparison can be used, as in the soft systems methodology (see Checkland (1981) and Checkland and Scholes (1990) for the theory and practice of soft systems methodology), as a way of determining expectations on a new information system.

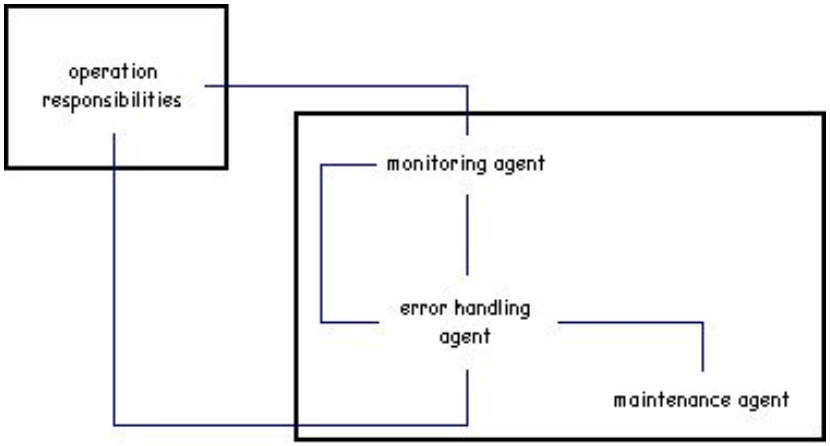
The positioning in this model of (intra- and inter-) organisational boundaries is key to effective error recovery. This will be discussed in the next section.

Organisational Boundaries

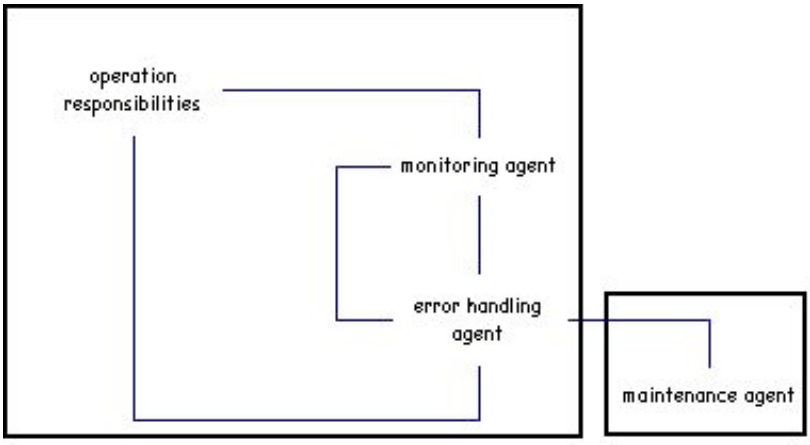
In order to discuss the problems arising when responsibilities cross organisational boundaries, we start by taking a slight simplification of the previous figure.



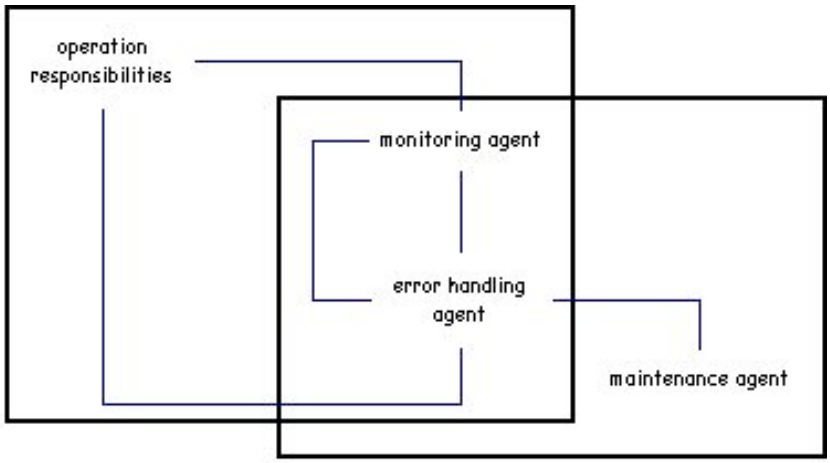
If maintenance responsibilities are in a different enterprise from the operation responsibilities, where exactly does the boundary lie? It could, for example, be like this:



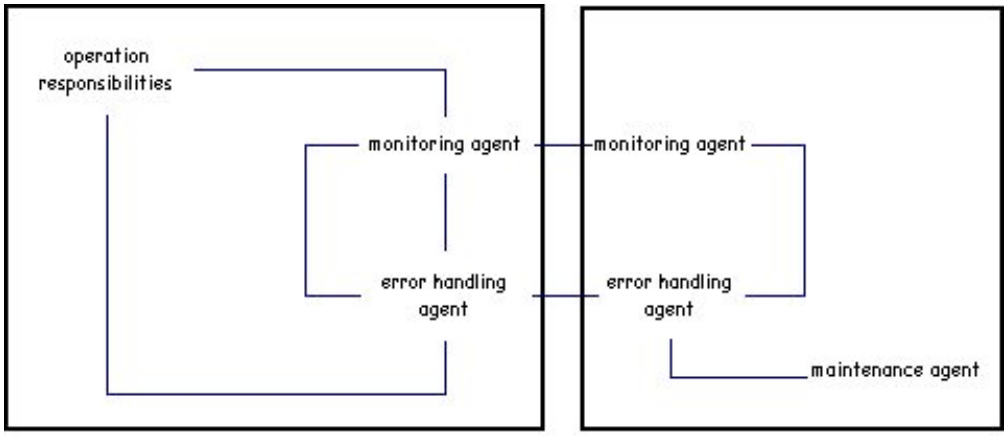
Here, system maintenance is carried out either by a separate organisation or by a separate division within the operating enterprise. As part of the maintenance, all the monitoring responsibilities can be transferred, but the operator is then dependent on another organisation for critical management information; there are a number of possible organisational failures associated with such a critical dependence.
An alternative that is theoretically possible but in practice would be defective, is shown below:



but in practice, maintenance will include at least some monitoring and therefore some error handling:



so that monitoring and error handling responsibilities are shared between the operational organisation and the maintenance organisation, Such shared responsibilities require good communications channels and some way of resolving conflicts in priorities, because this model is equivalent to the following:



The problems here are clear. Inter-organisational conversations are required to coordinate shared responsibilities; but the media and channels required for such co-ordination may be unclear and the supposedly communicating processes may be mutually opaque as indeed they were at Thames Trains and Railtrack, as the Ladbroke Grove enquiry shows.

Boundary Objects

Where, as here, communication occurs between to parties as a result of some extended relationship between them or because they are engaged in some common enterprise, it is necessary to examine not only the medium and channel of communication, but also the shared objects –sometimes called boundary objects– to which the communications refer. There has been much discussion on the concept and nature of boundary objects, and indeed whether they can easily be determined. In simple cases, however, the idea is a useful one. A boundary object is one which is visible on both sides of an organisational boundary, but which has different connotations on each side. For example, to Railtrack a train is something which uses their infrastructure; to Thames Trains, a train is something which delivers their transport services to their customers. So information about boundary objects is generated in two distinct contexts. Normally, such information is interpreted in the context in which it is generated, though parts of the context may be shared (e.g. the whereabouts on the line of the train).

In the presence of failure, however, when shared responsibilities are actually called upon to be exercised, information generated on one side (e.g. about the training of drivers) has to be interpreted on the other (e.g. was their driver appropriately trained *from our point of view?*)

In addition, two other things tend to happen:
i) what constitutes the boundary object is re-articulated (e.g. trains are now seen to have drivers)
ii) things previously not seen as boundary objects now take on that significance (e.g. signals have now to be treated as boundary objects; as indeed have drivers because Railtrack now realises that it has an interest in the driver's training and experience).

There are three distinct, but related, information management problems that now arise:
1) What one party judges to be a failure of, or implicating, a boundary object might not be so judged by the other party (e.g. the fact that drivers had difficulty in reading a particular signal was initially treated by Railtrack as a form of driver failure, not signal failure).

This is a point which is not always appreciated by those who consider a failure a state or behaviour (i.e. something which all competent observers can agree upon), because although it is a mismatch between what actually occurred and what a specification says should occur, there might be more than one valid specification. Clearly, a train crash is a failure, as is a SPAD: but wherein lies the error(s)? And of what is the failure that gave rise to the error(s)? This is the problem of socially determined failures, i.e. consequences whose subsequent characterisation as failures is a process of social agreement.

2) Information on one side of the boundary – including its context of interpretation and generation – might not be visible on the other side.

This undoubtedly occurred at Ladbroke Grove. The report comments unfavourably again and again on the way that information passed across the boundary but was not acted on for reasons that were obscure.

3) A shared approach to recoverability or repair might well be hampered by the invisibility of relevant information or its processing.

These problems are deep-rooted and give rise to a number of issues in the design processes for an information management system, to which we now turn our attention.

Some Implications for Design

Agency: The binding between individuals and responsibilities is a complex many-to-many relationship. We structure this using two distinct concepts. We have introduced the concept of *role* to classify the relationships held by an individual: an individual can hold many roles simultaneously, and a role may imply many related responsibilities. The concept of role is related to the structure of an organisation and is defined in organisational terms. *Agency*, on the other hand, is abstracted away from any actual organisational structure and is simply a collection of related responsibilities; how it maps onto work roles is a matter of organisational design and will change as the organisation changes. In particular, one particular agency may span inter-organisational boundaries, such as the consequential responsibility for a collision.
The concept of agency allows a conceptual separation to occur as organisations change. For example, small organisations often combine the sales agency and the marketing agency into the same role; as the organisation grows, this is often a desirable separation to make. Since agency is a more stable concept than role, an information system based on agency rather than role is more likely to be capable of change as the organisational structure changes.

Conversations: Dealing with multi-organisational failure and its consequences requires communication and cooperation. This implies that information, as well as being about the state of a boundary object, is also the (partial) state of a conversation between the communicating and cooperating parties. This means that an information system is sometimes better reconceptualised as a communication system, and this in turn requires a reconceptualisation of communication and conversation), one that provides a basis for undestanding failure modes.

Conversations and the relationships that they define, sometimes fail. The purpose of a theory of conversations is to explain the failures associated with the intentions of the participants. It is clear that the bringing together of obligations and responsibilities can create conflicts of interest as well as synergies. It can also create overloads and imbalances which could lead to failure in operation. In addition to failures of organisational policy and design, we have operational failures due to a lack of correspondence between the expectations of the participants. We have developed a theory of the attributes of roles and conversations that provide a basis for analysing such situations. We have also developed an analysis of failures to perform the intended role by failing to generate correct information, by misinterpreting presentations or by proffering incorrect or inappropriate resources. These failures would be accounted for in a theory of instruments. Finally, failures in reading, writing or transporting data are the province of a theory of communication.

The need to record: Responsibility modelling raises three important information questions: What do I need to know? What do I need to do? What do I need to record to show what I have done? It is this last that becomes of importance when the possibility of failure raises questions of answerability. Recording can be seen as an anticipated conversation between a responsibility holder and an investigator. For example, one possible organisational policy is that consequential responsibility following a technical malfunction rests with the person who chose to deploy the malfunctioning device. This answerability could be mitigated by providing recorded evidence about the soundness of the choice process.

Boundaries and boundary objects: Problems with systems are particularly likely to arise if the systems are intended to cut across organisational or professional boundaries. One reason why such problems arise is that the responsibilities on each side of the boundary are either undefined or are incompatible.
One design implication is that need for explicit representation of the nature of relationships across the boundary, identifying boundary objects, conversations and communication, and shared and non-shared responsibilities.
Because boundary objects have differing interpretations on the two sides of the boundary, there is often a need for two distinct representations of the object. For example, to the train operator, a train is a unit of schedulable resource, which requires other schedulable resources, such as a trained driver. Essentially the need is for a static representation. But for the infrastructure provider, a train is a dynamic unity defined by a head code and a moving point on the track; the need is for a dynamic representation. Tying together the two different representations is not, to be sure, an insuperable problem, but it does present a certain complexity in system design.

Monitoring: Not everything need be monitored. Obviously if failure is likely to be catastrophic, fault tolerance and recoverability measures are important. But if the consequences of failure are likely to be merely inefficiencies, resources for planning for and implementing monitoring are best spent where the structures of the system or its associated responsibilities cross organisational boundaries, since it is there that disputes are both more likely to arise and difficult and costly to resolve.

Audit trail: It is unusual for information systems to have the capability to record things that happen in the social domain, such as delegation. It is in the interest of agents who hold consequential responsibility that the audit trail is correct, reconstrucable and complete. For example, one possible organisational rule is that consequential responsibility following a technical malfunction rests with the agent who chose to deploy the malfunctioning device. This answerability could perhaps be mitigated by providing evidence about the soundness of the choice process, including those aspects of it that took place in he social domain.

Design for recoverability: There are two main classes of strategy for recoverability: backward recovery is the restoration of a previous state known to be safe, usually saved as a checkpoint or archive; forward recovery is the forcing into a known safe state. Backward recovery is not always possible for systems that are closely coupled to the world, since although the system can be rewound, it is not always possible to rewind the world. We shall therefore concentrate on some design considerations for forward recovery.
There are many risk assessment methods available for analysing possible vulnerabilities (weaknesses in the system) hazards (potential threats in the environment) and risks (decisions concerning what to do, or what not to do, about the vulnerabilities and hazards).

However, one important strategy for recovery after a failure is diversity: trying not to put all you eggs in one basket is as important during recovery as it is during normal operation. Remember that independent systems of the same functionality may well not fail independently (e.g. having a second driver in the cab may not help if both have been on the same defective training course).

Summary and Conclusions

Responsibility modelling: Focussing on responsibility is to make a distinction between who is responsible for performing a role and who (or what) actually executes a role. We advocate that looking at responsibilities is a better guide for designing information systems than looking at executions, since it allows analysis of problems that can potentially arise when responsibility has not been clearly allocated, those responsible do not or can not actually perform the role, responsibility cannot be enforced because of lack of corresponding authority, communication between jointly responsible actors is difficult or impossible, and many other causes of organisational failure.

We now look at the acquisition of information about responsibilities. Clearly one way of finding out about responsibilities is direct enquiry: asking informants (and their managers), looking at their job descriptions and contracts and so on. But the direct approach, although necessary, is also limited. People's interpretation of their responsibilities are often nuanced, and this nuancing is often better determined as a result of observation and subsequent elaboration, since direct questions are usually answered directly.

It is one of the roles of ethnography to observe the manifestation of interpretation of responsibility. It can do this by explicating social aspects of work and considering the relationship between local practice and formal procedure – how procedures are enacted and how practice is related to or explained as or accounted for in terms of formal process. It can probe into aspects of safety culture as these are enacted in the organisation. Because ethnographic interpretation considers systems in a broad socio-technical sense, it is particularly useful for analyses of 'systems' where computers are minor players. Ethnography is also useful in identifying boundary objects and the ways their interpretations differ on each side.

Ethnography can also be useful in failure mode analysis. One particular use is in situations where response to potential or actual failure is a preventative or recoverable action, ethnography provides a description of what actions actually occurred — as opposed to what actions were supposed to occur. (It is hardly necessary to stress how often these differ.) It can show how fault-error-failure chains are investigated and examine the nature of interactions across organisational boundaries – how processes are brought into alignment, and who or what does the job of translation. (This is particularly important for recoverability.)

Uses of responsibility modelling: So far, three possible uses for models of responsibility seem to be emerging:

1. During planning/procurement/requirements when there is a need to clarify the responsibilities of the different actors in the system, especially where multiple organisations are involved.
2. During an enquiry, when there is a need to find out who takes the blame and (perhaps) who should have done something.
3. During system operation, when a problem arises and there is a need to find out who needs to know and what they need to know.

Organisational complexity: Organisational complexity requires an ICT system design method which recognises that multi-organisational systems need to extend current methods in the way they deal with failure. Three examples seem particularly important.

1. Procurement processes which are based on the single organisation assumption may not work too well.
2. Information is often best regarded as a partial state of a conversation and understanding the nature of the conversation is needed to construct the multiple contexts of generation and interpretation.
3. Failures which can be traced back to errors in the sharing of responsibility are going to occur and the recovery procedures also have to be designed for a multi-organisational context of use. Where consequential responsibility is unclear, the social and legal processes require more information than just that immediately prior to the triggering event. the nature of the contract between the parties may have implications for existing (or non-existing) systems.

Acknowledgements

We wish to thank all those who have participated in discussions with us at Lancaster (particularly Mark Rouncefield Guy Dewsbury and Ian Sommerville) and Newcastle (particularly Mike Martin and Ros Strens). This work is currently supported by the EPSRC through the Interdisciplinary Research Collaboration in Dependability (DIRC).

References and Further Reading

Box, S (1983) Power, Crime and Mystification. London. Tavistock.

Checkland, P. (1981). *Systems Thinking, Systems Practice*, Chichester, John Wiley.

Checkland, P. and J. Scholes (1990). *Soft Systems Methodology in Action*, Chichester, John Wiley.

Clarke, K.M. and Hardstone, G.(2005), *Trust in Technology*, Kluwer, (in press).

Clarke, K., Martin, D., Rouncefield, M., Sommerville, S. (2002). "Going Through The Usual Rituals": coping with system failure in a hospital setting.

J.E. Dobson and R. Strens (1994) "Responsibility Modelling as a Technique for Requirements Definition", *Intelligent Systems Engineering*, vol. 3, no. 1, pp. 20-26, 1994.

Garfinkel, H., (1967) Studies in Ethnomethodology. Englewood Cliffs, N.J.: Prentice-Hall

Harper, R. and Hughes, J. (1992). '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 G. Button (ed), Technology in Working Order, Routledge.

Hart, H. L. A., and Honore, T., (1985) Causation In The Law (2nd Edition). The Clarendon Press: Oxford.

Heath, C. and Luff, P. (1991). 'Collaborative activity and technological design: Task coordination in London Underground Control Rooms'. Proceedings of ECSCW '91. Dordecht: Kluwer, pp. 65-80.

Hughes, J., Randall, D. and Shapiro, D. (1992). Faltering From Ethnography to Design. In Proceedings of CSCW '92.

Jonas, H. (1984) *The Imperative of Responsibility*, Chicago, University of Chicago Press

Ladbroke Grove Rail Inquiry http://www.lgri.org.uk and
http://www.hse.gov.uk/railways/ladbrokegrove.htm

Law, J (2000) 'Ladbroke Grove, or How To Think about Failing Systems' - published by the Centre for Science Studies and the Department of Sociology, Lancaster University at
http://www.comp.lancs.ac.uk/sociology/soc055jl.html

Lucas, J. R. (1995). *Responsibility*, Oxford, Clarendon Press.

Luff, P., Hindmarsh, J., and Heath, C. (Eds.) (2000). Workplace Studies: Recovering work practice and informing system design. Cambridge: CUP.

Martin, B, http://en.wikipedia.org/wiki/British_Rail Brendan Martin - british rail privatisation – what went wrong?

Martin, D., Bowers, J. and Wastell, D. (1997). The Interactional Affordances of Technology: An Ethnography of Human-Computer Interaction in an Ambulance Control Centre. In Proceedings of HCI ’97.

Perrow, C (1984) Normal Accidents. Basic Books. New York .

Randell, B. (2000). Turing Memorial Lecture: Facing Up To Faults. The Computer Journal, Vol. 43, No. 2, 2000

Reason, J. (1997) Managing the Risks of Organizational Accidents, Hampshire, England: Ashgate Publishing Limited; 1997

Sagan, S.D. (1993) The Limits of Safety: Organizations, Accidents, and Nuclear Weapons. Princeton University Press, Princeton, NJ.