N72-25964

DATA REQUIREMENTS ANALYSIS IN SUPPORT OF SYSTEM SAFETY

by

Mr. Irving Pinkel

Chief
Information Services
and
Publications Branch

Aerospace Safety Research
and
Data Institute
NASA-Lewis Research Center

Presented at the

NASA Government-Industry
System Safety Conference

May 26-28, 1971

Those friends of George Mandel who are wondering why it is that I am here in his place, I am happy to report that he is recovering nicely from a heart attack.

On the matter of the Aerospace Safety Research and Data Institute, about three or four years ago, after the Apollo fire, NASA realized that its safety organization could use a center where safety information accumulates and is validated and interpreted for use by the Aerospace Industry. Our group was set up in the Cleveland laboratory to serve all of NASA and the Aerospace Industry. Three years ago I was a lone member of this group and I spoke to this conference about our hopes. Now I am here to report how we have proceeded, what our points of view are, and where we stand at this time.

Let's review for a minute and take a look at our objectives. First, to support NASA and its contractors with technical information and consulting or safety problems. To identify areas where safety problems exist or where voids in technology exist, and to initiate research programs both in-house and under contract to fill these voids; to prepare state-of-the-art summaries and other publications of use in our area. The key to all this is to establish and operate a safety data bank.

It is my purpose today to go through this quickly to give you an idea of our thinking and where we stand. I might add, as an overall remark, the emphasis we have given in our efforts is to keep the user of the information in mind. That user isn't necessarily a safety specialist as you are, but can be any one of the engineers in the total system of engineering support. There are decisions being made at all levels. Many of our users are competent engineers who are being called on to make decisions involving technical information for which they have poor background.

In order to maintain contact with the user population so that we do a useful job, we stay in detailed contact with the entire industry and all institutional centers of NASA where problems are apt to arise. We also have membership on a host of committees. Obviously the space shuttle is prime to NASA's interest at this time and I might add that in setting up this data bank we try our best to do the work in those areas of immediate interest to NASA and then broaden our interest as time allows. Space Shuttle is being controlled at this stage by a variety of committees within NASA and we have panel membership on each of these committees. We worry about cryogenics and low temperatures systems because we deal a great deal in propellants which are liquified gases, and we have membership in the Compressed Gas Association where much of this work is done. I won't detail all of these things but point out that in addition to all else we deal with assorted NASA committees dealing with space-borne radioactive materials. If you are wondering how it is that NASA deals in radioactive materials for space, I will remind you that the largest space station which will orbit the earth will carry electric producing systems which will not use the sun as a source of heat but either a nuclear reactor or radioisotopes. This is a real concern to us at this time. The final committee we serve on is NASA's Spacecraft Fire Hazard Steering Committee which I chair. This grew out of NASA's concern for fire problems on spacecraft, particularly manned spacecraft.

The question is, What is Safety Information? We had to ask ourselves, we are going to collect information, but what? What is it? Is it that body of information that has a safety label attached to it in someway? Well yes, it is that. But is it something else as well? Here is what we feel constitutes the boundaries of safety information and I am sure this is an inadequate detail of these boundaries. First, safety information is a body of technical matter drawn together from various disciplines in support of a safety problem. This information is often indistinguishable from engineering, scientific or medical information. In a sense, what we are saying is this, that safety information can be drawn from any part of the technical and scientific literature and we have to be prepared to do just that. Safety information is also information on hazard management techniques, and where equipment is involved, the associated equipment. It deals in failure advisories, accidents, reports, and then the legal aspects of safety, codes and standards.

Now, where we are dealing with a user-oriented system, the user generally comes to a safety problem with certain categories of questions in his mind. He would like for example to recognize when hazards exist, and understand how he can detect the build-up of a hazardous condition. And so we like to organize our information that way. Or he would like to understand how to reduce the probability of a failure or an accident. So we organize our information this way as well. He would like to be able to assess the consequences of a failure. Oddly enough, when we look at the literature for assessing the consequences of a failure we don't go to the safety literature, we go to the anti-safety literature. We look to the demolition expert and say, "what do you know about what would happen if we had an explosion". He would like to be able to reduce the consequences of a failure and he would like to have the information so structured that when he comes with this question in mind he can find that kind of information.

Then there are certain scientific and engineering fundamentals he has to have in order to apply what information exists. We feel that here is a kep weakness in the communication of safety literature, information from the literature, and that is the interpretation of what the literature tells you. We feel that in many areas, we, the Aerospace Safety Research and Data Institute, shall have to prepare documents which show how the existing information in the literature is interpreted in terms of real problems. We haven't begun this process yet except in a very limited way. It is a difficult thing to do, but I think it is a vital step. And we also, since the legal aspects of safety are so important, have to make our engineer who is dabbling in a safety problem aware that there are certain legal aspects to the safety problem.

When we took a look at the existing information in safety and decided to create a safety data bank, we were first faced with what shall go in the bank? We are proposing to have a largely computerized bank and the first thing that hit me forcibly in this whole business was the fact that if you use a computer as a bank, as a place in which to store information, you discover how enormously, enormously costly it is to do a proper job of putting information into a computer. We said we have to be careful what goes in, not only from the standpoint of cost, but from the standpoint of credibility. Can the people get information out of our system and depend on it? They are surely going to use this as an authority for the actions they take and if we give them the wrong information or poor information, it is our responsibility. Also, we looked at the quality of safety information. Most of you are old pros at this and I think you'll want to disagree with what I am going to say next.

In the safety information that we reviewed, we often found that important portions of the safety information are misapplied laboratory data. Data that was gathered not with a safety problem in mind but simply a study of a discipline, and somebody is using that information improperly in a safety document. Safety reports often deal in opinions masquerading as fact and this is all too often the case. I think many of you understand this. A large body of literature exists in some fields and little or none in others, and sharply focused information is difficult to find for both reasons. There are times when you query an information system about a certain aspect of a safety problem, you get snowed with 2,000 documents. That is as good as giving you nothing unless you have enough discretion in the field and are inquiring enough to pick that which is useful from that which isn't.

Much of the literature contains incremental contributions and a large mass of reports must be reviewed for answers to the safety questions. This tells us that somewhere in our system we have to boil down the information into review and summary reports and let that be the input to our system and cut out the chaff of a large number of incremental reports. And too, a point I alluded on before, information is couched in scientific terms which are unfamiliar to engineers. In other words, the information isn't user-oriented. If you want to touch on this at all, give an engineer a man-machine problem, the business of integrating man into a machine system, and let him look at the data the psychologists put out and try to make some sense of it for himself. I'm not saying that psychologist's data is no good, but the psychologist's data is so couched in jargon that the engineer is hopefully confused.

The present retrieval systems often lose the relevant information and cite many irrelevant references. When this happens, obviously there is a degradation in the service being provided.

Here is what we said the components of a safety data bank system ought to be.

First, we should use a computer, should be document references. These should have an appropriate abstract so that the person looking at a document reference doesn't have to go by the title. Authors of reports are notoriously poor in titling their own reports so we prefer to have an abstract which helps a little. In the work which we are going to be doing, which we ask people to review literature in specialized fields, we ask the reviewer who is an expert in the field to write his own abstract in addition to the author's abstract, if he thinks the author's abstract is misleading. Computer information should have references to other repositories that specialize in information, and I want to bring up the point that we don't think we are the only safety data bank in existence. We know there are many. We hope to be complementary with them; not to overlap them, and in no case to totally absorb them unless it's worthwhile to do that. We do have to know where the other information resides and to have the computer point it out as an answer to a query on information requests. It has to be able to store systematic accumulations of safety data and what I mean by systematic accumulations is this. Much of the information that a safety engineer, or person involved with safety problems needs to use, have never been published. It has been garnered from research, completed in private places and these are available to us as curve and graphs etc. plots, formulas -- we have to be able to include that in our system so these come out. We can't rely entirely on documents. We then need a list of specialists in safety and safety-related fields and this goes back to our role of consulting. We ourselves don't feel that we are capable in every field to give consulting. This would be ridiculous for a group of about 16 technical people, and we couldn't hope to cover all fields. What we hope to do in providing consultation is to find an appropriate person somewhere who can serve that role, but we can't.

We don't intend to supplement the standard reference library with on-shelf references. There is no sense in sticking the normal materials of a good library in a computer. That's on the shelf and the standard library techniques work very well. We hope to microfilm all the information that is referred to in our computer-stored information so that if the person wants the reference we can slip him microfilm. We next hope to set up a Safety Information Analysis Center for consolidating this act of boiling things down and having only a few reports in the place of many; validating, in other words getting rid of the junk that isn't correct; and updating, getting rid of old stuff and making sure you're getting the latest in safety information and then prepare safety reports and advisories, much of which would be done under contract.

Now where are we in this matter of establishing the bank? First, our basic computer components have been acquired for the Lewis Center and they are being up-graded which makes me unfortunately say to you that we can't give you red-hot service quite yet because this up-grading step makes the computer unavailable to us for long periods of time. We have now completed the computer programming to give us a very flexible storage and retrieval system for information. First of all we give random access to documents and data citations in the computer storage, in other words very speedy retrieval. We can reach into any part of this storage immediately and pull out the reference without having to spin all tapes through a monitor to pick up the information we are looking for. This reaches in and pulls it out in a fraction of a second. We can fix the retrieval of information by author, by content, in which we use an elaborate system of key words so that you can get sharply focused information, by document origin and number, and I might add by the contractor or other Agency that did the work, by the program name that created the work and so on. There are many ways in which we can find documents under this system. We believe in continuous key-word Thesaurus development. These key words are the descriptives that describe the contents of a document. We know that as documents appear, any fixed Thesaurus will not cover the contents of an evergrowing field, and so the Thesaurus that we are developing can continue to grow with the literature as it comes in and we can always have an up-dated Thesaurus. When a searcher comes to the computer and says I want to find something, what word shall I use. The computer gives him the very latest list of words. The system is very flexible in that if we feel that having enlarged the Thesaurus and the descriptive terms that we allowed ourselves to use, we did an inadequate job of the existing citations in our files, we can go through and change the key words attached to that citation. In the end we hope to be free of any business of a Thesaurus and use free language for characterizing citations. In other words you have a freedom from the constraint of using specialized terms. This is one of the difficulties of finding information in a computerized system. The systems, if they are limited to a Thesaurus, have a rather strange number of constraints.

Let me give you an example of this: Suppose you were interested in cats. And in particular, since you are domesticated, you want a domesticated cat, you want a house cat and you want information on house cats. There are some retrieval systems that would say, "Okay, you can use the word house and you can use the word cat. Because the C in cat comes before H in house it will go into the computer with the word with the C first so it goes into the computer, not as house cat but as cat house. Now who would think of looking for house cats under that. You can do a lot of games with this of course. Try venetian blinds for example. This is true, some systems are this way and give the searcher quite a game to play to try to find the information that exists. We hope to break this block.

We will include a file of document abstracts and reviewers comments in which the reviewer will say the report is pretty good for this area of work but don't believe the title, it just doesn't have very much information in another area or, this is old stuff and it's wrong in this respect. We hope the reviewers comments will be tagged to most of these citations. As I said before, we would have a method for accumulating incremental data in terms of tables and formulas etc. and also the computer has devised within it a means for assisting the searcher in going through the strategy of the search. It keeps assisting him with clues and if he doesn't know what to do next, he asks the computer, "What next?" and the computer tells him.

Here is a view of what we are trying to do now. First of all we find that there are some excellent safety information files. Many of them are computerized, some are not, many of them have this so-called interactive--let me say it this way--we are more or less unique in having this easy interactive scheme of search and retrieval that many do not and where it's justified to absorb a given file or information on safety so we can have this nice access with our computer, we do this. In particular, an excellent file of safety information, which has already been put into our system is a file of about 35-36,000 documents in the nuclear safety field. The files of the Cryogenic Engineering Center and the National Bureau of Standards has already been placed into our system. The FAA Aviation Safety files, we are negotiating on. Recall we said that a complete information system would also include component failure rate files and here is the key word--IDEP--it is an information exchange program amongst the various segments of the Government. It deals with failure rates in the testing of components for aerospace devices--airplanes, spacecraft. By putting this file, which exists on paper, into our computer we can maintain an up-to-date record of all failure rate studies going on, that have gone in the past and those which are current. This will keep some branch of the Government from repeating a failure rate study on a piece of equipment which is already in progress by another Agency. You'll see a sample of the kind of print out this system gives.

Within NASA, following the 204 fire four years ago and then the Apollo 13 accident, both involving oxygen, and other oxygen accidents within NASA, we undertook a complicated and rather involved study of material compatibility with oxygen. This file is going into our computer so that one can find information more readily than the turning of pages in a book, which becomes very difficult.

Here is some safety information that we are asking others to gather with our support and our help. Oxygen System Safety, this grows out of the Apollo 13 accident, in which we are collecting meaningful literature and data and then we are collecting the practices of others in design and operation of oxygen systems. We are trying to put together the fire technology as it applies largely to spacecraft and aircraft and ground test facilities in support of development of either of these. The National Bureau of Standards has a contract with us to do this. They have a fire safety technology group who are charged by Congress to conduct work in this area. This portion of it is a cooperative effort with NASA now.

Human Factors, with emphasis on flight vehicles and especially the space shuttle. This study is going forward under the guidance of the Human Factors people at our Ames laboratory in California and it is to be a major effort. This Nuclear Isotope safety I mentioned earlier has to do with on-board nuclear materials. The business of non-destructive testing and diagnostic techniques with structures on machines safety codes and operating practices for aircraft, fracture mechanics data for structural alloys with special emphasis on low temperature applications of metals and let me cue you in here. NASA has found that every time it took on the use of a high strength material, particularly those which retain their high strength at low temperature, it found it had problems in fracture mechanics--the two ran together. When you try to grab the advantage of a red-hot material that had a high strength to weight ratio and good toughness at low temperatures particularly it had a fracture mechanic problem. The thing wanted to crack easily, which appears to be a contradiction of terms, but this is the way it works. Mathematical techniques in safety analysis, that is only beginning for us.

In an effort to organize our information so that the user can find his problem, we did this. We said, the user comes with certain questions in mind, very often he is concerned with the causes for failure in his systems and we are taking as our illustration this cryogenics fluid safety grid and a means for characterizing the information that exists in a given area and in this area on cryogenic fluid safety, what are the causes of failure? and we say the causes of failure under what conditions. When you are transporting, where you are storing, when you are handling the fluids in systems. These are the blocks which represents an intersection between this term, transportation storage or system handling, and failure causes. Each of these blocks constitutes a range of problems of interest and these then are the categories we create, this range of problems of interest, and place them in this chart so that a person with this problem on his mind under these conditions sees what has been done here. Not only do we do this but all these words that are descriptive terms for describing the literature that exists in this area will appear in this block. That was a simplified view of things--I think you can read the rest and appreciate its relevance to some of the remarks I made before. This is a simplified view only for our purposes here. If I were to show you a true chart, the one that developed for the fire problem, I think you can appreciate that it is a fairly involved chart. The hope that is on perusal by the user, the person who has a problem in mind and then comes to our system and says where can I find information and we give him this, he gets a first clue into how to interrogate the system to find his information. What words does he use to the computer to say give me information along these lines and the computer will begin to formulate a form.

This chart is also used by the people who input the information into the system and any key words that they develop to describe the contents of the documents they review go into those blocks so that the user, the guy searching sees the words that the inputter created to describe the information that exists there.

With regard to the IDEP record, this is the business of putting into the computer a record of the failure rates for equipment under test. The purpose of our computer handling of this is to tell a searcher where he finds the record on the piece of equipment he is concerned for. The address, because the IDEP system provides microfilms of all tests and there is where the information he wants resides. The question is, where is it. In all of the tapes that exist, all of the microfilms, in other words, he is looking for this address, the microfilm address code number. Once he gets that code number, he knows how to spin his microfilm to find out where the information exists. Now he can find the component he is interested in in a variety of ways. He knows the accession number, (I won't try to describe these terms in two much detail, I don't have time) the manufacturer, say the company, of the equipment, the date it was made or the date of the test, or the government part number or a description of the part. May be it is a relay, the contact rating in this area etc. He feeds this to the machine. The machine then prints out a page that looks like this and he can check and see whether this is truly the piece he wants, and is this the correct part number if he has the part number of the Vendor's part and so he says, Yes that is the right one and he knows where to search in the microfilm.

From time to time NASA puts out alerts on parts and this we hope to have in the machine and the key issue here on the alert, not worrying about anything else is this, that people have in the machine a system of alert. If somebody is concerned about what the latest alerts are, he simply queries the computer from a remote station, a console remote from our system, by telephone lines and asks what is the latest alert. He gets a statement which says failure analysis conforms something about a part and the trouble with the seal, etc. and he can identify what the alert is trying to tell him.

With regard to other data centers, we have identified about 150 data centers which we think are useful in our business. There are probably more. We hope to have them within our computer and we ask for certain information and say what data centers would have information on particular things. The computer would print give a print out: which would give them the name of the information center, say Electronics Properties Information Center, and then what do they cover in that center. If you are concerned for liquid metals and hazards associated with these, this is the kind of coverage the liquid metals information center would give you. Not only do you get this, you get information on first, the name of the Center, where it is, how you get information from them, do you call them up, do you send them a letter, do you have a fee to pay, etc. We hope that our Information Center will be one of a network. There are many good ones that have capabilities like ours and we hope that we can tuck them all together in one network so that when you query the system you query everybodies data base. We are trying to make our system consistent with this point of view. If you want to be part of this system and you want to query the information that we have, do you have to call us. I hope not. We would be available for any calls or for any letters in inquiry. What do you have on some kind of problem but we hope that those who are principal users of safety information will have their own console substations which are reasonably cheap. A $5,000 or $6,000 investment gives you such a station. With this tie in, you dial the telephone, FTS or any other voice communication line will put in communication with our computer and gives you the opportunity to access it for information only. This doesn't give you the opportunity to change the contents, only to get the contents out.

It is made of three major components. First a TV screen on which the print-out of the computer is placed and gives you all the information regarding the document you are looking for; a keyboard for instructing the computer on what you want next; and if you see something on the TV screen that you like and want to preserve after making a search, you hit a button on the keyboard and a print-out, permanent record hard copy appears here. These are the three components. For an investment of $5,000 to $6,000, you get them all.

We hope that when our system is rich enough to justify others having remote stations. Our hope is that we can handle many queries, 40 people on the line simultaneously.

That then concludes my description of the work we are doing.

Thank you.