Official Transcript of Proceedings

NUCLEAR REGULATORY COMMISSION

Title: LSS Advisory Review Panel (ARP) 9th Meeting

Docket Number: N/A

Location: Las Vegas, Nevada

Date: December 12, 1994

Work Order No.: NRC-74 Pages 1-225 plus 1A

ORIGINAL

NEAL R. GROSS AND CO., INC.
Court Reporters and Transcribers
1323 Rhode Island Avenue, N.W.
Washington, D.C. 20005
(202) 234-4433

UNITED STATES OF AMERICA

NUCLEAR REGULATORY COMMISSION

+ + + + +

9TH MEETING

LSS ADVISORY REVIEW PANEL

(LSSARP)

+ + + + +

MONDAY

DECEMBER 12, 1994

+ + + + +

LAS VEGAS, NEVADA

+ + + + +

The Advisory Committee met at Bank of America Tower, 101 Convention Center Drive, Suite 450, at 8:43 a.m., John Hoyle, Chairman, presiding. Other members present:

JOHN GANDI

CLAUDIA NEWBURY

JAY SILBERG

BRAD METTAM

MAL MURPHY

KIRK BALCOM

JIM DAVENPORT

LORETTA METOXEN

LLOYD MITCHELL

DENNIS BECHTEL

LAS VEGAS, NEVADA; MONDAY, DECEMBER 12, 1993 8:43 A.M.

--oOo--

MR. HOYLE: I think we should begin. I sense that the people behind us are going to have trouble hearing today unless we speak up. The mikes on the table, I believe, are for the court reporter as opposed to anyone else in the room.

I was waiting a little bit for some other of our members, but I think they'll come on in. They told me they were coming.

Today we meet for the ninth time as the LSS Advisory Review Panel. This is the panel which meets in accordance with the Federal Advisory Committee Act rules. It's an open meeting. There are copies of documents on the table to my right from the DOE side and the yellow sheet is the agenda for our two-day meeting.

Before we go any further, I would like to have those of us at the table introduce ourselves. Let me start with Claudia or maybe John.

MR. GANDI: John Gandi, Department of Energy.

MS. NEWBURY: I'm Claudia Newbury, Department of Energy.

MR. HARDWICK: Roger Hardwick, Clarke County, representing Dennis Bechtel.

representing Nuclear Energy Institute and the Industry Coalition.

MR. HOYLE: I'm John Hoyle, from Nuclear Regulatory Commission.

MR. CAMERON: Chip Cameron from the Nuclear Regulatory Commission.

MR. LEVIN: Moe Levin, Nuclear Regulatory Commission.

MR. METTAM: Brad Mettam, Inyo County, California.

MR. MURPHY: Mal Murphy, Nye County.

MR. BALCOM: Kirk Balcom, State of Nevada, just in time.

MR. ~~SWANSTON~~ Davenport: Harry Swanston -- no, Jim Davenport, Nevada.

MR. HOYLE: Okay, thank you very much.

MS. MITOXEN: Mr. Hoyle, I'm Loretta Mitoxen.

MR. HOYLE: Hi, Loretta.

MS. MITOXEN: Hi, how are you?

MR. HOYLE: Fine.

MS. MITOXEN: I haven't been here for awhile, but I work with the International Congress of American Indians. And this is Lloyd --

MR. HOYLE: This is Lloyd? Okay.

MS. MITOXEN: And he's director of the environmental department with the United Nation in

Wisconsin.

MR. HOYLE: Okay, why don't we see if we can make room for the two of you.

Loretta, I think I left your name tag at home. I'll make up one in a moment.

MS. MATOXEN: Oh, that's fine. Thank you.

MR. HOYLE: We do have one for you.

At our last meeting in September, we had a -- the opportunity to have a preliminary overview of the actions taken by DOE since moving responsibility for LSS to the Yucca Mountain Project office. Today we expect to hear the outcome of their working groups' activity.

Before I ask Claudia to start us off with the update of DOE's working group activities, let me just briefly run through our expectation for the two-day meeting. We'll spend the morning being briefed and discussing DOE's update. After a lunch break, we'll hear from Tom Nartker about progress toward an efficient LSS overview of his work at UNLV. In the afternoon we'll talk about the proposal concerning operation of the LSS, and perhaps setting up an LSSARP technical working group, which has been an idea of Dennis Bechtel's for some time.

Roger, is he coming this afternoon, perhaps?

MR. HARDWICK: Yeah. It will be late this morning.

MR. HOYLE: Okay. Tomorrow we'll have a review update of the topical guidelines activity, an overview by the LSS administrator Moe Levin of participant compliance, and a discussion of the LSS on a pilot project basis. Tomorrow afternoon, we've been invited to see a demonstration of OCR technology at UNLV. Tom Nartker, as I mentioned earlier, will brief us after lunch today. I think he will give us some guidance on how to get over there and a little bit about what he expects to cover. I've put on the agenda 1:00 to 4:00. Perhaps it won't take that long. So, unless there are comments to be made by any panel members, I'll ask Claudia to start off the meeting.

Okay, Claudia.

MS. NEWBURY: I'm up here. Can you hear me? As you recall, at the last meeting we had a brief discussion about what the working group has been doing in terms of looking at the LSS, what are the requirements for an LSS and how can we efficiently put together a system that's workable and useable to everyone who wants to use it. That was in September. And at that meeting, I committed to putting together a more complete presentation of what the working group has been doing and I reminded you that we were still in the review process. Well, we are still in the review process and looking at some of the comments that we had on this working group report.

We have -- as you are listening to what the working group is talking about today, please bear in mind that DOE does have their own concerns with what's being presented. Some of those are, that we'd like them to redo the model that they've been working on with the change in the schedule, especially with the PPA. They mainly need to re-look some of the costs that are resulting from the change in schedule.

In the meeting in September, we had four work stations, I believe, and there were a lot of comments on that. We would like to look at having more work stations, and that will change the cost profile a little bit also.

Also, we did not have adequate volume projections which I hope to get today from the participants. I shouldn't call you -- the effected units of Government in the state, so -- and the NRC, so that we can pull that in. And, also, we are looking at combining some of the options. There are some better ways, perhaps, to do the work than the M-member options that were considered. So, as we -- as you're looking at what the working group presents, please bear in mind again the DOE still has comments on this.

Just a brief history -- a reminder of what's going on. The responsibility for LSS was transferred to YMSCO -- whatever we are now -- in fiscal year 1994. We chartered a working group to examine the LSS and develop an

implementation strategy that's consistent with our program approach. They submit -- the working group submitted a draft report in October of 1994. We sent it out for review comments. The comments came back by the end of November and we're currently asking them to incorporate those comments into the report. We expect to see the final report by mid January. This was the basic charter of the working group.

Okay, looking at the development history of the LSS, they wanted to identify the commitments and expectations for the LSS -- both the ones in the regulations and ones that have been articulated in the LSSARP meetings. What are the statutory requirements? An evaluation of the implementation options. And again, we want to look at a few more options, for the total life cycle costs and the expected date of volumes.

Today, what I'd like to do is have members of the working group present a lot of this information to you and these are the people who will be discussing these: An overview of the NRC hearing process, Stan Echols; the OCRWM Records Management System from Marty Cummings -- please remember that the records management system is not the LSS; that we may incorporate stuff from the records management system into the LSS, but we're not considering them the same system; what the LSS requirements actually are -- the LSS functional models and cost evaluations, then schedule and

cost profiles, the system level requirements document, the volume projects, and finally conclusions and recommendations with the people who are indicated on these sheets.

So, with that, what I'd like to do is turn it over to these people and let them tell you what they've been doing and at the end, we'll come back and answer your questions. Okay? So, we'll start off with Stan.

MR. ECHOLS: How does this thing work?

MS. NEWBURY: There's a little button down on the thing. Keep going down where it say "On".

MR. ECHOLS: Okay. You can tell from that portion I am not the computer guru.

I'm Stan Echols with Winston & Strong. We provide assistance to the M&O in meeting its contractual obligations to DOE with respect to NRC regulatory matters. The Department of Energy Office of General Counsel provides legal advice to the Department of Energy itself. I'm going to be very brief in this. There's a handout that provides the gory detail of milestone and provides an annotation about what each of the milestones involves. We want to get to the substance really, of what's going on. If there are any questions regarding the licensing process under NRC regulations, Chip Cameron of the NRC General Counsel's office will provide detailed information to you, I'm sure.

(Pause.)

MR. ECHOLS: I'm just going to touch briefly on the initial process. Part 60 and Part II of NRC Regulations' Guide, the bulk of this -- the license application under Part 60 -- I look at 10 CFR 60 21 and 31. Part II covers discovery, motions, (j) -- the one we're most interested in focuses on document discovery. The layout for the administrative proceeding itself, which is a trial-type hearing, is also in Part II, and the requirements for the construction authorization itself are covered in Part 60.

The license application itself -- went to -- there were several milestones of -- actually, there's one that precedes the LSS administrator certification, really, and that's about a half a year before the certification. There's a sort of a prelicense application presiding officer that is identified during the half a year or so before certification, or hopefully longer. Documents are submitted, the LSS itself is tested and the prelicense application presiding officer acts in an administrative capacity to try to resolve any issues that might come up regarding documents that would go into the LSS if there is a dispute as to what is privileged or not or what's appropriate, what's relevant or what could lead to relevant information. That sort of thing. It's that person's role to help define and resolve those issues.

Also, the NRC, I think, is supposed to provide

further guidance as to what is relevant or could lead to relevant information for including into the LSS itself.

After the certification -- the certification has to occur roughly a half year -- or a half year before docketing. If there is no certification, one of two events occurs: There is either a hold until certification in fact occurs, or there's an option to a docket outside of the (j) realm which is outside the use of the LSS under traditional licensing under (g). The note there, 60.18 footnote two and 60.18 (g) relate to the fact that prior to formal interaction with the NRC, there is a rather rich period of time pre-docketing in which the NRC and DOE interact to identify and reach closure at least at the staff level on a number of technical and regulatory issues. Such things as topical reports, if there's any rulemaking, and so on.

The contents of the license application broadly -- there's a class of general information. The safety analysis report which deals with all of the issues identified in (e) to Part 60 -- this is the technical and design information. There's an environmental impact statement the Department of Energy will prepare to support its recommendation of a site, if any.

Under the Nuclear Waste Policy Act and NRC regulations, the NRC is to adopt that EIS to the extent practical. And I presume there will be some close looking

during the years preceding that, to make sure that the NRC's concerns are conveyed to the DOE. The general information again: schedule safeguard certifications, physical security plan. The safety analysis report -- again, the technical information, performance assessments of models that are used, the QA program, the alternate designs of major features, the confirmatory design program. Also, it's anticipated that the license application will in fact not be complete in certain areas. This is spelled out in 10 CFR 60 and it's incumbent upon DOE in the license application itself to present a program for ongoing studies to resolve any open safety questions, or whether or not there needs to be a resolution in case -- for instance, all parties agree that they've been sufficiently bounded.

The Environmental Impact statement again, to be adopted to the extent practical by the NRC -- the NRC NEPA requirements are found in 10 CFR 51. The DOE requirements are in 10 CFR 1021. Then, not only is the LSS a document discovery tool, but once the hearing begins, the filings of the documents themselves throughout the process will be done electronically; the theory being that again, time will be saved.

Just briefly, what goes on here: During the discovery process which is covered -- for document discovery, again, this is covered in Part (j). The reason

we're here under the LSS -- other discovery reverts to the duly applicable requirements under (g) to Part II, which -- interrogatories which are questions answered under oath; depositions -- where a party is asked, again, questions under oath, and they're stenographically or otherwise recorded.

Request to admit Yucca Mountain is located in Nevada. Yes, I agree. It's really to clear out as much as you can of the basic kinds of questions, the easily identified issues that are not in controversy. Later on towards the end of the process when the discovery is nearly complete, you will have motions for summary judgment and other motions. These again are to narrow issues so that once the trial-type hearing itself begins, the process is focused on those issues that are in fact in dispute. Again, the electronic filing of documents.

I know this is buzzing through quickly, but is in richer detail and you can read that at your leisure through a much more detailed handout.

The first pre-hearing conference is interesting in that the -- those who wish to become parties to the proceeding will file with the NRC that desire, the basis of its contentions that it might have, and in the first pre-hearing conference, the NRC gets to work on consolidating some of these issues, seeing if it can consolidate parties,

sifting through some of the contentions to see which ones will survive. As a practical matter, earlier on, the NRC tried to tighten up a little bit under Part II, the admission of contentions and the requirements to support them, because in the past it was pretty much -- everything got in.

As a practical matter, just drawing back on some of our recent experience in licensing the enrichment facility of Louisiana, the process now appears to be -- everything gets in, so that very little has changed in that respect. The process seems to be -- and I would let the NRC speak to this themselves, but to assure that all issues are vested -- they appear to be very liberal in admitting contentions, and then towards the end of the process in motions for summary judgment, they begin to have a much more critical eye at the back end. So they give the parties full opportunity to -- through discovery, to explore contentions, and that seems to be the same under the new as opposed -- as well as the old rule, and then focus on the process towards the end in motions.

The second pre-hearing conference is a misnomer. It should be the last pre-hearing conference. This assumes only two pre-hearing conferences, when in fact there will likely be several over the time frame. This is when you do finalize the issues, identify witnesses, establish the

hearing schedule; the housekeeping chores involved prior to the hearing process itself. This is when the discovery has been completed and then you -- and you now begin the licensing process -- the hearing process itself. We just discussed the motions.

I just wanted to go to the reality check on the three-year schedule. The -- again, the enrichment facility is a fairly straightforward chemical processing type plant, zero release. More of -- a chemical environmental and safety impact more than radiological. We're already at the three-year mark on that and counting. We finished the safety hearing and the environmental hearing is scheduled after the first of the year. That is a very straightforward project. The technology has been in existence in Europe and is being applied in centrifuge technology. It is very ambitious to presume a three-year schedule for this project, which is why I suppose Congress, in its wisdom, allowed a one-year variation there so that the three years not be maintained.

As noted before, it's an adjudicatory or trial-type hearing with Part 60 and 51 covering the safety environmental issues. By that, we mean that as opposed to a paper administrative process where you merely file positions and the administrator would review those and come up with a decision, here you actually have witnesses that are examined

and cross examined. You have expert witnesses.

There is various ways in which experts can provide testimony. There's some work now being done at the NRC and elsewhere with expert elicitation, for instance, on climatology. I know there was some inquiry if that might be the only way to have expert witnesses in the hearing process. I think the answer to that is "No", that under the Federal Rules of Evidence which the NRC is not bound to, but generally follow, experts broadly come in, and then the credential of the expert -- of the process -- through an expert elicitation, or whatever, goes to the weight -- how much credence will be given to the expert. But I think they would -- they should tread lightly on --

MR. SILBERG: Do you want to just explain briefly what "expert elicitation" is?

MR. ECHOLS: Well, this -- the -- and I think the NRC might be able to respond to this more directly, but on the issue of climatology, for instance, they put together a panel of experts in the area -- it was almost like a many negotiated rulemaking, I suppose. It took about a year, I think, to come out with the -- with an answer with respect to their view as to the effect of climatology over the next 10,000 years or so.

Some of the lessons learned, as I recall from -- I think it was -- Margaret Federline briefing to the ACNW, was

that they needed some kind of guidance with respect to the ultimate application of the elicitation to put it into context. In this case, a regulatory context. What will that information be used for? What kind of decisions will be made using that information? That helps to set the bounds of the inquiry.

An individual expert comes in on the basis of his education and life experiences in a particular area and he can be challenged on the adequacy of that experience or education and background. But, again, generally, if an expert is proffered and they survive the original round of questions probing his area of expertise, that person's testimony will come in and the challenge will go to the weight of it.

Is one expert more believable than another? Perhaps as an individual, how does that person stack up to a use of panels or through an expert elicitation process -- but I seriously doubt that if -- an intervener who proffers an individual expert would be denied the ability to have that expert testify on its behalf. For one thing, expert elicitation is a fairly costly and time-consuming process.

MR. DAVENPORT: Stan, give us your opinion about the use of experts in written form who are not personally present.

MR. ECHOLS: How do you mean?

MR. DAVENPORT: If someone were to offer a report authored by an expert as that expert's opinion without the presence of the author.

MR. SILBERG: It has to be submitted by somebody. It has to be sponsored by some --

MR. DAVENPORT: Right.

MR. ECHOLS: This is not an unusual -- in a former life at the NRC as a project manager, we worked with National Laboratories and other experts with the licensing process of construction and operation of nuclear facilities, who would cover quite long time spans. And there would be reports and studies done in one decade that would be used in the proceeding in the next decade. And --

MR. SILBERG: Substitute "century" for that.

MR. ECHOLS: That's right. And you -- it goes to the administrative record as being preserved -- about that -- how was -- what was the charter of that expert? How did that expert perform his task? Was it done under the appropriate methods? The QA comes into play here to document the history of that -- the study upon which whatever party is relying. So, this is not unusual at all -- the fact that there is a fairly wide span and the individuals who started the process were not around for the end of the process. Particularly here. I mean, you're going to have ongoing studies for 50 years. We all better

take our vitamins if the -- if the originator of the study has to be around with the back end of these things.

MR. DAVENPORT: Your hair is grayer than mine, Stan.

MR. ECHOLS: That's right. and if you're here for the post-closure hearing, we'll all be in with our walkers and respirators and so on. So, no, that does not affect the weight to be given any documentation.

MR. DAVENPORT: But we all concur that a freestanding document without an expert to proffer it is not an admissible document.

MR. ECHOLS: I don't concur in that.

MR. DAVENPORT: Oh, you don't?

MR. ECHOLS: It would be -- you say an expert. The person proffering it might be the manager in an area where under a prior manager's supervision, a document may have been produced. And the current manager says "Yes, under the administrative process" -- like a business records kind of rule -- that this was produced in accordance with the processes that were in place. And I wasn't around then and I'm not an expert, I'm a manager, but we've reviewed the records of the prior manager and the prior expert and this was done according to the administrative procedures that were reviewed and concurred in, and here's the QA of the record and so on.

MR. SILBERG: But, someone has to be available for that document to be cross examined on.

MR. ECHOLS: Oh, surely. And there might be a --

MR. SILBERG: A -- manager or someone.

MR. ECHOLS: And there may be a -- and there may be a panel --

MR. DAVENPORT: Just in the case of a business record, that record has value as to the matter of fact of what it has on its face. And what we're talking about here is the expertise --

MR. ECHOLS: Oh, you have to have someone to cross examine.

MR. DAVENPORT: -- that's behind -- oh, excuse me.

MR. ECHOLS: Yes. There will be a -- there may be a panel, there may be an individual that's a technical expert in that area. I've reviewed the report and the document and it comports with the standard practices in the field, and I'm supplementing that report with new information that's occurred over the past 20 years. And that report, plus my supplemented information, constitutes the testimony of the agency now. That kind of thing. So, there may be building blocks. It may not be a stand-alone, but it may be a document of supplement. But there has to be a warm body to be examined on that document.

MR. DAVENPORT: Who is offering expert opinion.

MR. ECHOLS: Yes. Now, if it's a background published document -- for instance, if there are National Laboratory, USGS reports, textbooks, those kinds of things --

MR. SILBERG: Then you get into official administrative --

MR. ECHOLS: Yeah. I mean, you can get off into tangents here. You don't cross examine the author of a textbook that's certainly accepted. You might bring your expert in to say there's no information that would challenge the textbook, or whatever.

MR. DAVENPORT: Okay, I see.

MR. ECHOLS: That's getting in a lot more detail. I want to get the people up here to talk about the LSS. Like I say, it's in gory detail over there, and we're talking about the NRC licensing process. Chip is the resident expert, I think, and would be more than willing to answer all questions.

MR. CAMERON: This is a real setup.

MR. ECHOLS: It is.

MR. MURPHY: Been through a licensing hearing, Chip?

MR. ECHOLS: Just briefly, somebody -- the issues that are discussed as part of the hearing process -- safeguards, in this case under the requirements are really

certifying that the safeguard procedures used by DOE will comport to those used at other facilities. QA -- these are all some of the key topics that will be discussed, I'm sure.

After the trial-type hearing ends, and assuming that the board -- the hearing board reacts affirmatively, the NRC will issue a construction authorization. And more than likely, based upon the hearing record, that board will place upon the construction authorization certain conditions. And those conditions will relate to health, safety or environmental issues. And for those kinds of conditions, you can look at any construction authorization for a nuclear facility, for instance, and you will avoid construction in certain areas. You will -- instead of having a cooling pond, have cooling towers. "Instead of two pumps, we want a third pump. Here's the operating procedures we want you to beef up," and so on and so on.

And that, I think, is it, in a very brief nutshell. It's -- I think that the NRC anticipates that the entire hearing process will last 90 days. Again, I think that's a very optimistic schedule. And the -- all that is being done with the LSS is to focus on two aspects: One is document discovery, to move that along and to help crunch a lot of documents in a very short period of time; and two, during the hearing process itself, is motions, practice and just -- the distribution back and forth of documents during

this three-year period would be done also electronically. That's it.

MR. HOYLE: Thank you, Stan.

(Pause.)

MR. HOYLE: Martin Cummings.

MR. CUMMINGS: Good morning, my name is Martin Cummings. I'm Martin Cummings, I'm the records manager for the management and operating contractor in Vienna, Virginia, and I was a member of the LSS working group. What I've been asked to do today is to present to you the current status of the OCRWM, DOE OCRWM Records Management System, let you know where we are today and where we're going to be going in the very near future.

I do want to stress that this is our records management system that we're talking about, and although we are -- we're trying -- you know, very hard to be in sync with everything that's going on with LSS and everything, there are -- you know, certainly reasons why we have the system that we do today.

It started out, I think, basically as a litigation support system quite some time ago, and we still continue to do that with our system. We also have the requirements of managing our quality assurance records, and also -- because I think that we're kind of a unique organization in the sense that we have two very high-level bodies that provide

requirements to us. And I think we're unique in that sense that in a lot of your nuclear activities, they -- you know, they get the requirements strictly from the Nuclear Regulatory Commission and on the -- conversely, you have a lot of Government organizations that don't have those requirements, that have -- that get requirements from the National Archives and Records Administration, NARA.

Well, we're one of those programs. We have to satisfy both bodies, and so as a result, it can be somewhat challenging to do that. So, just from a good business practice, we have a rather significant records management organization.

Now, obviously there is a very close -- going to be -- and is -- and going to be a very close relationship between the OCRWM Records Management System and the LSS, for the main reason that all of the data that's going to be submitted to the LSS -- whatever that may be, is going to come from our -- that is, DOE's input will be coming from our records management system. We still talk about estimates of that being -- 85 percent of the LSS data would come from DOE.

MR. MURPHY: Do you include all the contractors in that?

MR. CUMMINGS: Yes, sir.

MR. MURPHY: That's DOE, that's everything?

MR. CUMMINGS: Yes, sir, that's correct. So, obviously we want our LSS and the records management system requirement to be very consistent.

The critical overlap certainly would be our header fields, our data formats and then what we include into our records management system. Just to give you a real quick description of the current record management system, it is, as I said, the way that we capture and manage all of our program records. And I'll talk a little bit about maybe what program records mean on the next couple of slides.

We have our own policies within the office of Civilian Radioactive Waste Management that we have to follow. We're also instrumental in helping develop some of those policies. We have several initiatives ongoing right now on the development of supplemental DOE orders. And, also, our system with regard to requirements has to support all of our sites.

Back in the Washington area, of course, we have several sites there with regards to the M&O contractor in Vienna, we have the Forestal Building with -- you know, from headquarters out here on the Yucca Mountain project. We have numerous contractors here in Las Vegas and then what we -- you know, call our externally based to Las Vegas -- our participants all -- you know, the labs and USGS. And, so, I -- it's 10 or so sites in all, and we've worked very

hard over the last couple of years to try to provide an integrated records management system.

Our architecture -- first of all, please, everybody with a copy of the brief, can you please draw a line through that firs, bullet? That is a major typo. It is a VAX Basis Plus System. That was for one of our other information systems and I didn't catch that. Right now, we do have a database, as I said, in Basis Plus, that is the -- you know, that supports our header system, and our images right now are microfilmed.

With regard to the operation features, as I said, the M&O contractor is tasked with managing our records system and we're doing that based -- as far as -- you know, including screening criteria, we're doing it based on NCFR2, which I'm sure everybody in this room is very familiar with. Also, the OCRWM quality assurance requirements document -- or the Yucca Mountain site characterization plan. We have other waste acceptance requirements and storage and transportation requirements.

Now, what we exclude from our system is all of our administrative personnel and financial records. And I wanted to make that point, because in the past, I've noticed that when we talked about the OCRWM Records Management System, that there was some understanding that every single record that we had was in that system. So, I just wanted to

clear that up -- that we have our own inclusion/exclusion criteria right now within our own system.

MR. SILBERG: Are those documents in another system?

MR. CUMMINGS: No, sir, I guess the answer would be no to that. It's --

MR. MURPHY: They have to be somewhere. You have to have personnel records sufficient to establish the training qualifications of people in your -- in the QA portion of --

MR. CUMMINGS: That's correct. I -- those records certainly, I would -- could be characterized as personnel records and administrative records. But I'm looking -- I'm talking more -- when I mention here about -- administrative personnel from the corporate perspective, not from the quality assurance area. Those training records are certainly in there. And we have our system of -- you know, protecting proprietary information. So, you might have a header field in there but it would take -- certainly, the access to that would be controlled to actually get the document.

MR. SILBERG: Does that mean that those documents only exist in hard copies and not in anybody's automated data processing system or automated system? The records have to be there -- some format. Is it in another computer-

based system? Is it only in a hard copy system?

MR. CUMMINGS: I would -- only in a hard-copy system.

Camille, you have a --

MS. KERRIGAN: I was going to say that -- you know, you have multiple contractors and a lot of this information is proprietary, so each contractor would have their own system. But, certainly, they maintain it, and at the M&O -- at TRW, at least, we are moving from the paper base for some of the financial management into an image base system. We're looking at that technology right now. But, there is no integrated system -- is what I think Marty is referring to.

MR. CUMMINGS: We don't put a lot of process -- I mean, obviously processing power is very expensive and so it's -- you know, that kind of information that we don't run through the very stringent records management system that we have now. And we don't put a lot of processing power behind that information.

MS. KERRIGAN: But it is retrievable.

MR. CUMMINGS: But is retrievable.

MR. GANDI: I think you'll find on the sensitivity of -- like personnel files, that the personnel organizations will have their records filing be a hard copy or a small database on the PC, but when we look at project-related

training records which are sensitive in nature that are under an automated system, they're more or less separate from the records management organization. And when Marty speaks of the record management system, we're talking about project-related documents -- programs.

(Pause.)

MR. CUMMINGS: Well, some enhancements that we already made to the system and about to make to the system would include our software for our indexing records in our system. We've gone to what we call the IRIS, the Interim Records Information System. It's a much more user-friendly system, it's -- it is what we index our records into. It's improving the integrity and accuracy of our data, certainly, and it's helping us out on our retrieval capability. And it is what we will be using to support the reprocessing of all of our records that are currently on microfilm. And, also, the current records that we have -- when we reprocess them, we will be -- you know, screening them in accordance with our new and updated inclusion criteria.

In other words, at one time we put -- you know, I -- we put everything into the system, so -- I mean, you know, there's rental car requests in there, there's -- you know, that kind of information -- that we certainly want to purge our database and get that kind of obvious excluded material out. And we don't want to put a lot of processing

power behind that kind of information.

MR. DAVENPORT: Want to tell us the process you're using? How are you going about this reprocessing?

MR. CUMMINGS: First of all, reprocessing includes three steps. It includes the screening of the information against inclusionary/exclusionary criteria, and then if needed, it would be re-indexed into the new system and then it would be electronically imaged so that we -- you know, can get out of the microfilming business, which is what the next bullet relates to. That's being done by the management and operating contractor.

MR. DAVENPORT: And the inclusion/exclusion criteria are published?

MR. CUMMINGS: Yes, sir. It's -- the exclusion criteria, though, is certainly based on 10 CFR 2 support, yeah.

MR. BALCOM: Did you say who the contractor is?

MR. CUMMINGS: TRW Environmental Safety Systems is the prime contractor for the M&O.

MR. SILBERG: When you say "supports reprocessing of 500,000 records", is that total size of the current --

MR. CUMMINGS: Yes, sir, that's correct.

MR. SILBERG: -- IRIS?

MR. CUMMINGS: That's correct. We have an average of around 225,000 records in Vienna and we have another 200

and what, Jan -- 25, 30 thousand or whatever out here. So it's around 500,000.

MR. MURPHY: Is that total over the life of the project or --

MR. CUMMINGS: That is the total number of records that -- from the life of the program so far.

MR. MURPHY: Backlogged.

MR. CUMMINGS: Yes, sir, that's --

MR. MURPHY: That's not a projected "This is what we're going to reprocess" and --

MR. CUMMINGS: That's existing today. Those are records that have already, in some form or fashion, been indexed into the records management system and put on microfilm.

MR. GANDI: Using the screening, did you see a reduction -- rate reduction in those --

MR. CUMMINGS: Yes, there is a -- just based on some numbers that we did the other day, I would venture that number -- probably around 75 percent of that would only be relevant after it gets reprocessed. And like I said, that's the obvious excluded material; administrative-type things. You know, like I said, rental car -- there's rental car requests in there.

MR. SILBERG: What's the schedule for this reprocessing? How long is it?

MR. CUMMINGS: Actually, we have a schedule in the LSS working groups' report, and basically it's about a two-year effort.

MR. SILBERG: Starting when?

MR. CUMMINGS: Hopefully, starting in the fourth quarter of this year -- of this fiscal year. And the reason for that is, because we're just now going to be going to our electronic imaging system. We currently have a system that's in testing and evaluation, and if that works out, then I guess hopefully around May we'll have our electronic imaging capability. So, obviously we don't want to start any reprocessing effort until we get that capability.

MR. SILBERG: This population of documents, does this include everything that's in the contractor's hands up to this point?

MR. CUMMINGS: Yes, sir, that's correct. I mean, that's --

MR. DAVENPORT: Kind of remember it's documents, not pages.

MR. SILBERG: No, I understand that.

MR. CUMMINGS: Yeah, this 500,000 records has an average of 13 pages per record.

MR. MURPHY: It's just OCRWM. It doesn't take -- doesn't look at records anywhere else at DOE; is that -- office, or the weapons complex or --

MR. CUMMINGS: That's not exactly true. We -- in fact, my office in Vienna is working with E&L, the environmental side of DOE that has some of the DOE-related records -- you know, like from Savannah River and -- so that -- I mean, this -- the record management system does include those records. But we really haven't gotten a whole bunch -- we're just now getting into what -- we're working with those people and laying out our plans and strategies of capturing that information.

MR. CAMERON: When you talk about life in the program -- or as it's -- as all the records in this particular program, I think that a lot of these questions go to -- what's the definition of the program in terms of duration? In other words, is it the creation of OCRWM? And you've already answered the question about OCRWM documents. When you say "this particular program", what do you mean in terms of the time element?

MR. CUMMINGS: This particular records formation system, I think, started in 1987.

MS. KERRIGAN: '88, '87.

MR. CUMMINGS: 87. Okay, but it includes records from back to like -- and we've got them in there from -- what, '82?

MS. KERRIGAN: The early '70s.

MR. CUMMINGS: Early '70s. Yeah, I guess.

MR. CAMERON: So, you do. Okay.

MR. CUMMINGS: Yes, sir. It does include that. I mean, this -- the current system we have today has gone through several migrations and transitions over the years.

MR. HOYLE: Is this conversion to IRIS part of a step in what we used to hear being called Info Streams? Is this Info Streams out now or is this taking its place?

MR. GANDI: IRIS isn't associated with Info Streams.

MR. HOYLE: It's not associated.

MR. GANDI: No, it's a stand-alone records management system.

MR. CUMMINGS: All right.

MR. GANDI: What's in that system?

MR. CUMMINGS: If -- depending on what the plan is for, an Info Streams-type system would be -- I mean, it certainly would be integrated with whatever that would be. But, I mean, it is the stand-alone system.

MR. LEVIN: Well, what part does Info Streams play, if any, under your RMS? Is it a modulant RMS or is it totally separate?

MR. CUMMINGS: Our current -- the current system that we have now for capturing records is primarily based on -- you know, capturing hard copy records. The Info Streams type of a system would be the way that we'd -- you

would capture the records electronically from the creation -- you know, from the author of the document creating it and -- you know, capturing it right from the get go. Whereas right now our system is based on people generating a hard copy record and then it gets submitted to our system.

MR. LEVIN: So, Info Streams is not a part of the RMS?

MR. CUMMINGS: I think that would be a fair statement. Not at this point.

MR. MURPHY: So, what happens when this stuff is all reprocessed and captured electronically? Do you have two systems then? Info Streams -- that's processing and has captured all prospective documents that were generated from today on, and RMS has all of the documents that were generated from yesterday back? Is that what we're talking about?

MR. CUMMINGS: No, I don't --

MR. GANDI: I think we're confusing --

MR. MURPHY: Well, the reason we're curious about this is that in September -- I think it was the September meeting, wasn't it -- well, we went through a process of coming to concurrence on the notion that at the NRC's request, the other participants -- and some of us somewhat reluctantly, agreed that Info Streams -- Info Streams would

be an appropriate vehicle on which to base the LSS.

MR. GANDI: I think we're talking --

MR. MURPHY: Not something else, but Info Streams.

MR. GANDI: -- of the Info Streams.

MR. MURPHY: What?

MR. GANDI: I think we are talking the same minds as the architecture -- the base architecture of Info Streams becoming an LSS design criteria.

MR. MURPHY: What's going to happen to this stuff, then?

MR. CAMERON: The data that's currently managed by Info Streams, I think, is the issue.

MR. GANDI: Yeah, I think if we look, the managed -- the data managed by Info Streams is very limited right now. There are several modules to Info Streams. Currently we haven't gotten past the creation and concurrence of that document with Info Streams.

MR. CUMMINGS: I think also, a lot of the capability that you're sitting there talking about with Info Streams -- I don't want you to think that it -- you know, we certainly have capability within this system. I mean, when we go to our electronic imaging system, we're going to have the capability of capturing the headers, okay -- and right now our header fields are based on the LSS header working group -- okay? We're going to have electronic image and

that electronic image will be run through the OCR conversion process. So, we're going to have header, image and text.

MR. SILBERG: Okay, but what's the difference between that and Info Streams?

MR. CUMMINGS: Info Streams is the ability to capture that record at the time that it's generated and capture it in electronic form. Right now, you know, I get it when somebody generates a record off of a PC and they print it out and then they submit it to the records system.

MR. GANDI: Just to bring this down to maybe a functional basis, regardless --

UNIDENTIFIED SPEAKER: What's the relationship?

MR. CAMERON: Just to bring this down to a functional basis, regardless of whether we label something Info Streams or whether it's part of IRIS or whatever, the Commission -- and we did have a discussion September of -- a year ago on this. The Commission was interested in building on whatever the DOE system was for capture, and I think that's the most important point, rather than arguing about Info Stream.

MR. CUMMINGS: Yes, sir. And those elements are in this system.

MR. CAMERON: I think what you're saying is, the basic architecture -- Info Streams is an application using a basic architecture, and the basic architecture could be

ordered over to another application. So, you're not throwing away all the effort that was done -- put in to develop the architecture.

MR. GANDI: No --

MR. CAMERON: Info Streams is just an application, is what you're saying.

MR. GANDI: It's more than just an application or conceptual view of how documents are processed, stored, retrieved and so on.

MR. SILBERG: I remember discussions a long time ago -- two, three years ago, where I thought that DOE was already at that time capturing documents electronically as they were being generated. Is that not the case?

MR. CUMMINGS: I'm going to let Michael Fisher, who has a lot to do with our -- some of our development --

MR. FISHER: I'm with TRW. I'm the deputy project manager for Info Streams. I'll try and --

UNIDENTIFIED SPEAKER: I can't hear you, sir.

MR. FISHER: I'm Michael Fisher. I'm the deputy project manager for Info Streams. I'll try and clarify some of this. Info Streams was to be built in four increments and the first increment was merely a pilot increment where we standardized the office automation -- COT's, and we tried to use some of those COT's to do some rudimentary document routing, et cetera, to make that process a little more

efficient over a -- you know, a geographically disbursed enterprise. Washington, Virginia, Las Vegas, Nevada.

The second increment of Info Stream was to move into real document routing and concurrence to help capture that information which was already being done in paper form. But, again, it was an efficiency across a geographically disbursed enterprise.

When you got into what increments three and four were going to be, it was more automation of the records management system. And increment four, in particular, was going to be image capture, which is what Marty is about to talk to -- graphics to be replaced by electronic imaging. And we currently have two systems that we've built that are in pilot test and evaluation. One's in Vienna, Virginia at the records center, and the other is here in Las Vegas in the records center.

Now, as I said, that is -- that last piece is in the records management piece to do -- they need to capture, and right now all you can do is scan hard copy paper and get it imaged and you get your indexing information out of IRIS. So you have a hook to your legacy indexing system and then it allows better retrieval, where -- you know, evaluating the use of full text as well as index retrievals.

This is -- increment two, which is in a pilot evaluation back in Vienna, Virginia -- again, it's just the

automated routing and concurrence internally. They still have to generate hard copy and turn those records over to the records center. Eventually it was envisioned that those things would be tied together. And we haven't even addressed what was going to be in increment three, which was automation of other indexing and records management functions, et cetera.

So, the thing hasn't all been brought together and it's not even clear that it will be, because these are in pilots that are under evaluation. But the point to be made, and -- from NRC, was trying to make it to is -- you know, architecturally, we're using some of the same back-end products so we're leveraging what we've been doing on -- in records management. And if the decision is made to integrate those all together and subsume IRIS or replace it or whatever terminology you want to -- you would want to use, everything would be common and you could integrate that a lot more easily than if you had had -- entirely, independently developed systems that had no commonality whatsoever and you'd probably have to go back to square one and do it.

MR. CUMMINGS: As I said, we currently have a microfilm-oriented system and we -- in records management, we would -- we have -- we know what the LSS requirements are and we're moving in that direction. But, I've got to tell

you that even if we didn't have any LSS requirements, I mean, this is just the -- this is the next step for our records management organization. Just like the guy that said -- you know, "Look at all this hard copy data. Wouldn't it be neat if we could microfilm it?" Well, now, we're going to the next step. We're -- we have the capability of having electronic imagining and we want that in our records management organization. So we will be moving to that, and hopefully in the not-so-distant future.

Now, because of the -- our new database and the fact that we want electronic imagining, we certainly anticipate enhancements from that. We know it's going to give us a better retrieval capability, it's -- we're getting better and more efficient in our indexing practices and after we get really good at automating our current system with the electronic imaging capability and we get really good at that -- because, you know, change is something that -- you know you have to be very careful about and right now our people are working in a manual -- more of a manual microfilm environment today, and tomorrow they are going to be going to this electronic imaging capability.

So, we want to get really good at that, and then when we do that, then we're going to really look very hard at our work flow practices and just take it to the extent of what -- electronic imaging capability can really help us.

So, we think there's an opportunity to do things that are even a lot better, a lot smarter, and we're really looking forward to doing that. But we don't -- we want to make sure that we crawl before we walk and walk before we run.

MR. DAVENPORT: Excuse me. What hardware are you planning to use for your conversion?

MR. CUMMINGS: Michael, do you want to take that?

MR. FISHER: If you're talking about --

THE REPORTER: Sir, could you come up to the front here if you're going to speak, because I'm having difficulty picking you up on the microphones. Thank you.

MR. FISHER: If you want to get into specific components, the image base records management systems in terms of what scanners are we using -- you know, RICO scanners, IS 520's -- you know, what are we using for oversized scanners -- scan graphics, we're using them for oversized -- that sort of thing. And then we could get into a list of --

MR. DAVENPORT: So a lot of different hardware.

MR. FISHER: -- that -- items. Yeah, you know, we're trying to make it as plug-and-play as possible because of -- one of the things that we're finding, and a lot of independent consultants are verifying some of this -- so some of the folks from DOE and from the M&O recently went to a document management conference sponsored by the Delfi

Consulting Group. And one of the things that they said with respect to getting into imaging, et cetera, is that what you have to do is, you have to go and get basically the best of breed that you can. Go and get the best OCR that you can today. Go and get the best scanner that you can today and try to get those integrated. Because technology is changing so fast and -- you know, you have to --

MR. DAVENPORT: A number of years ago when we were doing some work at the then system, we had discovered that one of the problems was that the people who were doing the field definition and the identification of document -- placing in fields, were the least sophisticated people in the system and they were frequently making low-grade choices. Now, in your imagining, you're going to be getting into verification that the image document is accurate to the microfilm or hard copy document.

MR. FISHER: Right.

MR. DAVENPORT: Have you got a plan to guarantee that the people who are going to be involved in that quality review are of sufficient competency and knowledge about the program that they're going to be making high-grade decisions there?

MR. CUMMINGS: Yes, sir, I am. Because we're also talking about the same people who are QA trained today. I mean, they operate in a quality assurance environment today,

and so we have very stringent procedures and I think -- you know, certainly high quality personnel that are following those procedures. I have a lot of confidence that that will be a very good process.

MR. BALCOM: So, you're also in -- as part of your reprocessing and re-indexing into a search and retrieval database now, is this the VAX Basis Plus kind of interim step for you or are you just capturing the images?

MR. CUMMINGS: Let's see. How do I answer that?

(Pause.)

MR. DAVENPORT: Well, I kind of --

MR. BALCOM: Then you asked about hardware, so --

MR. CUMMINGS: You know, right now -- I mean, our database is VAX Basis Plus oriented. We don't know whether it will always be that way or not, but whatever way it is, we -- it will be an electronic transfer. I mean, we're certainly not going to go through any reprocessing again.

MR. BALCOM: Right, but you can go from microfilmed image to electronic image without re-indexing -- without doing any indexing of the data. My question is: Are you also going to -- as part of your next two-year plan, also build an index database of that newly captured data?

MR. CUMMINGS: Yes.

MR. FISHER: Within the pilot --

MR. BALCOM: That's Basis Plus?

MR. CUMMINGS: No.

MR. FISHER: No. If I could in part answer some more about your previous question, to -- the commonality to some of this, we are on a VAX platform as far as the server. It's client server technology. We're on our VAX platform. We're on IBM PC-compatible clients. It's the same architecture in the records management -- the image-based records management system -- as it is in the Info Stream system, thus the long-term integration issue being somewhat less. The IRIS system is VAX-based but it uses Basis Plus for its indexing tool, which is sort of a combination text management, database management system.

What we're doing in the pilot for the records management system -- the image-based is, they take the paper, they actually do the indexing into IRIS and then bring the -- that batch over to a scanner. When the operator starts scanning it, they look up that batch to get the batch numbers and the session numbers -- what's called the session numbers -- right out of IRIS, which has already been indexed. And then we go in and actually pull some of the -- right now, a subset of those fields out. And that was -- the basis of those was part of the application of IRIS, a quick indexing, quick retrieval feature. So, we don't pull all the fields. It's like the most commonly

retrieved fields.

I'm sorry, I don't mean to keep my back to people over here. I'm just trying to address your questions.

And that is put into an internal database to the records management system. But you can retrieve -- you can go back to IRIS and retrieve on any field.

MR. BALCOM: How about the full text of the data?

MR. FISHER: In the pilot, we are doing OCR and doing some full-text retrievals.

MR. BALCOM: Which engine are you using for that?

MR. FISHER: OCR right now -- we've actually integrated over time two different ones, and we've integrated Exposition Typereader first, and then we -- which we're not currently using, and we're currently using Xerox Text Bridge. Again, over time, that may change.

Going back to the plug-in nature, I know OCR technology -- I'm sure Doctor Nartker will be speaking on it at some point -- gets better each year. And maybe this year in the pilot, Xerox Text Bridge is better, but maybe by the fourth quarter when we're just about ready and we hope to go operational with this thing, maybe it's going to be somebody else's product and maybe there is a decision at the 11th hour that we really want to go with another product. So, we're trying to keep it modular from that aspect. But long term, the idea would certainly be to integrate the IRIS

system with the records management system, whatever you want to call that -- and if the Info Stream system stays around, integrate it with the Info Stream system and get those multiple databases down to one. That's the plan.

MR. GANDI: That is the plan, but specific dates and times of delivery haven't been put together. We're still looking at the pilot phase of our RDMS into February, at that time, looking at the software -- the particular software packages in general and seeing how they all fit into our process of how we do records. That's one area we need to address before we proceed with this rapid scanning of backlog documents. I think what you'll see is, during this process we may be changing software packages after all the bugs are worked out. As time proceeds, we'll be looking at new packages and what they can buy us. But the important part for us right now, we feel, is to get the process down.

MR. CUMMINGS: Yeah. We obviously want to go to the electronic imaging and we want to get very good at operating that way and then if they -- you know, whatever system we do land on, and that's the way we operate, then we want to reprocess our records. We certainly don't want to reprocess twice.

Yes, sir.

MR. METTAM: I've got two questions for you. The first one is fairly basic. One of the things about Info

Streams that most interests -- sort of elegant in capturing it electronically, since it was created electronically -- I assume that nobody is still typing these documents. and I -- it sounds like we've lost that a little bit. I realize that the imaging effort has to have occurred, because many of those documents were not captured when they were created and there's just this volume of material. But at least in my mind I saw future documents being captured electronically rather than that opportunity passing and then capturing them optically down the road.

MR. GANDI: We've talked this over, too, and a lot of that is process controllable as far as text. There's another question I think we need to answer such as electronic signature. The document is going to have to be scanned anyway with the signature on it. As far as the text file being verified, or the process of capturing that text file that was created to make that original, I don't see why we can't process -- load that through the system where we've captured that text. If I understand our records management system right now, the one we're in pilot phase, it takes that scanned document and runs it through the OCR patch process.

UNIDENTIFIED SPEAKER: Right.

MR. METTAM: But that process of doing the optical character recognition injects some inherit errors, and

if -- you know, the issue is basically a control issue. If you get a control document electronically, or on paper, it really doesn't matter if it's truly been controlled. You know, it seems to me that it makes a lot of sense to use the electronic media where it's available and properly controlled, rather than sort of -- you know, shoving paper through a scanner for the rest of your life.

MR. GANDI: I agree. I think how we process this 500,000 backlog we've got is going to become -- we're hoping that we're looking far enough ahead to -- like I say, capture the process. And that's being able to automatically capture that text file so there's no more verification process that has to go through to assure accuracy of the text file versus the scanned image.

MR. METTAM: Right. It would seem to be easier to verify an electronic copy against a paper copy if you had to -- you know, if your controls weren't there electronically -- than it would be to take it and scan it and do character recognition and then look for errors. It would be a lot faster just to verify and control it. That was my first comment.

The second one is, how does all this relate to something called Licensing Data Management System?

MR. CUMMINGS: Okay --

MR. GANDI: That's what I was referring to.

That's the system that we're evaluating right now and I referred to it as RDMS.

MR. CUMMINGS: Right. It's had a name change. Basically -- I mean, that is our electronic imaging system. I mean, that's the one -- and it -- like I said, it's going through testing and evaluation now and if everything goes well, then we hope to be on track somewhere around May or so, of starting to put that into operation. But that is exactly what that system is.

MR. GANDI: All right. I'll try to get close to -- I want to clarify one thing here and that is, that from a pure records management perspective versus this pilot of RDMS, Records Data Management System, they don't require full text search to go operational. And one of the things that could happen coming out of the pilot -- and I bring this up only because in the eventuality that the decision is made, "Well, gee. We don't need full text search right now. That's really an LSS functionality that would be required at some point." And there may be a decision in the Records Data Management System when they do go operational to not implement full text search and just have an index retrieving. I just wanted to clarify that so in the future, you know, "Hey. We thought we were going to get full text search in that system. It may not be there, but the plan is, we're trying."

MR. CUMMINGS: RDMS is really our way in records management to replace our microfilming system. Now, if we get full text, boy, that is going to be wonderful. We don't have a hard requirement for it. I just -- we just -- I do know in one of the recent litigation support cases we had, it sure would have been nice, you know, instead of going through the indexes that we did in our database and then matching it up with microfilm and so forth. It definitely would be nice to have, but our primary goal is to replace our microfilming operation.

Any more questions?

(Pause.)

MR. CUMMINGS: Thank you.

MR. DICKERSON: You're asking "Where are we? The answer is, we've completed phase one.

MR. HOYLE: Can you identify yourself, sir?

MR. DICKERSON: I'm Fielden Dickerson, and we've gotten through the background and we're going to spend the next 160 minutes talking about what the working group did. And you'll notice that we also have scheduled into the break too, so there is a rainbow in that cloud.

Now, Claudia had shown you essentially these points as the charter for the working group, and what I'm going to talk about up front is reviewing the historical background and identifying the requirements for the LSS. In

terms of doing this, we dug out the negotiating record. Jerry Salzman gave me sort of three linear feet of his files that I looked at. He talked to many of you in -- as historical figures, collected files -- we've looked at supplementary information that came out with the final rule and looked at the rule in detail.

Now, in terms of that, we're not going to be able to tell you everything that we learned about the requirements for the LSS in terms of expectations and commitments and so on. What I'm going to do instead is just give you a skeleton, a backbone for the LSS so that we can move on down through this in terms of identifying operational concepts, cost drivers and then look at the projected lifecycle costs associated with the LSS. And that's what George Hallnor is going to do right behind me here.

As I indicated, I'm Fielden Dickerson. I'm going to be talking about the requirements as defined in (j) and those are very much skeletal requirements. And the LSS is an information management system -- it's an electronic information management system. And there's some words that -- having to do -- about access, which indicate that there has to be some sort of remote facilities for getting to it. And then, finally, it says that there's something about communications that have to be associated with it. So

it's the information management system, remote access -- and has a communications capability.

As you move along through (j), you get into Part 1003 and you start to get some more definition of what's in that. In that, it tells us that we're going to have text, we're going to have image and we're going to have headers. And if you read this 1003 in detail, it has some disclaimers in there for material that's graphical of nature so that you put in headers and if there's other things that they consider inappropriate, then you can simply put in descripters saying where the material is. And they describe the text, the image and the bibliographic headers that go into that.

And then we move on down in this and they start talking about access, and they break access into two categories. One is public access and one's access for the participants, the parties or potential parties to the proceeding. And in that, they start to identify some of the remote facilities, and one can count them in terms of that. And the access is defined -- you know, prelicense application submittal of a particular character for the public, after a license application of another character.

MR. MURPHY: When did you come up with those remote access --

MR. DICKERSON: I'm sorry, what?

MR. MURPHY: Where did you come up with the locations for the remote access?

MR. DICKERSON: They're out of (j).

MR. MURPHY: (J) doesn't just mention Lincoln County and ignore Nye and Inyo county.

MR. DICKERSON: Yes, it does.

MR. MURPHY: No, it doesn't.

MR. CAMERON: Well, it ignores Inyo but it doesn't ignore --

MR. DICKERSON: This MAL is right out of (j). I'll show it to you.

MR. MURPHY: I'd be -- well, you better show it to me.

MR. DICKERSON: You know, we'll have a break, I'll -- I have it in my briefcase.

MR. MURPHY: Well, that -- I think that's wrong, but in any case, you're going to have to change it -- your remote locations are going to have to go in all of the effected units and local Governments at a minimum.

MR. DICKERSON: I'm telling you I came out of (j).

MR. MURPHY: Well, I'm telling you what you're going to have to do, at a minimum.

(Pause.)

MR. DICKERSON: I did not find that written down anywhere.

MR. CAMERON: The intent of the rule was for the access locations to be --

MR. METTAM: Yeah, this is not what --

MS. NEWBURY: I think Fielden is just showing you what he pulled out of (j). That is not necessarily where we're going to put remote locations.

MR. MURPHY: Let's not worry about it. That's just wrong.

MR. DICKERSON: Let me read you something that was a recommendation of the working group now, but put your mind a little bit at rest in terms of this. We had gone through the requirements of this thing, as I say, trying to identify commitments, requirements, expectations and so on, and one point we came --

MR. MURPHY: Well, let's -- let's not spend any time on this. That's just flat wrong. You're going to have to put a terminal in at least Tonopah and Nye County. You're going to put a terminal in Inyo County, you're going to put a terminal in Esmerelda County, you're going to put a terminal in maybe other areas where Nye County has contractors such as my office in Olympia, Washington. Let's just -- and Claudia understands that. Let's just not spend a lot of time on it.

MR. DICKERSON: Access is defined in particular fashion, and this is for the participants. And after the

licensed application is submitted, this is also the form of access that the public will have.

Now, these were the requirements -- the skeleton, if you will, that we put together, in which we started talking about options for the LSS. We were not designing an LSS, we were looking at conceptual options so that we could go in and do a cost sensitivity analysis associated with it, and out of this we wound up with this sort of model in which we identified hex dissemination. That's the whole text that we have there, and that is simply electronically transmitted to the remote stations by a network.

Imagine dissemination we broke into two categories. One was that of an electronic transmission and the other was something else. And that you might be talking about mail or however it gets there. And you'll notice under electronics, we also had CD-ROM as well as just downloading of the text from the wire. And out of this we built six options, and you can mix and match off of these things and -- but what we've done is, up here on the text dissemination that was coming back -- the question had been raised earlier in terms of the OCR and human verification. No for half the options, yes for half the options. Then in terms of network for the image dissemination, you see the options five and six and so on. And these are the six options that we wound up handing over to George, who is

going to be describing costing to you.

Now, there have been three costing activities associated with the LSS -- at least three that have been documented. One in 1989, and it was predicate -- this is the original one that was done and was based on the technology available at the time. And the difficulty was that it didn't give us a separation between records management costs and LSS costs.

And then in 1992, there was a calculation directed in specific issue of using record management technologies to support the LSS. However, that wasn't looking at cost sensitivities, it was simply saying "Let's do this and see what the cost result is." And what George is going to be showing you up here is making use of these six options that we've described for you. He's going to be demonstrating sensitivities and trade-offs for those various options and giving us -- you know, some view as to what we might -- how we might make decisions based on that.

Thank you. George Hallner.

George?

MR. HALLNER: Like Fieldsen said, I will go through the cost analysis, and more specifically, the models that we used to arrive at that. First up front, there was a number of assumptions made, mostly based on historical data. The fact that 85 percent of LSS holdings are anticipated to

coming from DOE and the participants in the OCRWM program, the remaining 15 percent of material will be submitted directly to the LSS -- to the NRC, which was my assumption in saying that LSS is a cost element separate and apart from the records management system that DOE maintains.

The -- okay, the records holding in the records management system will be supplied to the LSS in a form of copies. The records management's responsibility for DOE and participant records remain with DOE. That's an important thing, because there are certain costs associated with maintaining the records, so LSS would have a bona fide copy of that.

All dissemination of information will be done directly from the LSS. There will, of course, be the Freedom of Information Act requests coming to DOE and there may be also requests coming based on findings in material and LSS and governing the record -- to find clarification for records that are kept by DOE. Case in point may be proprietary records for contracts initial -- and stuff like that that would be taken by DOE.

In terms of facilities -- Now, I want to say that this is a model assumption. This is not going to be viewed as: this is the way it's going to be. We see that being two DOE server facilities. One in Dun Loring, which is the M&O facility in Vienna and one in Las Vegas, in terms of -- and

these are the records managements facilities. There will be one LSS server facility. The server facilities -- all the data stored -- and this is the facility where the request for information comes in, and that one will be located in Las Vegas. There will be a total of three -- sorry, four CAPTA (phonetic) facilities associated with these systems. DOE will have three for its records management system; one at DOE headquarters, one at Dun Loring M&O and one at the M&O in Las Vegas. The LSS will also have to have a CAPTA facility before -- because of materials submitted directly to the LSS, and that, we anticipate to be in Las Vegas.

Also, the model assumes nine public access facilities with up to three terminals or connection points, if you want at each, and they will be located in the Washington, D.C. area, and in the state of Nevada -- Las Vegas.

MR. DAVENPORT: Excuse me.

MR. HALLNER: Yes.

MR. DAVENPORT: May I go to your capture facilities?

MR. HALLNER: Yes.

(Pause.)

MR. DAVENPORT: Do I presume then that the LSS Las Vegas capture facility is the facility that we would -- the

state or any other potential parties would bring their materials for --

MR. HALLNER: That's the -- that's --

MR. DAVENPORT: -- delivery into the system?

MR. HALLNER: And it may not end up in Las Vegas. It maybe end up somewhere else, but for the model purposes, we use that as a -- as our --

MR. GANDI: I think the important thing is to know the mere existence. These are the assumptions used for the cost model itself.

MR. DAVENPORT: Yes, I understand that.

MR. GANDI: Okay. A lot of things could change.

MR. MURPHY: What about those nine public access --

MR. HALLNER: This is just based on account more or less from the list that Fielden had, and -- but normally, it should be taken that that is going to be it. Because -- facility turns out the relative inexpensive item. So it's --

MS. NEWBURY: Now, as I mentioned in my introduction, the number of stations can change and we want to reevaluate the cost model based on the change in the number of stations.

MR. HALLNER: That's a total of 27 --

MR. MURPHY: Yeah, I -- no, I understand that.

But I guess I didn't complete my question. Are you talking -- is that nine non-DOE access facilities or nine facilities to which the public --

MR. HALLNER: These are the nine public locations.

MR. MURPHY: But not -- you know, not participants, but are there going to be nine facilities where Joe Blow could walk in off the street --

MR. HALLNER: Local libraries -- you know, whatever.

MR. MURPHY: Yeah. All right.

MR. HALLNER: Okay. When I set up the model, I did to the extent possible, use real numbers in terms of what time does it take to index? What's the labor cost of an index? Or what is the cost of transmission times and so -- transmission -- and so forth. And using the data we had available, we came up with a -- three areas where the system has particular sensitivity. One of them is clearly the data volume that would be in the records management system and consequently LSS. And specifically, they are the intake of data -- the capturing of data is a cross driver that is directly related to the volume of pages being put into the system. And we did some cost analysis of that that I will look at a little bit later.

The other one is correction. Human corrected text versus OCR accuracy. The OCR can be augmented with some

intelligence, if you want, so there is some minimal type of mechanical correction that can occur. Somebody said that whenever a human gets involved and actually does text correction, that is a -- human corrected text here. And we defined options both with and without the text correction in there, and we looked also to see what the sensitivity of the OCR accuracy had -- what effect that was on the cost. And the other one was a dissemination cost for the LSS with varying volumes of pages disseminated over the life of the system. And we did two analyses: one assuming a mix of 80 percent text and 20 percent image versus 20 percent text and 80 percent image, to see what the trade-offs were with varying volumes. And I'll show you that.

MR. CAMERON: How do you define dissemination there, George?

MR. HALLNER: This is -- if I go out to query -- and I said, "You got to resolve this" and I said, "I want to see this document or these pages" -- is the number of pages that you actually bring back in electronic form or as printed material. Something that you can see as --

MR. CAMERON: For any number of --

MR. HALLNER: -- for research purposes.

MR. CAMERON: -- participants?

MR. HALLNER: Any number. Yes. This was a --

those -- a number -- a million pages per year on the average, regardless of who asked for it, okay?

To evaluate the extreme values here in terms of volume, we worked with a number of what we assumed to be LSS-relevant pages, and we have three cursor, and Jim Boone will talk about them a little bit more detail later. The lower curve is based on -- is what we historically achieved until fiscal '95 and then assuming a 10-percent gross rate as being what is submitted. And that curve, I think, is probably unrealistic, but it gives a little bound of numbered documents that we will see in the LSS.

The upper-most curve is -- I constructed out of thin air, sort of, because I said, "We have spent so many millions of dollars on this program today, we have produced this many pages, so they cost X dollars per page. We gonna have projection spending this many dollars, so this is the number of pages we will generate." That is probably a little --

MR. BALCOM: Did you take into account and go back to the old SAIC projection?

MR. HALLNER: Yes, I will mention that.

MR. BALCOM: Okay.

MR. HALLNER: The middle curve here is something that reflects what Jim Boone did in terms of the polled knowledge of an individual -- said "What do you think the

volume of pages submitted to the LSS by year will be?" and he came up with this curve. And as he will show you later on, that curve falls very closely to what SAIC assumes to be the case. So, I think the opinion of the experts have not changed that dramatically.

MR. BALCOM: Except there is, I think -- we were up around 30 million pages.

MR. HALLNER: Ah, but they weren't further. They went to 2010 I believe.

MR. BALCOM: Oh, okay. All right.

MR. HALLNER: I think if you draw out the line there, the lower level -- you end up with about the same level.

MR. BALCOM: Okay.

MR. HALLNER: Then, looking at the -- oh, I see option one shows up in this. That option one was something we used as a reference point. That was a -- option that is not in agreement with the (j), because it assumes a microfilm-based LSS. It was a cross reference for us to say that cost model was accurate anyway, so options two through seven here shows the allocation of costs. The lower part of the figure there shows the DOE record system cost. Those are the costs that would be borne by DOE to maintain their record system which would include all material that is LSS bound and some other records.

MR. CAMERON: Now, that would be a cost that would have to be assumed, whether --

MR. HALLNER: We have it --

MR. CAMERON: -- it was LSS or not.

MR. HALLNER: -- the other way. Right.

MR. CAMERON: So, anything above that is the only -- the incremental cost for the LSS.

MR. HALLNER: Right. We'll get to what incremental cost is, okay?

(Pause.)

MR. BALCOM: Same as alternative one through seven in the working group?

MR. HALLNER: Yes. Yes. Yeah. The next layer under each pile here is what I call the implementation costs, and that's the implementation cost as it relates to RMS. We have assumed that -- in this study, that any components of the LSS would also be components of the RMS, so this is essentially as to hardware costs associated with the LSS. So, that is really the DOE implementation costs from the various options here.

Then we have a black bar going through some of these, and that is the cost that has to be borne by DOE in that record system to do the additional processing required under that option to get the data into a form that can be exported to the LSS. And some of that may entail, for

example, text correction. The big black areas there was text correction, because the RMS does not need full text. So, that would be a cost borne by the --

MR. CAMERON: Would that include the use of the thesaurus and --

MR. HALLNER: Whatever that means at that time.

MR. CAMERON: -- it would be filling in the headers and --

MR. HALLNER: Right. Right.

MR. CAMERON: -- all that sort of stuff. Yes. Whatever the technology and the implementation is at that time. And at top level, the stripe with -- bars there indicates the actual cost of the LSS, which is the procurement cost, insulation cost, the operations cost and dissemination cost. Everything you have involved in this. Looking at specific figures, the LSS specific costs for the various options are shown here. And as you can see, they're not that much different. That being -- the reason for that is, that most of the data is digested by the RMS into an LSS-compatible format in the beginning, so you wouldn't expect very much difference between different options here.

I will also go through some -- what I call supporting information, to give you some more detailed insight here on some items. We looked at -- for option six, which is the one we seem to have been most agreeable that

that would be an RMS LSS-based system. The cost by fiscal year looks roughly like this. What you see in '98 and '99 is the actual procurement of the LSS and the loading of the LSS from the RMS, and then you get down to an operation condition from then on.

The RMS cost profiled by fiscal year looks something like this: And we see number of very large values in '96 through '98, and that is the installation of all the components for the records management system and also the reprocessing costs that we talked about here earlier today. And then we go down to an operational level.

(Pause.)

MR. SILBERG: These operational costs include not just the hardware and --

MR. HALLNER: Oh, this is --

MR. SILBERG: -- inputting all the --

MR. HALLNER: -- inputting, everything is -- this is to get the material into a system, into a form that can be used. But, as I said, this is a model, so we -- you know, we learning as we go along here, too. Also looked at the sensitivity and page volume, and here is where I used the upper or middle range and the lower cost -- lower number of volumes coming in. As you can see, not unexpectedly, the higher the volume for each option, the higher the cost. But as you see, the difference between the low and the high

value -- that is by option, and specifically, the scenario -- the options -- scenarios not connect -- the options where there is human corrected text has the biggest variance between the high and the low, which is not to be that unexpected.

MR. SILBERG: Are the LSS cost increases or sensitivities roughly the same as these?

MR. HALLNER: Oh, yeah. Well, this is the capture part. So that is all to deal with the original input coming in so that is really relating to 15 percent of the volumes. I think that if you scale this to -- so that it would correspond to only 15 percent of the volume or the total holdings, you would get something similar, yes.

MR. SILBERG: Now, wait. This --

MR. HALLNER: This is the RMS cost, which is 85 percent of the --

MR. SILBERG: -- is 85 percent.

MR. HALLNER: It is 85.

MR. SILBERG: Not 15 percent.

MR. HALLNER: The LSS is similar in shape when it was, of course, scaled appropriately.

Sensitivity to OCR. What I did there, I took a comparison between option seven which does not include human correction, and -- with option -- sorry, option six, which does not include human correction, with option five, which

does include human correction. And as you can see, as the OCR accuracy goes up, there's going to be less correction -- down and the cost goes down. It -- at the 94, 95 percent you can see there is tremendous cost penalty to do correction by human means in the sense of having someone look on the screen for suspected words, errors -- and do the manual correction. But as the OCR is moving up towards the 97 and above percentage points, the differences get smaller and smaller, and hopefully the -- get less.

In terms of dissemination services, the first -- okay, look at distributing the requested tables; 20-percent image and 80-percent text. And we look at various volume like 30 million distributed over the life of the system, 60 million, 240 million, 480 million.

What is the right amount that's distributed? To be very frank, I don't know. I think it's probably towards the higher end than the lower end here, but that's my own opinion. And as you can see, dissemination cost is not really all that outrageous and also, interestingly, the cost of distributing by CD-ROM is higher for the low volumes than for the high. So, depending on which way you're looking -- sorry, I'm looking in the wrong -- sorry, I made a mistake there. The cost of CD-ROM is always going to higher the text, because it's less efficient for text.

And then looking at the last one, the

dissemination -- 80-percent image, 20-percent text, same volumes. You'll see there's a flip between 60 and 240 million, but CD-ROM gets more cost effective because you have use of image there. And distribution of the image with telecommunication we discussed is very expensive, quickly. However, still, the costs are not really all that outrageous considering the volumes.

MR. CAMERON: Could you just -- I guess I'm still trying to get a handle on the dissemination, the 80-percent image or 20-percent image versus text. Are you talking about the -- whether the participants using the LSS are going to be --

MR. HALLNER: Our vision -- to -- our vision is that the participant will be able to request text or image.

MR. CAMERON: When you say "request text", for example, if you put it in terms of the -- someone on the staff of a participant -- okay, sitting down at a terminal and accessing the LSS, when they do a full text search and -- when does the dissemination start? When they request a download of the electronic text? I've been trying to figure out what you mean by --

MR. HALLNER: Okay. Dissemination in this sense starts when you have -- you've done your text search, you've got your list of hits.

MR. CAMERON: Yeah.

MR. HALLNER: You said "I'm interested in these 10 different things here. Give me the text or give me the image for those things." That's why you have the --

MR. CAMERON: Give me the text on my screen or give me -- does it matter whether --

MR. HALLNER: It doesn't matter.

MR. CAMERON: -- it's on my screen --

MR. HALLNER: You can direct --

MR. CAMERON: -- or whether you download it?

MR. HALLNER: -- the printer, you can look at -- on the screen or you can put it onto a mass storage media of your own.

MR. CAMERON: And the same thing with image?

MR. HALLNER: That's correct.

MR. CAMERON: Okay. And on image, do you assume that it's an electronic image or -- I guess that's what you're talking about CD --

MR. HALLNER: Yes, electronic image. Yeah.

MR. CAMERON: Oh, okay. Thank you.

(Pause.)

MR. BRADKON: Can I ask a question? My question --

MR. HOYLE: I'm sorry, sir, you're going to have to come up closer to the table to get this question on -- thank you.

MR. BRADKON: No problem. Steve Bradkon with a very quick question on the capture of header information and key words. In your costing, did you assume that the Department of Energy would capture full LSS header information and key words at the time they captured their --

MR. HALLNER: Yes. Yes. Yes.

MR. BRADKON: Okay.

MR. HALLNER: That's spending about 25 minutes per document doing so.

MR. BRADKON: I understand.

MR. LEVIN: Could you go back to the chart you had on LSS cost by fiscal year and explain that a little bit now? I think you said that this was the cost of loading the LSS --

MR. HALLNER: That -- in '95, '96, '97, that's more or less just getting ready for preparing facility areas and getting all the things in place. '98 will be a year the LSS is installed and loading starts. We're loading in this transfer of data from the RMS.

MR. LEVIN: So, this -- in 1998, this 10 million dollars is basically the cost of transferring the data from RMS to --

MR. HALLNER: No, it's also equipment --

MR. LEVIN: Oh, okay.

MR. HALLNER: -- you know, getting the right managers in place, getting the facilities, getting --

MR. LEVIN: Okay, so this wasn't just the cost --

MR. HALLNER: No, no. This is totals.

MR. LEVIN: Okay, do you have any projections from the cost of just transferring the cost from RMS to LSS?

MR. HALLNER: Yes, I have that. I don't have the absolute figure right off the top of my head, but it's going to be low, because the RMS is digesting the data in a form that's going to be -- you know, compatible with the LSS. So, essentially, it relates to -- there's virtually no cost in doing the electronic transfer, if you want. It's a matter of how much quality control do you want on your side? Do you trust us or don't you? That's the issue.

MR. LEVIN: Do you have -- and can I get and of the figures for the incremental cost of having the LSS as a separate system, versus if LSS and RMS could be one system? Do you have those numbers and incremental costs?

MR. HALLNER: I think the components are there and I could probably figure it out. I mean, I --

MR. LEVIN: I'd be very interested.

MR. HALLNER: -- I think the components are there.

MR. LEVIN: If you could, I'd be very interested in those numbers.

MR. HALLNER: Yeah. Yeah.

MR. CAMERON: Just one clarification, George, on the relationship between your total option cost graph and your RMS cost by fiscal year.

MR. HALLNER: Right.

MR. CAMERON: Looking at the total option cost graph, what components of that make up the RMS cost by --

MR. HALLNER: That is up through the black line.

MR. CAMERON: Up through the black line.

MR. HALLNER: Yeah.

MR. CAMERON: Okay.

(Pause.)

MR. BALCOM: George, I had a question about the 25 minutes per document that you're using for your cost models.

MR. HALLNER: Right.

MR. BALCOM: What does that include?

MR. HALLNER: That is flipping through it, deciding -- Marty can describe this better, but I'll --

MR. CUMMINGS: It's a screening and the indexing and the imagining.

MR. BALCOM: When you say "indexing", what do you mean? What do you include in indexing?

MR. CUMMINGS: The header fields that --

MR. BALCOM: Okay, including all the subject terms related to that document?

MR. CUMMINGS: Yes, sir. Yes, sir.

MR. BALCOM: And this is a 30-page document and you have an elaborate way -- once again, are you planning to use a thesaurus for that or what kind of control techniques are --

MR. CUMMINGS: Yes, sir. We do have a thesaurus, we do use it, it is the cataloging and indexing.

MR. BALCOM: Is it anything like the old thesaurus that we saw during the SAIC studies, which was a couple of inches thick?

MR. CUMMINGS: Yes, sir.

MR. BALCOM: Okay.

MR. CUMMINGS: It was derived from that.

MR. BALCOM: All right, I see Jan shaking her head back there. All right, thanks.

MR. GANDI: I think one of the questions that this brings up is, how many header fields do we need if we're going to have a full text information base? The elaborate 25 minutes of index -- of going through a document to pick out key terms is such that it seems to be lost when you have the capability of scan -- or searching the full text.

MR. BALCOM: Well, that's an old, old question --

MR. GANDI: Yeah.

MR. BALCOM: -- that we've gone over many times here and anybody who builds one of these systems also -- I don't know. It's set up for --

MR. GANDI: I think in the future --

MR. DAVENPORT: You have two charts here that I wish you would integrate for me. The one that's called "LSS Cost by Fiscal Year", will you show me where those numbers are on your graph called "Total Option Costs"?

MR. HALLNER: Okay, the LSS cost for fiscal year is the cost portion on option six, the uppermost section on that bar.

MR. DAVENPORT: So it would be six -- it would be the uppermost section.

MR. HALLNER: Yeah.

MR. DAVENPORT: Okay, so that's the option that they're -- the cost.

(Pause.)

MS. NEWBURY: John, did you want to take a break now?

MR. HOYLE: Is this a good time for break?

MS. NEWBURY: That sounds good.

MR. HOYLE: We're on break. Fifteen minutes.

(Whereupon, a brief recess was taken.)

MR. HOYLE: Okay, Marty, I think we're ready to begin phase two.

MS. NEWBURY: Okay, John, before we get into the next presentation, I wondered if Jim Boone could get up and kind of give us a reminder of what the differences are

between images and text and why we really care about an 80-percent image based system, just to help clarify.

Jim, do you want to say something?

MR. BOONE: Sure. Claudia brought up a good point, that maybe not everybody in the audience is technically astute here when it comes to these computer terms. But when the working group talks about text dissemination, what we're really talking about is the dissemination of ASCII text, if you will, in its current form, but letters going across a serial line to whatever work station you're working on -- you could read those letters on the screen. The current form is -- ASCII is very common these days. So, when you say text dissemination, it's strictly ASCII.

Image dissemination can take several different forms, but what you're distributing is a visual image, a visual likeness of the original document. That could take the form of a bit map image, it could be a post-script file. There are various ways of displaying a visual image. But the bottom line is, it's something that you would see on the screen or that could be called up and viewed on the screen.

MR. CAMERON: Is that really the distinction, though, between the text and image dissemination here?

MS. NEWBURY: Well, I think the -- what I'm trying -- the point I wanted to really make is, that there's

a lot of image-based information that will be in the LSS that is not available in text format. In other words --

MR. CAMERON: Graphs --

MS. NEWBURY: -- when you're scanning all this --

MR. BOONE: Maps and drawings.

MR. CAMERON: -- et cetera, et cetera. Right.

MS. NEWBURY: Maps, drawings, engineer drawings --

MR. BOONE: Absolutely.

MS. NEWBURY: -- that are going to take a large volume of electronic bits to get across the line and will take a long time to disseminate in that form.

MR. CAMERON: And that's a big cost factor, isn't it? Transmitting images --

MS. NEWBURY: Yes.

MR. BOONE: Yes.

MR. CAMERON: -- electronically, versus CD-ROM or --

MS. NEWBURY: Correct.

MR. CAMERON: -- overnight mail, whatever.

MR. BOONE: If an image is captured as a bit map, it takes a lot of computer memory to store that image and that has to be transmitted across a communication line. So the expense involved with transmitting these images is very large compared to just straight text -- ASCII text.

MR. GANDI: Another point to be made on that text -- I think at one time there was a spoken requirement to do a cut-and-paste type process from these files such as including an interrogatory or something. To do that you have to transfer the text and I think that was another one that was being addressed with the text transfer.

MS. NEWBURY: That's right. You can't -- if you just have an image of a document -- a written document or typed document, you can't cut and paste like you would if you had the full ASCII file. So lawyers like to be able to cut and paste. It saves them a lot of work, and we like to help them.

MR. CAMERON: So you think that's how we do our work?

(Pause.)

MR. DICKERSON: The next item we have here is the projected LSS schedules. I'm going to show you a couple of schedules for the LSS. Backing up, we began the LSS critical manufacture, if you will, by starting at the back end of the process and asking what steps had to be taken to arrive at licensing application submittal in 2001. And so I've placed that right smack in the middle of 2001 and then backed up to the requirement that the administrator has to certify the loading of the LSS six months prior to that license application. And you heard this morning from Marty

and -- in terms of the records reprocessing as being a lengthy process, and so I've shown that in there, to show you some of our deliberations that are going into this building and LSS schedule.

And in addition, I have noted some things down below in terms of milestones that are going to be associated with the LSS, and then the funding profile in terms of -- one can talk about required funding profile or you can talk about possible funding profile, depending on what might be available, how rapidly one can respond to infusion of funds. And George showed you these two annual cost graphs for the RMS and the LSS, and this was a -- process. He didn't arrive at these annual costs without some sort of structured view as to what a possible schedule might be for the LSS. And we played off costing versus schedule and so what you're seeing is the final product here in terms of -- this is what -- yes, sir?

(Pause.)

MR. METTAM: Is this option six that we're looking at again?

MR. DICKERSON: Yes, uh-huh. That's right. And in the schedule, there are a number of sub elements that we can talk about and this is one I've just, just pulled up because it -- one that people have a good bit of concern about in terms of the steps that you have to go through

for -- to make an analysis. And I've just put "months" on there to demonstrate that something that stretches out becomes a component part of the scheduling process. And there are a number of those that are built into it.

This happens to be one in which we arrived -- said, "Let's assume that the build/buy" decision is that of a build and we wind up with this sort of schedule. What you've seen already is the records reprocessing down here. You've seen the bullets out here in terms of "Where in the world is certification?" and the license application. And then, I also put in some review periods for the administrator for this, and you see the flow-down from the regulatory requirements that -- the system requirements on down through here, and this is what's in your package that you have. This is build.

Yes, sir.

MR. CAMERON: Fielden, where would -- at what point on the build schedule would the LSS be, quote, "operational", unquote, which I would use as the -- trying to use as the point when access would be available to LSS participants?

MR. DICKERSON: Well, there are a number of us that might answer that. What basically you're doing down here in terms of development is that -- at least the picture I have in my mind shift, is that during that development

there's going to be access and iteration with the potential users.

Did I say that properly, Camille?

MS. KERRIGAN: Yeah. On that schedule that he shows there -- on the schedule that's shown up there, you can see the LSS development schedule takes place during '97 and '98 and depending -- it's hard for me to answer that question because a lot depends on the selection of the process that they're going to use for development. And based on that, that could be an iterative process where documents are loaded and then the users get a chance to use the system and determine whether or not they want to have modifications made.

MR. CAMERON: So, that could be -- the LSS development bar could be a period of break-in pilot testing -- I don't know what you would want to call it, but --

MS. KERRIGAN: It could be, but you see, right at the very front there is a process where you're going to make a decision on how you are going to actually do the development. Are you going to vend it out to people who might respond to an RFP, in which case they may come up with some very innovative ways of doing development.

MR. DICKERSON: But, we're giving only one answer to your question.

MR. CAMERON: But it --

MR. DICKERSON: For the build --

MR. CAMERON: -- sounds like there's --

MR. DICKERSON: -- we would see iteration going on in here with the users.

(Pause.)

MR. CAMERON: Okay. Because I was just thinking back to one of the points that Stan made and I'm not sure that it was stated right on point or not. But the PALB, the Prelicense Application Licensing Board, is to be set up six months before access to the LSS is to be given. And I suppose even if we're just doing pilot testing, things like that, maybe that won't have to be so formal. I don't know, but I'm just trying to figure out where some of these other things are.

MR. DICKERSON: Yeah. Well, you'd -- ultimately, you're going to ask about -- when is it loaded -- the LSS. And this is backlog that you're dealing with down here, trying to get prepared to finish the loading. So, you're not going to be fully loaded until out there.

MR. CAMERON: It says "System turnover to NRC occurs before the LSS is fully loaded."

MR. DICKERSON: I think that's right. Correct.

MR. CAMERON: And I don't know what assumptions

were built into that.

MR. DICKERSON: We -- you know, that -- they are not very sophisticated assumptions. We were simply trying to give a flow here of responsibility and steps, if you will. Trying to get some understanding of whether the process would fit within the time that was available for it -- for 2001.

MR. CAMERON: And is that -- well, is the bottom bar on the records reprocessing -- is that captured as opposed to the bar that's for load LSS data, which is actually loading it into the search and retrieval system?

MR. DICKERSON: Yeah, it's the process that Marty was talking about in terms of making the conversion from microfilm to digital material.

MR. CAMERON: Okay. So --

MR. DICKERSON: Now, in parallel with this reprocessing, of course, there is the business of processing that is going on for the material that is flowing in in a continuing fashion.

MR. CAMERON: Was is -- is it true that under this chart -- this may be too much of an overview for you to answer this at this point, but you would contemplate the NRC loading the LSS data. I'm not sure what loading LSS --

MR. DICKERSON: Loading, I think, is not a big deal. At least as it has been explained to me, Chip. And,

basically, what you're going to wind up with is tapes or disks or some such thing that is the result of this reprocessing, and then they simply have to be read into the machine in some fashion.

MR. CAMERON: Thanks.

MR. METTAM: Help me with something. When I looked at load LSS data, I was assuming that was the non-DOE data from the earlier chart where we had the LSS entry that -- the 15 percent other data.

MR. DICKERSON: We have all of that encompassed in there in terms of loading, right.

MR. METTAM: So -- okay, so then the -- this bar -- load LSS data includes both DOE data and the --

MR. DICKERSON: Yes. Uh-huh, yes. Tried to get it all in there.

MR. CAMERON: And the bar on the bottom on records reprocessing really includes the capture of all data, not just DOE.

MR. DICKERSON: Well, what we were looking at here was the DOE, and that was the DOE reprocessing. And that comes back to the -- you know, the dollars that are required to do that, and so it was part of a coupling of the records management system in a timing fashion so that you could afford to do the reprocessing on a time scale to meet this.

MR. CAMERON: And should there be a similar bar,

although much shorter, before the turnover to the NRC that is reprocessing of participant data?

MR. DICKERSON: Yes, there should, and that's a very good point.

MR. CAMERON: Okay.

MS. NEWBURY: That's not a DOE function, though. That's something that whoever else is putting things into the LSS will do. It's --

MR. CAMERON: Well, I guess under option three it was going to be a DOE function.

(Pause.)

MS. NEWBURY: We will reprocess all the effected parties' records?

MR. CAMERON: Yeah. In other words, to make -- to utilize the DOE capture system, basically, as the cost efficiency to load the 15 -- not to load -- let me not use that term -- to capture, you have to reprocess here to capture the 15 percent of the other participants' records, including the NRC's, which may not be -- in some cases it may not be too much -- not too much reprocessing may be required. But that's a -- I -- that's a key point in terms of the commission, I think. So, we have to address it --

MR. DICKERSON: Certainly. Yeah, that's an interaction. What we were looking at in terms of this were the things that DOE controlled and that DOE -- things that

DOE was currently responsible for.

MR. CAMERON: Okay.

MR. DICKERSON: Now, that's the build schedule, and --

MR. CAMERON: We've got a question over here, too.

MR. DICKERSON: Oh, sorry.

MR. DAVENPORT: The first question -- all this costing and scheduling appears to be done with the choice of one of these options. Are we presuming that one of these options is the choice? Should we cease to call them options anymore?

MS. NEWBURY: No.

MR. DAVENPORT: Are you costing and scheduling all these options? And comparatively, how do they look?

MS. NEWBURY: They selected this option to show you because it does appear to be the most cost effective, but do -- as I mentioned earlier, DOE would like them to look at some other options as well. So this is just --

MR. DAVENPORT: Okay, so we're still open for options.

MS. NEWBURY: -- as an example. Yes.

MR. DAVENPORT: And the several that you're preferring, I take it, are three and six? Did that come through?

MS. NEWBURY: One and six and what's the other

one?

MR. BALCOM: Three was the original. Well, not the original, but it was -- I think the recommendation coming out of the working --

MR. DAVENPORT: Alternative three?

MR. DICKERSON: But, we're still in an option mode, though. It was the -- mode, and what we were trying to do was demonstrate here, Jim, whether this whole process would fit.

MR. DAVENPORT: Okay, and that gets to my second question. Excuse me. If -- your presumption here -- you choose an end date of 01 and you work back. NRC -- John, if the -- if we discover that this timing gets us to 02, 04, 07, something like that, is NRC prepared to assert the quality of the LSS as a determiner rather than the deadline? I mean, are we -- are we going to be stuck with 01 or are you guys going to have the fortitude to say the -- operating sufficiently enough. We're not ready to flash a green light of go on the licensing procedure.

UNIDENTIFIED SPEAKER: Or we'd go on (g).

MR. LEVIN: We aren't -- I'm not going to certify a system that doesn't meet the -- requirements and the quality requirements.

MR. CAMERON: And the question -- yeah, I think that Moe is very clear on that point. The question

becomes -- is, if you don't have the system that's working or if you can't certify that the documents are in -- remember, the certification requirement and the rule is related not to the system design and development and "is it working," it's related to DOE's and other -- DOE's document compliance. But, say in either -- say either of those things don't happen; the design is not working right, the documents aren't there. Would the Commission say that -- "We don't want to docket this unless the documents are in there and the design is working"? In other words, don't dock it under (j)?

MR. DAVENPORT: That's a political --

MR. CAMERON: With the Commission's docket under (g).

MR. MURPHY: I think chairman -- has made statements now three times on the record.

MR. LEVIN: That's right.

MR. MURPHY: Which would indicate to me that (g) is no longer going to be an option. He has clearly said no LSS, no license application. And I think there is even in one of the documents that we looked at -- you know, the -- we may be transposing the NRC's alternative three with option three. I think we're talking about apples and oranges there.

MR. CUMMINGS: That's correct.

MR. MURPHY: But, one of those alternatives included the -- a proposal to amend 10 CFR (j) to remove that option, or 10 CFR Part II to remove the option that if (j) isn't -- if the LSS isn't up and running, we can proceed under (g).

MR. CAMERON: Which to be perfectly frank, the existing commission at that time -- okay, which was a five-member commission -- the Commission did not want to eliminate the option.

MR. MURPHY: Which -- at what time, though? When the rule was adopted or --

MR. CAMERON: No, no, no. Just when the three NRC options were approved back in '93, the staff proposed that to the Commission and the Commission did not want to say at least at that point, that only (j) would be available for docketing. But keep in mind that there is another Commission paper from last year which we distributed to the panel that went through some of the -- for lack of a better word, compliance sanction options open to the commission in case the design and development of the system by the --

MR. MURPHY: That's the paper I'm thinking of. Yeah.

MR. CAMERON: -- did not happen. And one of the things that was left open to the Ccommission then, is that even though they did not want to change the requirement in

the rule on (g) docketing now, that they could always choose to do that in the future. So, that's still up in the air. And you're certainly correct to point out the statement on his views at this point, but I don't think we can speak for the rest of the Commission on that.

MR. DAVENPORT: Or the ones to come.

MR. CAMERON: The ones to come.

(Pause.)

MR. MURPHY: I guess my question and my point relates to the approach you take for your scheduling. Perhaps it would be better rather than to take an end date and work back, to get it in, to do like you did costing; to build it out of zero, to take the actual amount of time you think it will take to do these things and stretch them out and see where you end up. Because if they take an assertive role on quality, they're going to be putting you there ultimately, anyway.

MR. DICKERSON: Well, but another part of that, Jim, is, that what fixes your rate for a very large -- resources that are available to do certain things. For example, the reprocessing of records. And if indeed you want to start at zero and work forward, if you want to reduce the reprocessing resources to a tenth of what they are, you can stretch it out for 30 years.

MR. DAVENPORT: Well, yes, but you have to work

out knowledgeable assumptions. I mean, how much money are you going to have? Obviously, fund it at next to zero and make it last forever.

MR. DICKERSON: Right.

MR. DAVENPORT: But, given projections of what resources you will have, and the quality that you want to build, I think you're scheduling builds off of today rather than starts in the future.

MR. CLINE: I'd just like to point out that the schedule has gone through a number of --

MR. HOYLE: Could you identify yourself? Just identify yourself.

MR. CLINE: I'm Mike Cline with the M&O.

MR. HOYLE: Thank you.

MR. CLINE: And just a -- the schedule went through a number of iterations, and we did initially start from the out and work back, but we also started from the baseline and went forward. So it's -- and we took costs in all the -- on the various issues into consideration, so it went through quite a bit with --

MR. CAMERON: And I guess I'd just say on -- generally, that from our work with Claudia and the DOE M&O staff -- is that it seems like they have full attention of trying to get a workable system loaded -- developed by the time so that it can be used for the license application.

And I would also say that the Commission, not just the chairman, recognizes that the LSS is a critical path component in the licensing process and we want a fully operational and loaded LSS by the time the license application comes in.

MS. NEWBURY: We're taking your no LSS -- your license application very seriously.

MR. DICKERSON: Okay, the second test scheduled that we have here is that of -- instead of building the LSS, of acquiring it from an outside source. Now, if we -- you'll see that the front end of this is just as it was before. The back end of it is pretty much as it was before. But we get into the procurement process. Any number can play in the procurement process. This has caused more heated discussions among the working group, I think, than any other subject -- is, what is a rational time period to use for procurement. This happens to be a set of numbers that we have up here. I won't defend it.

MR. CAMERON: Fielden, can I ask you a question about the build versus buy? Certainly there's going to be procurement -- won't there be procurement involved in the build also?

MR. DICKERSON: Yes.

MR. CAMERON: And how much of a problem does that -- of the same type of problem does that present,

versus the --

MR. DICKERSON: The procurement type of -- that's for the build is that basically of going out and buying things that are well defined. What we're doing for this is going back and saying "This is a system that we want to purchase and we're asking you to design it in the sense of using what you have available as a manufacturer, and putting it together to give us a turnkey operation." Much was -- I'm probably saying that badly, Camille. Would you care to take a try at it?

But, basically, you know, you could -- for doing the building ourselves, one can go out and say "I need a number 3,000 pyramis", for example, and you simply -- you know how to buy that.

MR. CAMERON: But you're really -- when -- maybe the distinction needs to be made clear between build and buy. On one hand, when you're talking, it seems like the build is DOE going out and buying components.

MR. DICKERSON: That's right.

MR. CAMERON: Procuring components, okay?

MR. DICKERSON: Procuring components.

(Pause.)

MR. CAMERON: And on the buy, it's -- you have -- you give the contractor the functional requirements and they basically go out and buy the components.

MR. DAVENPORT: That's right. That's right.

MR. CAMERON: And they may be buying the same that they would buy up to the build.

MR. DAVENPORT: They could very well. But they are -- you know, putting it together as they will put it together to get us a turnkey operation.

MR. CAMERON: I think what you need to do if you haven't done it already, is -- the procurement process is -- the procurement process is complicated in -- is there anything difficult that should be in the LSS build schedule on procurement that isn't in there now, versus what's in the buy procurement process? In other words, it looks like a big problem under buy, but not a big problem under build.

MR. DICKERSON: I think that's because we're looking at -- like, saying a turnkey system is a procurement process for the total package. We're proceeding with our record system and we think we've got the best products available today, but this technology is changing probably monthly. I have -- if we get to the point through February testing our system, or a glitch shows up that it won't handle -- say, 40 million pages and we feel that we need to go out or at least solicit the marketplace and see if there's a cheaper way to provide all the functions that we've determined, I think would be appropriate for us to do that. At least solicit the marketplace and see if

something's available.

MR. CAMERON: But it is probably easier to procure under the build than it is the buy.

MR. DAVENPORT: Oh, yes, because all we're doing is procuring parts and pieces, rather than a total system in a lifecycle procurement process.

MR. CAMERON: And one thing to think about, I guess, for future discussions that we're going to have about operation and maintenance is how the build versus buy -- what the implications are for the operation and maintenance of the system in terms of the NRC taking over those functions of operation and maintenance.

MR. LEVIN: Under both schedules, you have revised functional requirements. Are you -- is anybody going to talk to the process of --

MR. DICKERSON: Yes, we are. A little later, I -- it's the very next topic, as a matter of fact. You're just one step ahead of us, Moe. Okay? Thank you.

(Pause.)

MS. KERRIGAN: I'm Camille Kerrigan and I'm with the M&O in Vienna. And Info Streams developed -- I'm the deputy manager for the Information Management Systems Organization in Vienna, and Info Streams is under that organization. And, also, the records management system in Vienna is under that organization. I was on the LSS working

group. So that gives you just a little bit of a background.

Fielden just showed you the schedule for what we would expect -- some of the considerations for getting to the LSS. What we as the LSS working group said -- "What would we have to do to complete this project called LSS?" and no matter which way we go, we have to look at what the requirements are. And earlier in the presentation you had Fielden giving you a review of what we saw in the regulations as to what the requirements would be.

However, there was a document called -- the name of it was "The System Level Requirements" document, and just to give you some background of that document, it was developed in 1988 by SAIC and it was last revised in 1990. We looked through the literature and could not find any formal documentation that this group approved the document in any formal way.

MR. MURPHY: We didn't -- it was never approved?

MS. KERRIGAN: Sorry? Yeah, it was never approved. But it represents --

MR. SILBERG: We were never asked to approve it.

MR. MURPHY: Well, but I think if we had been asked to approve it we probably wouldn't. I mean, there was a fairly strong feeling among the group at that time, I think, that it was -- it went too far. Don't you recall that?

MR. SILBERG: Yeah.

MR. MURPHY: I mean, it was much to --

MS. KERRIGAN: Detailed?

MR. MURPHY: Gold plated.

MS. KERRIGAN: Yeah.

MR. MURPHY: Too ambitious, too expensive.

(Pause.)

MS. KERRIGAN: And you'll see we find the same -- you'll see the -- this briefing shows what we felt were the inadequacies in the document as well. But there were -- I believe some very good reasons at that time as to why the document was developed the way it was.

The document was prepared subsequent to a prototype proof of concept that was performed by SAIC, and therefore, I think the document -- rather than representing a real set of requirements that is just slowed down from what the regulations say -- I think it really represented a set of requirements that could be extracted from their proof of concept.

And what we were looking at is, if we go out to the vending community to the vendors and say "What could you build if we gave you a set of requirement?" we were wondering, could we use this document to issue as part of -- let's say an RFP. So, we formed a committee. I was on that committee to look at the document and to see if it was

adequate for that purpose. And we found that it was not.

Again, it's not a criticism of what SAIC did at the time, because they did that as a result of the proof of concept. But from just a generic point of view, we felt that the document was specifying a design, and if you were going to go out to the marketplace and say "These are my requirements, what do you want?" you don't want it to specify design. You want to get the most current, innovative concepts that are in the marketplace of the day.

The document also has a mixture of system and procedural requirements; things that -- if you go out, nobody could satisfy, because it's a matter of setting up a procedure to satisfy those requirements.

Another shortcoming in the document that it was really difficult to determine what the hard requirements were -- because sometimes they would say "The system should do this." Well, should means maybe it -- you know, maybe you can and maybe you can get away with not doing it, as opposed to saying the system shall, which is a hard requirement.

The document contains extensive specifications of non-quantifiable things. You know, it has to be user friendly. Well, how do you test when someone turns a system over, whether it's user friendly? And the requirements are restated for different system functions in a way that makes

them conflicting.

The rest of my briefing is basically examples to show these deficiencies, and I can go through that very quickly. Here are some examples of specifying design. They tell you the system has to be built in four sub systems, so they're actually going through and saying, "You will deliver the system in this fashion", and that's very limiting if you're going to go out with an RFP to say "satisfy these requirements." And the document even goes on to tell you what the interfaces between those sub systems will be, which is very much design.

Here's an example of conflicting requirements. In one of the sub systems, they talk about hard copy of large outputs, and this is a direct quote, "greater than 100 pages should be available overnight." In another part of the document they talk about "wide volumes of hard copy". Again, this is a quote from the document. Now, they say "Thousands of pages should be available via remote print and distribution within five working days after request." Well, is it five working days or is it overnight? You have to resolve those if you're going to put that level of detail in the document.

Here are some examples of ambiguous and conflicting requirements. "Average clearing response time" -- they don't define what average means -- "shall be

less than 10 seconds." And it says "terminal response time to begin to respond shall be no more than a few seconds", and then "search time to identify all documents" I left a little part out -- "shall be completed within a few minutes." So, sometimes you find very specific requirements like 10 seconds and sometimes you find very ambiguous things like a few seconds, a few minutes. And that is all throughout the document. These are just examples.

MR. MURPHY: Even more substantively, I think they reflected requirements -- and I shouldn't speak for everybody -- but certainly, when I was interviewed, they reflected requirements that I never in my wildest dreams thought the LSS should do. I mean, I can recall being interviewed by somebody from SAIC and asked that question, you know, "How quick do you need this stuff?" Ten seconds is absurd. I can't conceive of any circumstances under which I'd need the 10-second response. Even if I was asking for it at counsel table in the middle of a hearing, I wouldn't need it in 10 seconds.

MS. NEWBURY: And if you go back to Fielden's briefing just right before this, you see a period of time where the requirements document is going to be revised. And then I believe there's a period of time in there where the LSSARP gets an opportunity to look at the document to see if that's what you want. Again, these -- you can't test to

these requirements. You know, to say it should be interactive and intuitive, yeah, it should be. But you can't put that in a requirements document where you're going to reward a vendor for doing this, because these are not hard and fast testable requirements.

And then there are scattered throughout the document, things that a system would not implement. They would be implemented through procedures. Like -- this is in the document. "All documentary material shall be submitted to the LSS capture systems in accordance with procedures established by the LSS administrative." And note the "shall". That is a hard and fast requirement in this requirements document, but it is nothing that -- I mean, it's a procedure that you set up. It's not something that a computer can assure or build into the system. And rather than read these, they're just other examples of things that could not be implemented from an automated system, but could only be implemented by setting up a procedure.

So, the conclusions after -- it's quite a lengthy document. We recognize that if we were going to use this document to issue an RFP and that's not to say that that conclusion has been reached yet -- that that will happen -- we couldn't use this document because of the reasons I just went through. So, we need to revise it to correct the deficiencies, remove the ambiguities and

conflicts and remove the references to the design.

Obviously, it would be on the critical path because you couldn't make your decision about the LSS until you got your requirements listed at least to everybody's satisfaction, that that is what you want your system to do. And this last bullet, I didn't have the discussions, but -- I don't know who did. I don't remember exactly -- we did discuss with someone at the NRC about revising the document and, as you said, the LSSARP had never approved this document.

(Pause.)

MR. MURPHY: Did DOE ever approve that? Did DOE ever accept that body, do you know?

MS. NEWBURY: I don't know.

MR. DAVENPORT: Not that I know of.

MS. NEWBURY: I've seen it as a DOE publication. I'm not sure if it was a DOE publication. I've seen a published version of the requirements

MR. MURPHY: Yeah, SAIC publication -- my memory is a little bit fuzzy, but it seems to me that the department never even accepted that because it was a lot of these problems. I mean, I don't -- you know, I'm sitting here wondering what is there to revise? Throw it away. Write a document that we can use.

MS. KERRIGAN: A lot of work went into it and

there are a lot of very good requirements --

MR. MURPHY: Well, a lot of work went into designing the vertical exploratory shaft too and we're not worrying about that.

MR. CAMERON: Yeah, I don't think there's any particular magic to the SAIC document now, and I think that we've been sort of assuming that you're taking every look at it to do something sensible.

MS. NEWBURY: And that really is --

MR. CAMERON: If that's what you mean about NRC as amenable.

MS. KERRIGAN: Right. And I think that's --

MR. MURPHY: Is it okay with you, Moe?

MR. LEVIN: I was going to say it's okay.

MS. KERRIGAN: We'll have to -- that's an important point to get across.

MR. MURPHY: I don't remember the conversation, but had I been a part of that conversation, I would have been amenable to --

MR. SILBERG: What are you doing to re-look at this and what is the schedule?

MS. KERRIGAN: The schedule was part of the schedule right before this. We are putting together -- and I probably shouldn't be the one to speak to this, because it is another organization within DOE that is chartering

themselves with the -- developing a requirements document. But I believe the schedule is, we're supposed to have one out at the beginning of calendar year '95.

MR. BOONE: Well, for that person who's part of the organization that chartered itself --

MS. KERRIGAN: Right.

MR. BOONE: -- my name is Jim Boone.

MS. KERRIGAN: Okay.

MR. BOONE: I'll just take that question and run with it.

MS. KERRIGAN: Okay.

MR. BOONE: My name's Jim Boone. I work for the M&O and the project and management organizations within the M&O, and I support Claudia directly with LSS and Technical Data Management issues.

Now, I was relieved to hear some of the comments about this requirements document, because we weren't aware if there were any sensitivities to the content of it and what we have done is, tried to look at this document and revise it to make it something useful and try to carry forward as many of those requirements as we could into the revised document.

Now -- so I'm refreshed to hear what you've told me, that you're not going to hold our feet to the fire for every one of these. But our intent is to generate a useful

document that the computer industry or information management folks take and build us a system that will perform the functions that we really want the LSS to do. That's currently being worked internally within my area. We will be turning this over to others within DOE to review by the end of this year. As expected --

MR. SILBERG: This calendar year?

MR. BOONE: -- this calendar year, and depending upon how -- what sort of feedback we get, it could be January or February of next year when we can bring it forward to the committee. And that's one item that I think the technical working group could look at as one of their first items -- is to take this draft -- the functional requirements document and make sure it has the right spin on it -- that it does what everybody expects this thing to do, so we can move forward.

MR. LEVIN: Now, are you saying that you're in the process now of revising -- coming up with this revised document?

MR. BOONE: Yes.

MR. LEVIN: what are the procedures -- how are you going about determining the functional requirements?

MR. BOONE: What we have done is taken (j) as the basis. These are the high level documents. Camille talked about flowing requirements down. We're trying to flow these

requirements down into lower level functions. Now, the actual hard requirements are fairly sparse in (j), so there has to be some interpretation made on what is expected. We know that it's a document discovery system, we know that the law community would be significant users of this, the law community is familiar with certain systems such as LEXIS. So that gives you a clue on what type of interface might be compatible with what people expect. So, there are some derived requirements, and that's a very high level one that this -- you know, have this sort of interface. So, basically it's a combination of regulatory requirements and derived requirements which spell out a paragon that we think should satisfy the majority of the users.

MR. LEVIN: When -- at what point do you plan to go to the users themselves again and ask their requirements?

MR. BOONE: I would view that as bringing the requirements before this group. Now, we -- I caught -- I hesitate because what we don't want to do is to create something that's gold plated. If we went and surveyed everybody, everyone has a different idea of what the system should be, and that's not the mode we want to get into here. What we want to do is to design a system and implement it in a timely fashion -- something that's practical -- and focus on getting something up and operational first. So we try to steer away from any of the extraneous sort of functions that

aren't really necessary to support the basic mission of the LSS.

MR. GANDI: I think another point that needs to be made is, I've been -- kept saying "We're going forward with our records system", realizing we've got to start doing this process and reprocessing, anyway. The requirements for our records system include Federal records management, our -- schedule, but they also include the LSS requirements. I, myself, would like to have a clear handle of what those requirements are so we don't have to get those damn things up again and reprocess them a second time. I think that's probably just an advisory review panel's responsibility to give us some clear direction as the developers, on where we go.

MR. METTAM: When do we get to see that?

MR. GANDI: I would hope you would get to see these documents in a preliminary format prior to any issuance of them.

MS. NEWBURY: I think it would be fair to say that when we come up with the requirements document that's a straw man, that we'll come to the panel for review and you can add to or subtract from as you choose. But you have to have something to start with, and what Jim will be producing will be the straw man requirements document.

MR. BOONE: Trying to take a regulatory spin here

and not create a monster.

MR. METTAM: I understand. You know, there's a difference between what people would like to see in the system and what people need to have in the system, certainly. You know, we would anticipate your filling the latter rather than the former. But the question still is, though, when do you think that would occur that we'd have an opportunity to look at that straw man?

MS. KERRIGAN: Jim, when did you say?

(Pause.)

MR. BOONE: I am shooting for -- trying to get this out for an internal review by the end of this year. Now, I'm not certain -- I can't judge how that's going to be received by the DOE or any of the other contractors on the project.

MR. SILBERG: Well, are you talking about the -- you're talking about the LSS requirements and I'm talking -- you're talking about the RMS and --

MR. BOONE: We're talking about the LSS requirements.

MR. GANDI: Which is a portion of the requirements in the system we're rebuilding.

MR. SILBERG: Your reprocessing, you said, is going to be done through a series of requirements.

MR. GANDI: Right.

MR. SILBERG: Where are those requirements reflected?

MR. GANDI: Jim's developing them and we are -- with management --

MR. SILBERG: So it's the same --

MR. GANDI: -- organization as --

MR. SILBERG: -- set of requirements.

MS. KERRIGAN: Right. What John is saying is, he's more anxious than the LSSARP to get what those requirements are.

MR. SILBERG: Okay.

MR. BOONE: And what we're really doing is, he wants --

MR. MURPHY: But he -- but you -- you're going to produce something hopefully by the end of the year for --

MR. BOONE: Hopefully.

MR. MURPHY: -- internal review.

MR. BOONE: Yes.

MR. MURPHY: And then you folks at DOE are going to take a look at it for some period of time.

MS. KERRIGAN: As quickly as possible, we will get it to the LSSARP.

MR. MURPHY: So, it sounds to me like in answer to J's question, sometime late winter, early spring --

MR. SILBERG: You know, I thought there were two

separate --

MR. MURPHY: -- until we can have a --

MR. SILBERG: -- sets of requirements documents. I don't know --

MR. MURPHY: Yeah. No, I guess this is the same thing that --

MR. CAMERON: No, this is the same thing that you said, that the working -- might be good task for the working group to review, and I guess it's up to the panel to decide whether they want the working group to take the first crack at that ad then report back with recommendations.

(Pause.)

MR. MURPHY: That was my next -- that was going to be my next point. So that we -- you know, we haven't gotten to whether or not to establish that working group yet, but that would seem to me to be the logical progression. To have the technical working group take a look at it again. I was going to say some time late winter, early spring, and then late spring have a meeting of the full ARP to --

MR. DICKERSON: Matter of fact, now that's the schedule. In terms of putting our schedule together for the LSS, we were anticipating bringing this all together so that this would be a topic for the March meeting.

MR. MURPHY: The March meeting of this group?

MR. DAVENPORT: Yes. Yes.

MR. MURPHY: Well, that's quicker than I thought, but that's fine. That would be great.

MR. BOONE: This is a critical path item so we're trying to move forward as quickly as possible.

MR. GANDI: Are the -- is the substance of this report or this document more on legal procedural or more on the technical side?

MR. BOONE: No procedures whatsoever. What we're looking at is a system design. And so the document is structured -- the current vision, unless people have severe problems with that -- but it's a layered document. So at the very top layer there are the general descriptions of what the system should do and it comes right out of (j), and some are prudent good practice requirements. There is another layer which describes those upper level requirements in more detail, and that's geared at -- say, a group such as yourselves, which don't want to get into the nuts and bolts but you want to have a good feeling of how the system is going to look. And the final level are the nuts and bolts, and these are what our computer friends will use to build this thing. So, the idea of the structure, at least, was to give different audiences a different layer to look at so they have a good feeling on --

MR. SILBERG: Well, it sounds like instead of sequential reviews, these could be going on -- the technical

folks could be looking at the lower level or the bottom level technical thing while the non-tech use -- the first couple --

MR. MURPHY: You're probably right.

MR. BOONE: Yeah, the theory is that all the lower level requirements would support the higher level ones, all the way up the -- all the way up --

MR. CAMERON: Let's hope not.

MR. BOONE: Okay, any other? Well, now to the other reason I'm up here. At our September 9th meeting, Mike Cline gave an overview of our activities and touched briefly on the data volume projections that the working group made. I'd like to go into those in a little bit more detail. Not anything real deep.

We have to first step back and say "Why are we even doing data volume projections?" and there's really only two reasons. Number one, we want to try to get a handle on what the total life cycle cost of this LSS would be. As George pointed out, data volume is a significant cost driver. Costs are almost proportional to the data volume. But, clearly, we need a reasonable projection in order to get a ballpark for the funding profiles we do, and we also must supply our system designers to some sort of capacity -- something they can design their system to accommodate.

Now, SAIC recognizes needs and they performed two data volume projections. The first one was in March of 1988 and it was this document, the preliminary data scope analysis, and this document was later revised in August of 1990, which is the revised data scope analysis.

Now, in the first document, the 1988 document, what SAIC did was, pull the various records management systems that were associated with the project at that time -- there wasn't just one as Marty mentioned, there was an -- a conglomeration of various records systems. So SAIC went out and determined what the historical growth rates for these were, they estimated a nominal number of pages per document, percentage of documents judged to be licensing relevant, accounted for duplicate documents, et cetera, and they made a projection from 1988 -- March 1988 to August of 1990. And it was believed at that time that the LSS would be partially loaded in August of 1990. Now, these estimates were made -- they were presented as a low and a high in order to get a bounding feel for what these numbers would be. And the numbers were extrapolated from 1990 to 2009 by basically assuming a compounded growth rate of 10 and 20 percent respectively for the low and the high projections.

The original numbers were revised, as I said, in 1990, after it was discovered -- significant error in the initial loadings down here -- magnitude off. SAIC

recognized changes in the program strategies and so forth at the time and basically presented a new set of estimates, which curiously ended up near the same volume that was estimated in 1988.

The primary difference between these two, as best I can tell, was that the percentage of documents considered to be licensing relevant increased from here to down here (indicating).

MR. METTAM: Jim, before you leave that chart, you should have some history now for a couple of those early years. Do you have age volumes for '90, '91, '92?

MR. BOONE: Yes.

MR. METTAM: Are they in here somewhere or can you throw them on this chart just to see if they fall in either one of those estimates?

MR. BOONE: I don't have them here. They're in the working group report.

MR. GANDI: Page 11 --

MR. BOONE: Yeah, but that's '94 and above. '94 --

MR. METTAM: I did a quick scan and couldn't find anywhere where the historical -- you know, just to see if the lines make any sense.

(Pause.)

MR. BOONE: We have those numbers in the working

group report which -- but I don't have them committed to memory.

MR. BALCOM: And where does the backlog figure end here, if it does?

MR. BOONE: The backlog as it currently exists -- are records that are already into the system. So, we -- processing -- just reprocesses existing records. There is no backlog, if I'm correct.

MR. CUMMINGS: Let me -- Marty Cummings here -- talk about terminology just for a second. Backlog is records that have not been processed into any system. And the reason we talk about the reprocessing effort and we don't use the term "backlog", is because it is records that are already in the system. And so, you know, right now the reprocessing effort as you saw earlier, is around 500,000 records at an average of about 13 pages per record. And we do have -- I mean, we do have that data there, Mr. --

MR. BOONE: Okay, as a rough -- adding up these numbers from the report, we have roughly -- at 1994, roughly right here --

MR. METTAM: What, eight or nine, something in there?

MR. BOONE: About seven million pages, I believe, is a nominal --

MR. CUMMINGS: Yeah, close to seven million rings

a bell.

MR. BOONE: Right around there.

MR. METTAM: Do you -- can I borrow your report to just look a couple of years earlier than that?

MR. BOONE: Yeah, this one's marked up. Can I get -- after the break, let me give you --

MR. METTAM: Sure.

MR. BOONE: Okay. Someone didn't do these very --

MR. BALCOM: But you're going to show your own estimate anyway, right?

MR. METTAM: Right. But service in '94 goes on.

MR. BOONE: But our estimate is based on historical information, so I guess that would be another starting point if you looked at it.

Now, the working group needed -- realizing we needed to update these projections primarily to reflect what our current program milestones and deliverables are -- used to be term the proposed program approach, it is currently the program approach. And we also -- Marty talked about efficiencies that have been realized in the records management system. We wanted to reflect these efficiencies in our estimate. Something that SAIC could not have accounted for.

MR. DAVENPORT: Before you begin that, didn't the

NRC say that they had some concerns about the program approach and as -- it falls to assume that it's going to work that way if it hasn't been determined that it's going to work that way?

MR. BOONE: Well, that's a whole different issue that I can't answer. All I know --

MR. DAVENPORT: Am I right or wrong?

MR. BOONE: -- is, we were chartered with -- assuming that this was the approach that was going to be carried forward.

MR. DAVENPORT: I understand that.

MR. HOYLE: The Commission hasn't ruled out the program approach. Yes, it has concerns. The Commission meets with Mr. Dreyfus again next Monday back in headquarters. And the Commission specifically asked him to come up with some more details, that it had been a little too vague in the past and they wanted to hear some real details this time.

MR. DAVENPORT: The last meeting I was at, I perceived them saying this dog won't hunt. Now --

MR. HOYLE: No, I didn't hear that.

MR. DAVENPORT: You didn't.

MR. HOYLE: No. I'll send a transcript of next week's meeting --

MR. DAVENPORT: Yeah.

MR. HOYLE: -- to the panel.

MR. DAVENPORT: In any event, are you venturing down a course which perhaps needn't be ventured? Or --

MS. NEWBURY: Well, at this point since we are -- DOE is following the program approach, it would be a reasonable assumption with volume projections, to use the program approach the DOE's five-year plan is based on. We --

MR. DAVENPORT: I understand what you're saying to your contractor. It makes sense. But the problem with the program approach as we perceive it, is that it defers work -- important work -- essential work out into the future which needs to be done pre-licensing, which would make your estimates' volume pushed out in the future rather than in the near term. It would seem to me that that would be a crucial enough variable that you would want to have looked at the volume projections, assuming the statutory approach, the one that's required by the statute as opposed to the one that DOE currently as on its mind.

MS. NEWBURY: So, you would like to see a model developed that includes volume projections that would assume that we have all the data we need at license application?

MR. DAVENPORT: Correct.

MS. NEWBURY: We can probably run another model that includes that.

MR. SILBERG: Does that really change anything, or all it does is push that license out?

MR. BOONE: Time phasing. That's all.

MR. SILBERG: Well, it pushes out the license application -- the data will probably look the same thing, it's just a question of whether the -- when the license applications are --

MR. DAVENPORT: Except that if you require that all the information be developed before licensing begins, DOE has an objective of procuring that data promptly, and if they can develop it as they want to get to it, there's no time constraint on them.

(Pause.)

MR. SILBERG: I always had the -- worked on the assumption that they were generating the data as fast as they could and I don't think they could speed it up any more.

MR. DAVENPORT: It's good from our prospective.

MR. GANDI: The data -- to size the system up -- excuse me, size the system in the costing model, facilities describes -- and such, and the system performance. That was one of our main concerns on the volume factor. How much it would change --

MR. BOONE: A preferable way, I think, to select the various options is to present a curve, basically, on

dollars versus volume, and then you could choose your intake. Because everyone has an opinion. I mean, I'm going to get to that, but there's no right or wrong answer. There's a thousand -- there's an unlimited number of ways to make projections. This is just one way. And what I -- I'll steal our thunder, but what we're going to see is that the volume -- that our volume projections are somewhat in line with SAIC's. We could come up with that answer many different ways. I feel confident in that.

Should I move forward?

MR. BALCOM: I've got one thing. The updated projections were needed to reflect refinements in record system screening practices.

MR. BOONE: Yes.

MR. BALCOM: I don't know why, but that terminology -- I think I'd like to know a little bit more about that.

MR. BOONE: That gets back to Marty's presentation earlier today, and perhaps Marty should expound on that.

MR. CUMMINGS: I'm not sure I understand your question.

MR. BALCOM: Page five -- I think it's the prior -- yeah, updated projections were needed to reflect refinements in records systems between the practices. I'm going to guess that that means fewer document.

MR. BOONE: That means -- he gave the example of travel and rental car vouchers and things --

MR. BALCOM: Right. Is that all that means?

MR. CUMMINGS: Right. Yes, sir. It really is. It's -- you know, back when SAIC did the study, I mean, there was the -- there -- we just didn't have the inclusion/exclusion criteria back then that we have today, and that is from lessons learned and experience over the years. We've been able to capitalize on that, and so the -- we have just gotten a lot better at doing what we're doing and we're not putting in the stuff that is not relevant.

MR. BALCOM: Because I remember a meeting a year or so ago where there was talk about -- there was a possibility of using automatic programs to determine relevancy for documents and I guess I'd like to know if that's still a possibility.

(Pause.)

MR. CUMMINGS: We have a small prototype. What -- we don't operate by it. It is still something that we're thinking about, but that also is determined by how the screening process does end up and how difficult we make it. If it does -- you know, if it stays or becomes a difficult screening process of what goes in, then we certainly want to take advantage of that kind of technology. If we simplify

the screening process by not having to put a lot of time in processing power and looking at what whether records should go in or not, then that wouldn't be necessary. But that is something that we're still thinking about currently.

MR. BOONE: As you'll see in a little bit, the working group purposely excluded the issue of document relevancy from our calculations. We separated that variable out.

MR. BALCOM: Okay.

MR. BOONE: The actual projections themselves have three components. The first component is DOE, NRC is considered the second and the various stakeholders the third component. Now, it's been suggested by SAIC and others that DOE would supply 80 to 90 percent of the verbiage of the documents that go into this system, so clearly that's where the majority of the focus was -- to try and get a better handle for the lion share of where the documents were coming from.

The DOE contributions considered the headquarters as well as the project record system -- we just lumped them together and treated them as one. As I said before, we separated out licensing relevancy all together, so what we're looking at is this total information going into our record system and we also attempted to couple our projections to the current program approach, as we were

directed to.

(Pause.)

MR. METTAM: And the project level documents include all the documents?

MR. BOONE: Yes.

MR. MURPHY: I'm sorry, I didn't hear that.

MR. BALCOM: I'm sorry. I said that -- I asked if the project level documents would include all the participant documents and the answer was yes.

MR. BOONE: Yes. NRC and stakeholder contributions were estimated in a much less rigorous way. We contacted NRC and we obtained two numbers. Historical data for 1990 and 1994, and there's a simple linear extrapolation. The NRC told us that this number included all the licensing relevant documents, and it was a -- assumption that the stakeholder contributions represented 10 percent of the NRC's. Now, clearly, you folks can supply us more information regarding that and we'd be happy to include that in our --

MR. MURPHY: And who was -- yeah, what was --

MR. BALCOM: This debate is meaningless until we have --

MR. DAVENPORT: It's coming.

(Pause.)

MR. BOONE: Yeah, I know. I know. Although,

I know what your reaction's going to be, but we'll see. Instead of simply scanning some sort of mathematical group on these records, what we tried to do was to poll various people that had been working on this project within the DOE and the contractor community for years and hopefully tap into some of their knowledge. We sent out a survey to roughly 20 people, senior people on the project and asked them to give us their projection of basically what the profile -- a relative profile would look like between now and 2010.

And I think a picture is worth a thousand words, but basically each participant was asked to generate a bar chart similar to this and their instruction was if they thought that, for example this one, that in 2001 is where the majority of the records will be submitted, then this curve is normalized to one in that particular year and all the other intakes are normalized relative to one.

Now, with 20 different people, we got 20 different answers on this, that the profiles were widely scattered. And we used some high-powered mathematics and combined these statistically using a harmonic mean to throw out the highs and lows and this is what you've got. I had to go back to some of my engineering for that. But this is what we ended up with. Now --

MR. DAVENPORT: You're not going to use that

number to compute anything off of I hope.

MR. BOONE: Yes, absolutely.

MR. DAVENPORT: Your salary.

UNIDENTIFIED SPEAKER: I'll take it.

MR. BOONE: Now, what we do --

MR. DAVENPORT: Now, wait a minute. What one?

MR. BOONE: We'll find out. Okay.

MR. DAVENPORT: No, tell me.

MR. BOONE: I'm going to tell you right now.

MR. DAVENPORT: But how could these calculations have been done if you did not know --

MR. BOONE: They haven't been done. This is an intermediate step. These are basically multipliers. Now, all we've done, we know -- well, given -- we have a good feel for what the 1994 records intake is going to be. So we basically calculated a multiplier based upon this value here in our absolute number that we have and applied that multiplier across the board to turn these into -- does that make any sense?

MR. DAVENPORT: It does and it's just backwards.

MR. BOONE: It's a projection.

MR. BALCOM: It's a good feeling. I'm just thinking back to what Camille said about the SAIC terminology. No, I'm just kidding.

MR. DAVENPORT: Now, my point is this: You

couldn't ask anyone what would happen in 1994, what percent '94 would be of 2001, is it, that you're --

MR. BOONE: 2,000, yeah.

MR. DAVENPORT: No. Your total one -- there, right. Okay. You couldn't ask a person what percent of that '94 would be if you didn't know what 2001 was going to be first.

MR. BOONE: I'm not sure I understand. Please.

MR. ECHOLS: It's asking the 20 different people what year do you think will be the peak production year of documents and that was -- a lot of people thought, well, that will be right at the licensing proceeding. And as you ramp up with site characterization information, suitability information, you know, this was a relative profile. Then they said, well, based on what, If that's five times what the current level is, what's the current level. That's how they've got --

MR. DAVENPORT: Yeah, and that's the most nonsensical approach I could imagine; it's just backwards. I'm sorry. I made my point. I don't mean to argue. It's just --

MR. BOONE: Well, it makes no sense to me and the way they had this intended.

MR. ECHOLS: This is just to try to construct, as I understand it, just a profile of relative numbers of

documents produced per year, what would be the drivers of it and is that based upon now as the budget increases and so on. So that's what they --

MR. DAVENPORT: Go ahead and do it if you want to.

MR. ECHOLS: As I understand it that's how they produced it.

MR. BOONE: Well, you can't really ask the people ask what is your total -- how many records do you expect in 2001 or 2010. They have absolutely no idea.

MR. DAVENPORT: Exactly.

MR. BOONE: Absolutely no idea.

MR. DAVENPORT: Exactly. That's what I'm telling you.

MR. BOONE: Agreed. That's why the relative approach was taken here. Okay. You don't have to know what the absolute --

MR. DAVENPORT: I not only have no idea what I'm going to need in 2001, but I know that I'm going to need precisely 28 percent of that in 1994.

MR. BOONE: This is not precise. This is a mixture. Well, like I said there are unlimited numbers of ways to do this and this is just one.

This next graph is a comparison to the 1990 SAIC projections. This top curve of -- let me back up.

When we talk to our records people, they feel that

if we excluded records specifically excluded by 2.1005, that section of the rule, then roughly 90 percent of our current record system intake would be considered relevant -- oh, I'm sorry. I said that backwards. Roughly 10 percent of the existing material in our records system could be excluded per the exclusion criteria given in sub part (j). So this 90 percent represents an estimated upper bound of what the volume might be. Fifty percent shows that if you adopt some more detailed screening criteria to try to reduce the number of documents going into the LSS, that's just an indicator of the effect it will have on the total volume; 50 percent is arbitrary.

But the point to be made is in terms of order of magnitude we're not that --

MR. ECHOLS: What is your estimate, I mean, relative to SAIC?

MR. BOONE: Pardon me?

MR. ECHOLS: When all was said and done, the number of pages.

MR. BOONE: Who said that?

MR. ECHOLS: Stan Echols. The question was once -- so you've got the SAIC production.

MR. BOONE: The SAIC is here.

MR. DAVENPORT: And you're where?

MR. BOONE: We're roughly 30 million here at the

high level versus, what, 36 to 41, 42.

MR. DAVENPORT: And that's based not on just the 10 percent discount but upon your reconstructing the estimates from '94 forward. Okay.

MR. BOONE: This is based on our methodology which I just discussed.

Now, before I show you numbers like I promised, you'll recognize that these are order of magnitude projections, are generally sufficient for system design here. So what we saw is a validation; our numbers versus SAIC, the same ball park.

We also recognize that there is much -- very large error associated with these as there would be with any projection because you're trying to predict the future, a very difficult thing to do.

And I believe that John has sent a questionnaire around polling the various people on what their expected volume would be and should you folks supply that to us, we'd be happy to incorporate that into this model for our latest revision.

MR. MURPHY: You didn't send a questionnaire. You just asked us to be prepared to tell you what the figure would be. Is that right, sir?

MR. BOONE: That's --

MR. DAVENPORT: Now, did you tell us yet what

NRC's number was?

MR. SILBERG: Fifty-nine thousand.

MR. BOONE: So roughly 60,000 per year.

MR. SILBERG: Although the suggestion that that would remain flat and would not ramp up as you approached the licensing proceeding makes no sense at all. It's a small number compared to the OCRWM number, but I can't conceive that they'd be operating in a flat level.

MR. BOONE: I agree. It makes sense. But then we fall into the same -- we're subject to the same assertions that it makes no sense regardless of what methodology we use there.

MR. HOYLE: Jim, who gave you those --

MR. DICKERSON: I checked with Moe and he ran it down with one of your staff or someone there.

Do you recall, Moe, we had a discussion and we talked with one of your staff --

MR. LEVIN: Yeah, projection -- we checked the high level waste, yeah, MSS, and that was their projections.

MR. MURPHY: Did they -- they came with a flat linear projection like that?

MR. BOONE: No. They gave us two data points.

MR. LEVIN: Yeah.

MR. BOONE: Now, we asked for some speculation. We -- to my knowledge, we did not get that response.

MR. MURPHY: Well, they just got lazy.

MR. BOONE: We would certainly consider any profiles that NRC or others would propose.

MR. HOYLE: All right. Let me take another look at it, but maybe the total is the right representation.

MR. FRANK: I'm Jim Frank. And I think one of the things that's important was when we've talked about the cost model, we used a whole series of volumes and, when you talk about designing a system, you would never use this volume for the system design. You would always put in a factor of growth just like any error factor that we would put into any design of any weapons system.

So that I think this is a good basis for numbers and whatever happens, you know, the real requirements are probably going to be a little higher than this. But the cost allocation will show you the impact of it. They were really big, weren't they, George, six times this size?

MR. HALLNOR: Yes.

MR. FRANK: Yeah. So if you look at that chart that George put up about the sensitivity of volume to cost, I think you'll find out that doubling these numbers probably doesn't do a whole lot to cost.

MR. BOONE: Now, one parting note, too, is we have been asked to provide the final version of this report to DOE by mid January. So any input that you have would be the

sooner the better.

MR. MURPHY: On volumes. We're going to do that in the next day and a half I thought.

MR. BOONE: Good.

MR. BALCOM: Do it now. A total of 500,000.

MR. FRANK: Let me intervene one more time. I'm Jim Frank still. I recognize his problem with trying to identify voices.

Remember, as you try to make estimates, we only capture a document once in LSS. So if every report that comes out of the DOE you count and NRC counts, that's not right, because the LSS will only permit that document to appear one time.

That's also why when Marty was talking about estimates changing is that in DOE we got smarter, too. So we, when we reprocess data, it will only incorporate a document one time as opposed to having it being attachments to several other documents.

The other thing that transpired is the documents that are in public domain do not find themselves in the records management system. So if it's a textbook, for instance, we don't copy the textbook.

MR. DAVENPORT: What's your definition of public domain?

MR. FRANK: Marty's got the definition of public

domain.

MR. SILBERG: That doesn't include DOE publications, I take it?

MR. FRANK: No. If it's a DOE publication, we'll put it in there but, you know, things like textbooks.

MR. MURPHY: Cited references.

MR. FRANK: Cited references that are in the reference libraries that we maintain are not included in the record system and that's why it got -- but that's just a caution, because if we don't watch out, we'll all count the same document and we'll have a little problem with --

MR. MURPHY: Well, but we let me add a caution to your caution. You're absolutely correct. And the estimates I gave, for example, are estimates only of documents and it's an ambitious estimate --

MR. FRANK: Sure.

MR. MURPHY: --only of documents that Nye County will produce.

MR. FRANK: Good. That's appropriate.

MR. MURPHY: But you will on instances copy or include the same document twice. If DOE produces a document and it goes to USGS and somebody in USGS writes marginalia on that document, that becomes a separate document and you better have a system which doesn't -- which permits you -- which does not permit to overlook that marginalia. That has

to go into the system.

MR. FRANK: That's correct.

MR. CAMERON: Two pieces of LSS --

MR. MURPHY: You could have six memos -- that six copies of the same memo with different people's handwritings; they're all different documents.

MR. CAMERON: But keep in mind is this two pieces of LSS Arcania (phonetic), if I could use that term; one is marginalia, only it's not required to be placed in the LSS of an issue. It only comes in through the deposition process. Okay.

MR. MURPHY: Right, right. So it could get in.

MR. CAMERON: Right. It still could get in.

Secondly, in terms of -- the rule does provide for exclusion of what we call judicial notice material and it specifically says textbooks and reference books, but I would want to be very -- I would want to, I guess, add a caution to mount caution is that we really should be in agreement about what that universe is in terms of you mentioned references that are cited. I mean, I'm not sure we have a problem there or not, but I think it's just something we need to think about.

MR. SILBERG: Jim, could I just make sure I understand. Your number of 500,000 pages is the cumulative numbers of pages added to some year?

MR. DAVENPORT: Give or take a page or give or take a year, yeah.

MR. SILBERG: Okay. And the year is 2004 or 2003?

MR. DAVENPORT: Yeah. And we'll be bound to that. We're willing to be bound to that to the same extent that DOE has been bound in the past to deadlines by this --

MR. SILBERG: Penny a page.

MR. DAVENPORT: Any of those things.

MR. SILBERG: A penny a page penalty for being over on --

MR. DAVENPORT: Right.

MR. SILBERG: Your number now was --

MR. MURPHY: The same thing, same thing. It's --

MR. SILBERG: No, no. I mean, yours was 300,000?

MR. MURPHY: Yeah.

MR. DAVENPORT: So those are shot-in-the-dark numbers, Jay, you know.

MR. SILBERG: Oh, yeah.

MR. MURPHY: How many -- what's your numbers, Jay? Remembering, of course, that they're all going to be rejected because you guys won't even be allowed to intervene.

MR. CAMERON: I believe the non intervener number.

MR. SILBERG: I think it would be close to -- on the order of 10 percent of yours --

MR. MURPHY: Yeah.

MR. SILBERG: -- and I think that's optimistic. Most of our documents are things we get from other folks.

MR. MURPHY: Chip brings out historically significant phraseology from the days of the LSS negotiations, and that reminds me of something I meant to bring up earlier this morning and forgot to do it and that is circulated draft. That's -- there's -- that's another key little thing that was negotiated at some length and the system has to somehow be able to take account for circulated, but not a finalized drafts that are covered by the rule.

MR. CAMERON: Yeah. The important part is not covered by the rule because we went through this exercise at one of our last meetings where the term "circulated" is a misnomer. We kept that language for some reason; probably because we didn't want to make any more waves than necessary at the time, right. But it really is a non concurrence draft.

MR. SILBERG: Yeah.

MR. CAMERON: So we should be aware it's not like every circulated draft of a document has to go in. It's only the non concurrence draft. But you're right, you do need to have a handle. But I would imagine that would be within DOE; it would be pretty small.

MR. MURPHY: Yeah. You understand what we're talking about, Claudia?

MS. NEWBURY: Yeah, we were just -- we still end up with a record package relative to that type of document. In other words, to close out anything that we didn't concur on, you have to still put a final piece of paper that says we don't concur with it and put together a record package so it is in the records system.

MR. GANDI: Now, this brings up another little thing. We've talked inclusion and exclusionary lists, yet we've never been officially blessed by anyone on these inclusionary and exclusionary lists, and left to the records' people, the computer people to define what does 10CRF2 mean and I get a little uncomfortable.

MR. CAMERON: I mean, that's -- it might be useful at some point to see what lists -- to compare what lists all of us are using --

MR. GANDI: We'll happy to --

MR. CAMERON: -- to see --

MR. GANDI: -- give you a list of ours.

MR. CAMERON: -- if there is -- how much of a deviation there is.

MR. DAVENPORT: You might want to -- are we prepared to do that at this meeting?

MS. NEWBURY: No.

MR. MURPHY: We can put it on the agenda for a future meeting, though.

MR. SILBERG: Well, if there is an existing DOE list, it would be nice to send that around early before a meeting and we can look at it and, when we come out to the next meeting, we can talk about.

MR. DAVENPORT: And the list, sir, that you were describing of reference -- general reference materials and those kind of general exclusion things, if you could get that into a paper format that we could look at to make sure that we're all on the same wavelength.

MR. CUMMINGS: Okay. I would like to respond to that. We will get it to you down on the paper. But just to respond to it right now, the way we chose to handle cited references, I believe the wording in sub part (j) is "readily available." The way we've chose to define "readily available" is in our technical information centers and technical libraries and we have a cross reference from our record system that tells us it's in there. Our technical information center and technical library personnel understand that they can't do anything with that, you know, till somebody from records tells them.

But what it allows us to do is, for instance, taking a textbook or something, breaking that down and indexing it and imaging it into our system and all is a very

expensive processing power that we don't want to have to do if we don't have to.

MR. MURPHY: No. And it's not necessary to do it because the experts, the scientists who are going to be interested in that textbook already have it on the shelf in their office. If they're an expert on some, you know, many of the esoteric disciplines that we deal with in this program everyday, and somebody cites, you know, the basic text on climatology, the climatologists have that book and it doesn't have to be put into the LSS.

MS. NEWBURY: John, I think this is really a good topic for another meeting because it does bring up all sorts of things like copyright infringement and if you scan a copyrighted document into your system, what have you done to yourself, and I think it would be an excellent topic to spend a half day on actually.

MR. GANDI: I have several of these little --

MS. NEWBURY: And we can certainly provide our inclusion/exclusion lists to people to review prior to that meeting.

MR. SILBERG: A lot of these I don't think there is going to be any dispute, but it would be nice to know before everyone goes far down the road that, in fact, we're all on the same scripts and I think we will be.

MR. GANDI: See, we have some even simple little

questions. Like we have the concurrence ladders and when we talk about concurred drafts or whatever, we've been scanning that concurrence ladder, which is our yellow page, to the document in the system. Now, is that a requirement for normal correspondence to have the whole 20 people that concurred on that in the system, or is the person that signed that original letter as --

MR. SILBERG: You mean the separate sheet --

MR. GANDI: Yes.

MR. SILBERG: -- or just one sheet that shows each of the concurrences?

MR. GANDI: It's usually the first page of the document, but it's in yellow copy; so right after the first page of the document that's right now microfilmed comes the concurrence sheet.

MR. SILBERG: I think that's an important bit of information to know who signs off on a piece of paper.

MR. GANDI: Okay. Which brings up the -- my question is how are we going to handle it in our imaging system. So we've essentially added a concurrence page to our image text and header file. I just wanted to bring that up. We've got several of these little questions that we seem to stumble on.

MR. CLINE: We better move on if we can to close out at this point.

I'm Mike Cline. I was the -- or I am the lead for the LSS working groups for this exercise. I must say it's been a very interesting and often challenging exercise.

I'd like to introduce two other members that are not giving presentations in this morning session. Jim Frank who is a -- stand up Jim -- is a member of the working group, and Doctor Tom Nartker from UNLV is also a participant.

What I'm going to do is just summarize conclusions and present the significant recommendations for the working group. Many of these things you've already heard; we'll just go over them.

The working group recognizes that a certified LSS is a very important -- is very important for a timely submission of the license application, and a timely LSS, operational LSS is also part of the DOE's licensing strategy. We recognize that the LSS must be made operational and then certified six months prior to the LA and that the LSSARP expects the LSS to be operational well in advance for that time.

Also I want to point out, as we said earlier, the DOE records management system is a separate and distinct system from the LSS, but provides the majority of data for that system.

Our exercise on volume recognized that we have a

significantly less -- or that the estimates are significantly less than previous studies suggest and that the total number of pages entered as well as those disseminated are key cost drivers.

We also recognize from Camille and Fielden that care must be exercised in using the historic LSS record.

As a result of the exercise, the working group identified a smaller set of design requirements than were previously documented, and that the -- since the LSS is going to be operational for 50-plus years, there is clearly a need to monitor the technical issues as we move through time. This should minimize the cost over time as well.

Reevaluation of the LSS cost structure has identified significant cost reduction than previous estimates and these are due to some identification of some selected costs related to the RMS, our records management system, they've been moved over to the records management system; clearly improvements in technology and the reduction in total labor cost as a result of improved technology.

Our current estimates for the 10-year cost for the LSS is less than 80 million dollars. However, with -- we are going to consider an eighth option or a variation of options as suggested by DOE, so that may cause our cost to vary somewhat from that estimate.

MR. CAMERON: Mike, can I ask you one --

MR. CLINE: Yes.

MR. CAMERON: -- clarification there. In terms of the -- what you consider as cost for the LSS, is that --

MR. CLINE: Total.

MR. CAMERON: -- going back to the previous charts that were presented; that's the incremental above the DOE records system?

MR. CLINE: Yes, that's the incremental. Above the DOE is DOE records management system and that is total cost for the LSS; that's both development and implementation operation.

MR. CAMERON: So that's both development and implementation.

As which the charts that you had up, George, that increment was only operation and maintenance?

MR. HALLNOR: As procurement of --

MR. CAMERON: Okay.

MR. HALLNOR: -- and the implementation --

MR. CAMERON: Okay.

MR. CLINE: This is wasteland costs for the LSS in six.

MR. CAMERON: Okay. Okay.

MR. CLINE: The schedule for implementing or developing and implementing the LSS has been prepared and that was presented earlier.

Also presented were the point that we identified six LSS options and all six options have satisfied sub part (j). A full range of options has been presented in the analysis and also a set of evaluation criteria were selected to determine or discriminate among the options.

Of the six options evaluated, options five, six and seven clearly ranked higher. Options two, three and five utilized the human verification text for correction purposes, which is a significant cost driver. And options five and six offer significantly greater value to the user. Of those two, five and six, option six has a lower cost; it does not use the human verification.

Yes?

MR. METTAM: How about option four?

MR. CLINE: It fell out.

MR. SILBERG: So that doesn't have that factor, mail distribution --

MR. CLINE: Yeah. Let me show you how they rank. Option four ranks fairly low in the ranking.

MR. DAVENPORT: We're working with some ignorance over here --

MR. CLINE: Yeah.

MR. DAVENPORT: -- and that is the definition of these options --

MR. CLINE: Sure.

MR. DAVENPORT: -- the six options.

MR. SILBERG: You have to go back to one of those earlier presentations and there's that -- yeah, that's the -- it doesn't quite you give everything.

MR. DAVENPORT: It doesn't tell what's in each option.

MR. BALCOM: How closely are they related to the original alternative? When you say "working group," is this -- this is a different working group than the original NRC DOE working groups.

MS. NEWBURY: This is the DOE work -- chartered working group that has been presenting all this information all morning.

MR. BALCOM: Which has been in place for what, like six months?

MS. NEWBURY: Six months.

MR. BALCOM: Okay. And then the alternatives that the joint NRC DOE working group, which is what it was called, are they related to your six options, those alternatives, or totally out of that?

MR. CAMERON: They may be. There probably is some overlap, but I think that the NRC DOE working group was sort of a reality check at the time and I don't think that it has any viability or vitality --

MR. BALCOM: Okay.

MR. CAMERON: -- at this point, so I wouldn't worry about that.

MR. MURPHY: Well, those were management development kind of alternatives, anyway, whereas these are designed options.

MR. CLINE: Yeah, these are operational options.

MR. MURPHY: Operational options.

MR. DAVENPORT: Do you have a definition statement of these options, a --

MR. CLINE: Yes.

MR. DAVENPORT: -- published paper that says option one is defined as follow?

MR. CLINE: It's clearly in the report; we defined what they are in the report. Okay. And I can -- let me summarize very briefly the commonality of all the options.

MR. DAVENPORT: Okay.

MR. CLINE: Options two through seven are all compliant with sub part (j) as we mentioned.

These bit map images for record storage, okay, they have on line searchable headers, they have on line searchable full text and they provide for transmittal hard copy image. Okay. And that could be by mail or fax. Okay.

The differences --

MR. BALCOM: Are you reading from something we --

MR. CLINE: Yeah, I have something. I can -- if

it's all right with Claudia, I'll pass it out.

MS. NEWBURY: What's it from? I don't --

MR. CLINE: It's from one of our earlier presentations on -- LSS scenarios it's called.

MS. NEWBURY: Okay.

MR. CLINE: So just a description of the scenarios.

MR. DAVENPORT: Keep going.

MR. CLINE: The differences, okay, between two and -- of the six options there, image available only by mail or fax for scenarios two -- or options two, three, four and seven.

Okay. On line transmission of electronic images is five and six.

Human corrected OCR full text is options two, three and five -- are options two, three and five.

And machine corrected are four, six and seven.

Okay. What we did here is there is some questions that with this ranking you can see here that we have -- we first went through and ranked our criteria and in that exercise cost came out at the highest and image printing came out the lowest. In the backup material you'll see a description of this presentation; there should be a description of each one of these.

So there are some criticism as to -- with respect

to cost, cost being the major driver. So we dropped cost out and this is how the ranking comes out. You can see that five and six still rank above the others.

MR. CAMERON: Just one clarification, Mike.

MR. CLINE: Yes.

MR. CAMERON: In terms of the options on image availability, you talked about hard copy, fax, express, and you talked about five and six having on line transmission of images.

Is it CD Rom -- would you distribute images through CD Rom in three and seven? In other words, it's more than just express mail, on line transmission; you do have CD Rom transmission.

MR. CLINE: CD Rom would have image.

MR. CAMERON: And you would -- that option would contemplate distributing CD Roms to participants so that they could call up the images through playing those CD Roms?

MR. NEWBURY: I've asked them to evaluate another option which is that CD Rom is distributed on a periodic basis with all those images and then you have on line text for search for more --

MR. CAMERON: Yeah, I think that would --

MS. NEWBURY: -- recent updates.

MR. CAMERON: -- be the important option to consider, at least from what I've heard from some of the

people back at the NRC, if that option hasn't been considered.

MS. NEWBURY: I asked them to look at it because it's how we are planning to distribute the technical database within the project and eventually to the NRC and to outside parties, and that's within the next year or two. And so we'll be looking at it in terms of the technical database and see if it works sufficiently that we could use something similar for the LSS.

MR. CAMERON: And how is that different from the CD Rom scenarios that you had in existing three and seven?

MS. NEWBURY: I believe they were just strictly CD Rom distribution without any on line access. If you combine the two --

MR. CAMERON: Oh, okay. You wouldn't have on line full text access.

MR. CLINE: Full text access will -- four, three and seven. Okay.

UNIDENTIFIED SPEAKER: For the text.

MR. CLINE: For the text. Okay.

MR. CAMERON: Okay.

MR. CLINE: All the scenarios -- I'm sorry. All the options looked at full text search.

MR. CAMERON: Right.

MR. CLINE: Okay. But in options three and seven

you would only have CD Rom for the image, okay; CD Roms will be sent out for the image.

And I think what you're suggesting, Claudia, is that we have image on --

MS. NEWBURY: On line for between updates of the CD Roms, so we get --

MR. CAMERON: Oh, I see. You would always have an -- the idea is --

MS. NEWBURY: You would always have a full set available, but because of the cost of transmitting images over the line, you want to minimize that as much as possible.

MR. CAMERON: Oh, okay. Okay.

MR. BOONE: Yes, this is Jim Boone. What Claudia is basically suggesting is a combination of --

MR. CLINE: If you go to the last --

MR. BOONE: -- number six and seven.

MR. CLINE: Yeah. If you go to the last view graph of my presentation, you'll see kind of the wiring diagram of the different options. Okay.

MS. NEWBURY: And they'll be considering that in the revision to their draft.

MR. CAMERON: That's good, because this is -- you know, a lot of times you get people saying, well, why aren't they considering this, that or the other thing, and it's not

a question of whether the analysis shows that that's the best way to go, but being able to demonstrate to someone that, indeed, that has been considered.

MS. NEWBURY: Jim?

MR. DAVENPORT: This is sort of reminiscent of the multi-attribute decision based methodology of --

MR. CLINE: It's a little simpler. It's a little simpler.

MR. DAVENPORT: Let me ask you on --

MR. CLINE: Sure.

MR. DAVENPORT: On this page 9 what are these numbers? Like in column six you have the numbers 82, 63, 63, 48, 54 and so on. What are those numbers?

MR. CLINE: Okay. I'm going to let -- let's see, who wants to take that one. Jim? Jim. Okay.

What it was is the working group -- there are two parts of it. First of all, the working group went through and identified a set of criteria that would discriminate among the options. Okay. And that's what you see on the left-hand side over here. Okay.

MR. DAVENPORT: And that number in parentheses is what?

MR. CLINE: That was a relative ranking of those criteria on the left-hand side.

MR. FRANK: I'm Jim Frank. In this pretty

standard way these are the weighting (sic) factors, if you want to look at them this way, these are the scores that were generated times the weighting factor.

MR. DAVENPORT: So relative cost, it was weighted 10 and --

MR. FRANK: Right.

MR. DAVENPORT: -- and option of risk weighted seven.

MR. FRANK: Right. So cost is more important than operational risk, operation risk being can you really put a system together that will do this.

MR. DAVENPORT: Uh-huh.

MR. FRANK: And down here, for instance, as Mike pointed out, to take away the bias of cost, we put another set of numbers together; it was just a limited cost number.

MR. CLINE: We just subtracted the cost numbers.

MR. FRANK: And it says --

MR. DAVENPORT: So how did you derive at the numbers of the cost --

MR. FRANK: Even without it you still come out the same place.

MR. DAVENPORT: How did --

MR. SILBERG: Which cost numbers did you subtract, the first two categories?

MR. CLINE: Yeah, the first two, the top two.

MR. DAVENPORT: How did you come up with the numbers in the boxes?

MR. CLINE: The values.

MR. FRANK: Oh, these.

MR. DAVENPORT: Yeah.

MR. FRANK: We all sat down as a group and used, I guess you would call it the Delphi technique. We sat down in a group, everybody rated each one of the options independently, then we made everyone stand up and justify why they rated it that way. We talked about it as a group of people and came to a single number that represented the evaluation of the community relative to that particular factor.

MR. ECHOLS: Could you also go into the fact that they started off with maybe two, three times that number of --

MR. FRANK: Oh, yeah, yeah.

MR. ECHOLS: -- and when it was down.

MR. FRANK: Yeah.

MR. CLINE: We had 20-some factors or criteria. But we found many of the criteria we had would not differentiate among the options.

MR. DAVENPORT: I'm sorry?

MR. CLINE: We started out with many more criteria over on the left. Okay. But as we went through the

exercise we found that they would not differentiate among the options, so those were eliminated. These are the only ones we could come up with for our deliberation. Discriminate.

MR. DAVENPORT: You don't have lawyer useability on there; I can see that.

MR. FRANK: Well, that was -- we did. We did. As a matter of fact, we have one of those fellows with us, Stan over there. And --

MR. DAVENPORT: Uh-huh. Oh, well --

MR. FRANK: Oh, yeah. No, no, seriously.

MR. DAVENPORT: Yes.

MR. FRANK: There was some very, very serious discussions about responsiveness. There was some very serious discussions about images. For instance, you know, most of us were of the opinion when we started before we got legal advice and counsel, that everybody would want to see images and we discovered that most lawyers, most people who were in the system are most apt to search text first and when they finally get around to wanting an image, it's because we're going to court or something. But so the ideas did evolve as we talked to the community and particularly Stan relative to how the data would be viewed and how it would be used.

So, yeah, there was -- and responsiveness was one

of them; search accuracy was another one which you're going to hear more from Doctor Nartker later on. You know, all of these factors were bias based upon how we believe the community would use it.

MR. MURPHY: And you assigned the most importance to -- over data accuracy or responsiveness or anything else, cost was the --

MR. FRANK: Right.

MR. MURPHY: -- most important factor.

MR. FRANK: Right.

MR. DAVENPORT: Yeah, but then they did a permutation without it.

MR. FRANK: No, we took it out.

MR. MURPHY: Well, I understand that. But in the first cut cost was the most important factor.

MR. FRANK: I guess because --

MR. CLINE: The feeling was that data accuracy was the improved technology. It was not -- you know --

MR. CUMMINGS: Mike, excuse me. I think this is a very important point here. What they're reading into that chart is that cost was the most important factor of the LSS. That's not it at all.

MR. MURPHY: No, I know. I'm not reading that in.

MR. CUMMINGS: It's a discriminating -- it was the most important discriminator between the options and --

MR. MURPHY: I'm not reading that in. Well, all I'm trying -- all I'm -- I'm just going through an exercise here to see if we can't change Jay Silberg's vote from 1988. I think Jay wants to move to reconsider this.

MR. CAMERON: Jay, were you going to suggest that? Maybe we'll get some champagne.

MR. MURPHY: Well, we're going to reach a consensus here by the end of today about that.

MR. FRANK: Okay. But anyway, that's how we did it. We did it -- just to go it one more time, is people in the group evaluated them, each one of these factors independently. We argued about them among ourselves, cussed/discussed and then Jim Boone normalized them and multiplied them by the factor, the weighting, and that's how these numbers occurred.

MR. CLINE: Well, we normalized this way.

MR. FRANK: We normalized this way. We --

MR. CLINE: And we also normalized this way.

MR. FRANK: We were on a normalization kick on that.

MR. SILBERG: By the way, you defined all of your evaluation factors except the last one, your backup material. I don't know if that was intentional or not.

MR. FRANK: This one?

MR. SILBERG: Yeah.

MR. FRANK: Image printing. I don't know why it's not defined in there, but it's a relatively -- if it was relatively unimportant, so I guess --

MR. SILBERG: Right.

UNIDENTIFIED SPEAKER: Have the ability to print the image at the station.

MR. FRANK: Yep.

MR. CAMERON: I've heard this twice today and it may not make any difference at all, but I keep hearing people talking about this system like it's primarily or almost solely going to be aimed at lawyers, that maybe we can make a lot of jokes about how you just define the user interface, if it's aimed at lawyers, but I won't go into those.

Part of the purpose of the LSS that's set forth in the supplementary information was going to be the use free license application by not only the legal folks, but also the technical and policy folks to not only prepare for a hearing, but also to do their ordinary job. And I guess that I would just want to say keep in mind that, you know, if there are significant decisions being made that are being made one way because of the fact that lawyers are the primary audience, think about that and --

MR. MURPHY: I think that we concur on that.

MR. CAMERON: And also I just would want to make

sure that if you made a decision about the availability of images based on the fact, well, lawyers don't want to have images, that may be an example of something you want to think about and maybe the technical folks at DOE and NRC, Nevada and Nye County, maybe they're not worried about availability of images. But I would just say don't put too fine a point, from my point of view at least, on the fact that this is a lawyers-only system.

MR. CLINE: It was overstated.

MR. FRANK: That's a very good point. But if you'd just look at the scoring for a minute, we thought bringing in images were very important because this is basically, you know, these two are the high scorers and they scored maximum number of points in --

MR. CAMERON: That sort of washed out, then --

MR. FRANK: Yeah.

MR. CAMERON: -- that distinction.

MR. FRANK: And we also, we really learned probably for our own benefit, or at least mine, that there was a greater interest in using text for searching to identify things than perhaps I thought.

MR. CAMERON: Yeah, that's true.

MR. FRANK: But the reference people moderated my ignorance.

MR. BALCOM: Tell me what operational cost risk is

because it seems to be a substantial number -- would make a substantial change between options five and six, that second item.

MR. FRANK: Yeah, it was --

MR. DICKERSON: Should I just read it?

MR. FRANK: Yes, that's a better way.

MR. DAVENPORT: Yeah.

MR. DICKERSON: Fielden Dickerson. Operational cost risk, this is the risk that the operational cost will be higher than the original estimate. Operational cost estimates are based on assumptions. Operational cost risk considers whether it's a high, medium or a low cost impact if one or more of the principal assumptions associated with each option should prove to be incorrect during the operation phase.

MR. BALCOM: So adding human verification then would --

MS. KERRIGAN: But human verification, you could make an estimate, but the chances of you getting that right are not very good. So when you go operational, it's very risky that today in 1994 that we could do a good estimate of that, whereas, we know technology, things that don't involve human beings will improve, those costs would go down so that the operational risks of that your cost elements were correct are more safe, if you will, than things that would

involve human beings.

MR. DAVENPORT: That's a real --

MR. MURPHY: Doesn't that make you feel good?

MR. DAVENPORT: Oh, it makes me shiver to even hear that.

MS. KERRIGAN: Right.

MR. DAVENPORT: We strongly abdicate a human-verified system. To abandon that is we're not prepared to do.

MR. CLINE: Well, may we suggest that you at least wait until you hear the presentations from the UNLV.

MR. DAVENPORT: Yes.

MR. CLINE: Okay. Because --

MR. FRANK: But let me -- I'm Jim Frank again. Just --

THE REPORTER: I got you.

MR. FRANK: Okay. Well, I've suffered through this gentleman's problem several times so I'm sensitive.

The term human verification in this sense means that after a document is scanned and the OCR system has turned it into text, then we will look at the text to correct whatever errors occurred in OCR. The original document image is exactly as the image would be.

And I think you're right. I think Doctor Nartker's experiences with the ability to retrieve documents

successfully even with some of these idiosyncracies that could occur in a conversion and the accuracy that is currently reflected in the technology will probably make you a little more comfortable that you will be able to find the documents you're looking for with text search that relate to whatever you're trying to -- whatever you're trying to search for.

MR. BALCOM: So this is just cleaning up OCR?

MR. FRANK: Yes.

MR. CLINE: OCR text.

MR. FRANK: It's -- that's what it means.

MR. CLINE: Remember, the image is always going to be the image.

MR. BALCOM: All right.

MR. CLINE: Okay. Whether you get it on line or by CD Rom or whatever.

MR. SILBERG: Is there a distinction made between verification for documents that are created electronically versus those in the reprocessing? Do you need that kind of verification for the original -- for a document that's composed electronically and just automatically --

MR. GANDI: I think you need the process to assure that you're -- that that text copy follows it or matches that image copy when it's imaged and that's where we haven't --

MR. CLINE: Yeah.

MR. GANDI: -- really put in the system yet.

MR. CLINE: In concept it sounds easy, you know, you generate the document on the disk and you transfer it over to the records management system. But what happens, in fact, are there are changes as it goes through concurrence and it's questionable whether you get the, you know, whether you're getting the original text or not. So the procedures aren't in place yet for controlling that. Ultimately DOE would like to have that.

Moving on --

MR. SILBERG: By the way, when you -- this backup information I just noticed two of your definitions, two of your features are not defined.

MR. CLINE: Okay.

MR. SILBERG: And you might want to just make sure you have the right -- versus the right --

MR. CLINE: All right. Recommendations. We had a number of recommendations. We feel that these are the most significant to DOE. We believe that DOE should immediately proceed toward revising the systems requirements document, which is, in fact, being done now, and conducting the analysis of benefits and cost, that's the make versus buy.

We also recommend that options five and six be given further consideration for development. Both have

greatest benefits to the user and considerable flexibility for expansion and development.

We also feel that DOE should proceed expeditiously in converting the current microfilm based records management system to an image system.

MR. SILBERG: Is that first recommendation going to be expanded to include this CD Rom distribution with on line --

MR. CLINE: Yeah.

MR. SILBERG: -- for the increment?

MR. CLINE: That was our recommendation in the report.

Now, since the DOE have reviewed, they have suggested considering another option. So we need to go back, consider that and do another ranking. Describes that one and do another ranking. Does that answer your question?

MR. DAVENPORT: Mike, when we had talked about using the system, before we talked about the ability to mark documents, to interact in the documents so as to use them in case preparation. Will there be any live access into the system to be able to use it, or are talking all about downloading the information?

MR. CLINE: You mean like a LEXIS/NEXIS type of system or a cut and paste or --

MR. DAVENPORT: Well, no. I was thinking of it

being -- of the system being interactive.

MR. CAMERON: It was sort of the idea, I think, at one -- and this may be in the -- if you're setting out your requirements documents, that an ability for a user to have a tailored search, in other words, if they're using documents X, Y, Z, G all the time, that instead of that user, for example, saving those and using them on some sort of personal full text system, will the system provide the capability for a user to come back in and use that sort of tailored search again?

MR. DAVENPORT: Right.

MR. HALLNOR: The answer is yes. Yeah, that's the short answer.

MR. CAMERON: We're all hungry. That's all we want are short answers.

MR. BALCOM: One last question.

MR. CLINE: Yes.

MR. BALCOM: At page 4 one of the conclusions is the working group identified a smaller set of LSS design requirements than previously documented. Will we get a chance to see, you know, more of what that means? Is this something other than just real simple things you've already talked about?

MR. SILBERG: I think that's the document. This is what we'll provide you.

MR. BOONE: This is the systems that will function. The functional requirements document, we spoke about it previously, about all of that information.

MR. BALCOM: Okay.

MR. BOONE: Including the ability to go back and reuse your queries.

MR. CAMERON: Maybe all that means is that some of the things that Camille pointed out that were in the SAIC --

MR. CLINE: Exactly.

MS. KERRIGAN: Right.

MR. CAMERON: -- report are out at this point.

MR. CLINE: Yes, exactly.

Any other questions? Claudia?

MS. NEWBURY: That concludes everything that we have to say. I wanted to reiterate that everything that was presented is still draft; it has not been approved by DOE yet. We expect to have a final document out in January and want to distribute it to the LSSARP or give it to John for distribution for the LSSARP as soon as it's completed and approved.

MR. CLINE: I'll get copies of what I was reading from.

MR. HOYLE: Thank you very much, Claudia. All of your team did a super job of presenting their material and responding to questions. Let me make a couple quick

announcements before we break for lunch.

First is please return at 2:00 o'clock. We'll start with Tom's presentation promptly at 2:00.

There's an agenda change. I've been asked to swap the last item of today and the first item of tomorrow. So that would be talking about the LSS Topical Guidelines at the end of the day, today. Let me ask Roger whether he thinks that would be okay.

MR. HARDWICK: That would be great. No problem at all.

MR. HOYLE: As far as you know, Dennis will come in the morning also?

MR. HARDWICK: Yes.

MR. HOYLE: Because I know he's interested in both of those.

All right. The third is after hearing discussion today about when we should meet again, I'd like all of you to start looking at your calendars for dates for the next meeting, looking at mid March. So this would be the weeks of March the 13th, the 20th, March 27th and maybe even the first week of April the 3rd.

MR. SILBERG: Where would the meeting be?

MR. HOYLE: I guess we would need to talk about where that should be. If there's a reason why many of you are coming East in that time, we would have that one there

in the East. We had talked earlier about having every other meeting there. But let's talk about that at the end of the day tomorrow where. Okay. So start looking at your calendars.

If any of you know of meetings that are going to be taking place that we could team up with --

MR. MURPHY: I can do that just --

MR. HOYLE: -- the day before or after.

MR. MURPHY: We just went through the six-month interaction scheduling process last Tuesday down, you know, here in Las Vegas. So I can give you all of the formal NRC/DOE interactions that were scheduled between January 1st and June 30th.

MR. HOYLE: Okay. That would be very good, because I always find I'm in conflict with somebody's meeting at one time or another.

MR. MURPHY: I'm going to take care of you, John.

MR. HOYLE: Okay. 2:00 o'clock, please.

(Whereupon, at 12:58 p.m., the hearing was recessed to reconvene this same day at 2:10 p.m.)

AFTERNOON SESSION

MR. HOYLE: We'll start the afternoon session. The first item we have is a briefing by Professor Tom Nartker. The title of it is "Progress Toward an Efficient LSS, Overview of Work at UNLV." Tom has handed out some copies of his slides. So at this point I'd ask Professor Nartker to lead us through.

DR. NARTKER: Thanks, John. As John said, I'm Tom Nartker and I teach at UNLV in the computer science department and starting at about 1990 a group of us at the university undertook the task of trying to study and improve LSS critical technologies and to establish experimental research facilities and hopefully to save some money and improve performance of the LSS.

So what I really want to talk about today is how have LSS critical technologies evolved in the last four years or so, how will these changes affect LSS cost and what UNLV proposes to try and lower costs even more.

Of course, what our critical LSS technologies? As we all know, they are optical character recognition and information retrieval, text retrieval systems. What do we mean by critical? In OCR what is critical is, of course, OCR accuracy.

And the key question is what accuracy is achievable, how good an OCR device can we acquire for the

LSS. And the key question in text retrieval systems is what accuracy is needed; how good must OCR output be for text retrieval systems to maintain effectiveness; how will you notice as a user of the LSS, if there is noisy data that's not cleaned up, how will you notice when you do a query, what will you see and will you be affected.

There actually is two aspects of these questions. One is how good is OCR technology; what's the best you can do. And another aspect to the question is who produces it, what device is it, what manufacturer should I buy.

The same thing is true of text retrieval, what accurate -- not only how good can I do, but what text retrieval system is the best and, especially perhaps, what is the best on LSS on DOE data. Maybe different OCR systems do better on other kinds of data that are not characteristic of LSS-type documents. We want to know not just what is the best achievable, but especially we'd like to know what's really the best for LSS documents.

So the big question, the magic question is if the accuracy needed by an information retrieval system is greater than the accuracy achievable by the best OCR device, then manual correction will be required. That's the main bottom line and it's the bottom line because manual correction is very expensive.

So the questions I am going to talk about are in

1990 what was the best OCR device; what accuracy was achievable in 1990; what accuracy was needed by text retrieval systems in 1990 or '91; and what were perceived costs of closing the gap in 1991 or '90; and what's achievable today; what accuracy is needed today and what can we say about the changing cost between 1990 and 1994.

I should say before I leave that slide that we say right from the start that in 1990 there was virtually no solid information whatsoever indicating what accuracy were needed by information retrieval systems. No one had ever done an experiment. No one had any notion of what -- there was some speculation about it. There was a lot of talk about the problem but, in fact, there had been no real research on that question. There was nothing known.

So, in fact, in the early SAIC documents and published in several places the study that SAIC did in 1988 and '89, they had to assume a number and the number was dictated pretty much by sort of by folklore; they chose the number 99.8 percent. And I think that's sort of driven a little bit by folklore that we perceived that we needed 99.8 percent character accuracy in order to maintain reasonable retrieval effectiveness, and I think that was sort of a number that was accepted in the government communities at the time as what could be achieved if you did manual -- OCR plus manual re-key of documents. The exact history on it I

don't know, but at least it is quite clear that what was perceived to be needed in 1990 was 99.8 percent.

Well, what was achievable in 1990? In 1990 UNLV purchased this box right here. This is a Calera Recognition Systems RS-9000. In fact, it was pretty much the best OCR product available in 1990. We paid 22,500 dollars for this box. It's accuracy we very carefully measured on a fairly large sample of LSS documents at 97.4 percent.

So the accuracy perceived to be needed in 1990 was 99.8 percent and the accuracy achievable was -- using this Calera RS-9000 was 97.4 percent and that 2.4 percent difference was perceived to be very expensive. Remember that what was needed in 1990 was assumed. We very carefully measured the accuracy on LSS documents for the RS-9000.

So what can we say about the cost of closing that gap? Well, in fact, that subject was studied fairly thoroughly by Lois Dickey who was part of the SAIC project and she published a paper in 1991 where she documented in 1990 or 1989 dollars the cost of labor for converting a typical LSS page assuming a 99.8 percent accuracy requirement, and she said the cost of preprocessing and scanning was 33 cents a page and zoning the images was 62 cents a page, a quality control step was 28 cents a page.

The important step here is this $2.56 a page, is this number right here and if, for example, if you assume

the current estimate from the working group report of about 17 million hard copy pages will have to be OCR'd in the LSS at $2.56 a page, to close that gap is 43 and a half million dollars. So just to close that gap is no trivial amount of money. Just that step alone is 43 and a half million dollars, and those were 1989 or 1990 dollars, probably costs more now. Furthermore, I think Lois Dickey's report was accepted then to be fairly conservative.

Now, we are going to talk about how we can lower this and how we at UNLV have tried to lower these costs and I'll tell you ahead of time that we, in fact, have made an attack on both of those, on this cost and that cost as well.

MR. HARDWICK: Excuse me, Tom.

DR. NARTKER: Yeah.

MR. HARDWICK: That 43 and a half million dollars, was that just based on the 2.4 percent difference? Is that what --

DR. NARTKER: Yes.

MR. HARDWICK: -- it costs for the difference?

DR. NARTKER: Yes. You wouldn't have to do the step at all if you had output that you perceived to be satisfactory, wouldn't have to do it at all.

MR. HARDWICK: So that 2.4 percent cost 43 and a half million dollars.

DR. NARTKER: That's correct.

MR. ECHOLS: Now, wait.

DR. NARTKER: That's one way to look at it, yes.

MR. ECHOLS: All documents, once you have the gap you have to do all documents.

MR. CAMERON: It's an all or nothing, right?

DR. NARTKER: Well, that -- we're --

MR. ECHOLS: In other words, it wasn't 2.4 percent of the documents had to be done at 43 million dollars.

DR. NARTKER: No. That's correct. No.

What Stan says is correct. It is not 2.4 percent of the documents; that's correct.

Well, so the big question now is what accuracy is perceived to be needed in 1994 and what accuracy is now achievable in 1994 and what we can say about the cost in 1994. And to start off with I should say the answer to the achievable question has two parts, has two answers.

First, technology has progressed, sort of speak on it's own, and in 1994 this same company, Calera Recognition Systems builds a follow-on product to this. It looks very much like this, as a matter of fact. It is now a shrink wrapped software package. It now costs 595 dollars, not 22,000 dollars and, by the way, it's 98.5 percent accurate instead of 97.4 percent accurate.

Now, as long as I'm on the subject, I really can make you a special deal on this machine. Anybody who's

interested, I can make them a deal they can't refuse.

But in fact it's true. The current technology is the shrink wrapped product is noticeably improved and noticeably cheaper over the original system. So, in fact, accuracy has improved on its own.

What else can we say about accuracy? Well, as you'll see tomorrow if you come take a look at our lab and come see the demo which we'll conduct tomorrow afternoon, UNLV has spent quite a lot of effort not just learning to do experimental research with these technologies but, in fact, automating the task of doing that experimental research. And all of these measurements of accuracy of OCR systems that I have described and you'll see tomorrow are completely automated and, in fact, the 97.4 percent that we measured here in 1992 on a sample of 460 pages all done on LSS prototype database contained 817,946 characters. The Calera RS-9000 in 1992 made 21,000 errors. The upgraded 1993 Calera product we also measured carefully in between and it made 16,000 errors and improved to 98 percent accuracy. In 1994 Calera Words Grand 3.0, this product, in fact makes 12,459 errors or 98.5 percent accuracy.

But during this time each year our group has tried to conduct its own set of experiments to see in what way we might make simple improvements on these technologies and lower the error rate and thus the cost even more, and one of

the things we have done is built a voting machine.

What we do is we acquired the best OCR engine from each of four or five different vendors, we hook these things together in parallel and hand these images to all five engines at the same time, synchronize the character output, and in the beginning we did simple majority voting. If three out of four systems said it was an "A," then we defined our engine as saying the answer is an "A." Okay. And we had our own juristics to break ties and, indeed, we've done this for three and a half years straight we have built an ISRI voting engine.

And each year we have been successful at reducing the errors -- the number of errors made by the best participating engine by almost 50 percent. In this case it's perhaps more like 40 percent. We reduced -- this engine is a participant in the ISRI voting machine. And by using our voting scheme that's been improved with a set of juristics, we actually are able to produce an OCR engine, that we'll show you running tomorrow, that reduces these errors, this number of errors from 12,000 down to 7,200, which you'll notice we're up to 99.1 percent in character accuracy now. This is on DOE data, this is on LSS data and it's a fairly non-trivial amount on 817,000 characters.

You can come out tomorrow, Frank Jenkins will show you his data preparation lab. The character set we're

preparing for test next year is about double that size, will be over a million and a half characters to further verify this work.

But the point here is that 7,200 characters is 66 percent fewer errors, and while it may not imply completely of 66 percent reduction in cost, certainly the savings is proportional to 66 percent of that.

MR. SILBERG: Why is that? If you have to 100 percent review all the documents, why does the cost go down at all?

DR. NARTKER: The cost of manual work is dictated by the time that a person has to spend and if a person would have to spend half the time correcting half the number of errors, the cost goes down by half, but there's some overhead involved.

MR. SILBERG: Is the time in comparing the two, or is the time in correcting the errors? Does the gap compare the two for every entry --

DR. NARTKER: There are at least three levels of manual correction that we could talk about and perhaps it's a good time to mention the three. The original SAIC reports were actually based on a complete manual re-key of all documents based on OCR'ing the document and then just manually re-keying ever character on the document. It was not a correction step; it was a manual reentry. And the

correction step then involved comparing that manual re-keying with the OCR output and wherever they differed having, you know, a verifying person pick the correct answer.

A second way to correct output is simply to accept all the output produced by an OCR system except the characters that that system flags as being reject characters or the characters that that system flags as being suspect characters and to manually just check those characters. So you can go in and just manually check one, you know, the characters that are either rejected or marked as suspect markers. Now, obviously, you don't get near as good an output from just doing that as you would if you did a manual reentry step, but it's much cheaper.

What I'm getting to here in a minute is yet a third way, and I'll get to that, which we think will be even cheaper, perhaps will compete at least as well with the manual re-key as does the character-at-a-time correction.

So we know that what's achievable in 1994 is 99.1 percent in character accuracy.

What can we say about what's needed? Well, this turns out to be the hardest of these questions to answer and the most complex.

Let me talk for a minute and digress, sort of, about how you do research and information retrieval to

determine whether OCR error has an affect on -- affects you as a retriever of information; how do you experiment with information retrieval systems. Let me just give you a little bit of a peek at the kind of things you do and then I'll tell you what this has led us to and what we propose and what we think we could contribute to the LSS.

First, most information retrieval research is based nowadays and has for several years on the concept of relevance judgments. Relevance judgments are very subjective kinds of things, but they are widely accepted by the researchers in the field as being the best kind of metrics to experiment with these systems that exist. While their drawbacks are recognized, they are, in fact, widely used.

And the two widely used measures of performance of an information retrieval system are recall and precision. Recall is the ratio of relevant documents retrieved for a given query over the number of relevant documents for that query present in the database. Precision is the ratio of the number of relevant documents retrieved over the total number of documents retrieved. So they're slightly different. Both recall and precision take on values between zero and one, so they're both fractions.

And for a given set of queries -- of example queries, a typical test retrieval system in the research

literature is characterized by what's called a precision recall graph and it's simply for -- averaged over a set of queries. One can simply plot for a large number of queries the average recall precision that is characterized for that set of queries by that particular information retrieval system.

And the thing you might notice here is that neither precision or recall ever get very high. For example, if recall were a .5, that means that when you -- on the average when you issue a query that says I'd like to find documents that are like this or something or have this kind of information in them, that means that on the average of all the documents that you're interested in for this query you get half of them.

Okay. You'd like to think, boy, I'm going to get half the documents I'm interested in every time I do a query. No, huh-uh, the world isn't that way and it turns out that there is an enormous amount of literature in the information retrieval field to back this up. The literature's been done.

The most recent set of studies in the United States called TREC I and TREC II sponsored by the National Institute of Standards and Technology managed by Donna Harman who is, in fact, on our advisory board here at UNLV have showed over and over again that text retrieval systems,

all of them in existence, in fact, exhibited behavior that looks pretty much like that curve there.

I should point out that, in fact, what this means is that when you are doing information retrieval you are pretty much operating inside of this box since you seldom get anywhere close to even half of the documents that are in the system that you're interested in in any one query because you have to work very hard with several different queries to begin to find them all. And so precision and recall typically tends to be between, say, 15 and 25 percent in both precision and recall.

Note that this big black spot in the middle says that that's where you'd like to be; you'd like to at least get half the pertinent documents every time you issue a query and you don't come close to getting half with any available system.

So the way people do information retrieval research, and we have done quite a bit in the last three years of research with the LSS prototype documents that we inherited from SAIC and with the LSS prototype set of queries, we have conducted -- we have constructed not only sort of a ground truth version of this set of documents and issued these queries, but we've also constructed a noisy version of those same documents with OCR error in it and issued the same queries to see what would happen.

And, of course, if this line were shuffled up when we average those queries, then we would say, ah-hah, our performance has gotten better. If the line were shuffled down, then our performance would get worse.

This is a precision recall graph characterizing higher retrieval effectiveness, and obviously the same thing like this shows a precision recall graph characterizing lower retrieval effectiveness.

Well, to make a long story short, we found in conducting experiments like this with a number of different text retrieval systems using LSS data that, well, what we did we learned a lot of things. We learned some things we knew when we started, but we see them different, things that we didn't appreciate the importance of some things before.

Perhaps the simplest thing is that it is clearly not character errors that affects retrieval effectiveness in an information retrieval system, it's non stop word error. You want correctly spelled non stop words. The stop words are words in information retrieval systems that are typically not even indexed. They're words like "the," T-H-E, or "and" or "or" or "but." These stop words aren't -- they're not even stored. They may be in the document, but you would never do a query to find a document that had the word "the" in it. You would, of course, get them all. So stop words are words that have essentially no

retrieval value and are not too important.

What are important are the other words that are, we might say, non stop words. And there are typically a list of a hundred or so stop words that most retrieval systems don't even bother to store or index at all. So that it is, in fact, words that we're concerned about much more than characters.

And, in fact, in studying the output of a number of these experiments and we will show you tomorrow this. Doctor Taghva can show you some of the specific publications which we have gotten in the last couple years on specifically these problems and I'll leave it up to Doctor Taghva to tell you about -- more details about that research.

But what we've essentially concluded is something we knew when we started and that is that incorrect characters don't affect retrieval effectiveness, that misspelled words do affect retrieval effectiveness and that we should concentrate on words and not so much worry about the characters.

And so what we have evolved in the last year and a half is, in fact, a special system, a software system which we have written which has been designed by Doctor Taghva and implemented by several people. We originally called it Rummage, but the group has now chosen the name Manicure for

it.

Manicure is a -- it's a semiautomatic expert system based document editor which corrects words at the document level instead of characters at the page level, never bothers to look at characters at a time, go right straight to the document, build the document and start worrying about words, don't bother with characters. Characters we pick up when we do that. We're not going to leave any characters that are clearly in error whether they're part of a word or not. We're going to focus on doing word correction at the document level, not character correction at the page level.

And we do one more thing. If you do enough of these experiments, one of the things you find out is that OCR devices frequently generate a lot of junk characters; for straight lines, they'll generate a series of dash/minus signs and all kinds of junk characters and they quite clearly have nothing with do with retrieval of information at all. I mean, you're never going to say give me a document that has 10 minus signs in a row or something like that, you're going to ask it for a word.

So what we immediately built into Manicure is an ability to accept automatically zoned pages, because the difference between manual zoning and automatic zoning is you get even more junk when you do automatic -- when you use the

automatic zoning feature in these devices. So, in fact, that adds another 62 cents a page if you can get by with doing automatic zoning instead of manual zoning in savings.

MR. DAVENPORT: What do you mean by automatic zoning?

DR. NARTKER: Please?

MR. DAVENPORT: What is zoning?

DR. NARTKER: Excuse me. Zoning is the operation of identifying the text body zones on a page and separating them from the --

MR. DAVENPORT: The white space.

DR. NARTKER: -- pictures. No, not white space. From photographs or from graphs or from other non-text objects on the page.

MR. DAVENPORT: Marks that aren't words.

DR. NARTKER: Tables, tabular information.

MR. METTAM: Makes the forms into two columns.

DR. NARTKER: Zoning is the act of preserving the correct reading order when you have a newspaper style multi-columned document. So even if you have no graphs or photographs on a page, you're going to have to zone the left column because you want to read it down the this column and so forth.

So in order for the LSS to work properly, all assumptions have been up to now that manual zoning would be

required, and Lois Dickey's figures that I showed you said 62 cents a page and that's probably a fairly conservative number.

We have built Manicure and tomorrow Doctor Taghva will demonstrate its operation for you and talk about its design and you can ask questions and see it operate. We hope you'll all come out; please come out. We have -- our group is very much looking forward to trying to show off tomorrow because we've worked very hard at some of things.

Now, the point of all of this is that if you really can save 62 cents a page in zoning costs and a fair fraction of that $2.57, or whatever it was, a page in error correction costs by doing word correction at the document level, we believe that the effective savings from that would be something like $2.62 on 17 million pages or about a 44 million dollar savings over original projected cost.

So, to summarize, in 1990 the accuracy achievable we measured to be 97.4 percent, the accuracy assumed to be needed was 99.8 percent, the cost of correcting that in 1990 was perceived to be greater than 54 million dollars. In 1994 we can demonstrate an achievable accuracy of 99.1 percent. So we're not at 99.8 yet, but you can see we're closing in fairly quickly.

The accuracy needed, well, I haven't put down exactly what number it is because we haven't worried about

measuring an accuracy needed. What we've worried about is using actual DOE documents and actual LSS problems in building up an editor that we think we want the LSS to use.

We want to contribute this editor to the DOE. We would very much like -- we don't think there's another system like this on the market anywhere and we very much would like to have this editor be used by the Licensing Support Systems, and we think the reduced costs just of these two systems by themselves is in excess of 44 million dollars.

Now, these technologies are increasing. We have measured each year in our annual tests an improvement of, oh, probably in excess of 10 percent a year. I think the technologies probably as a rule of thumb are improving by about one percent a month. We expect that will continue and, in fact, in three days we start our 1995 round of testing. All of the vendors have been invited to submit their latest and greatest technology to us for the UNLV 1995 round of tests and the deadline for that submission is, in fact, December 15th. And the devices have started to arrive. I think we have three or four submissions as of today and there will be another three or four more in the next three days.

On the horizon the interesting thing to us is that this year for the first time we are going to be testing gray

scale OCR engines. No one has produced a commercial gray scale engine before now and so this year we're testing engines. I think we have two already in house. And we anticipate another five percent or 10 percent improvement in accuracy just based on gray scale engines, because --

MR. LEVIN: What is a gray scale engine?

DR. NARTKER: -- you use a gray scale scanner and you convert a black and white page not just into white spots and black spots in the bit map image, but you convert it into shades of gray when, in fact, there are shades of gray. You have preserved the appearance of that page much more faithfully and, indeed, have more information. You convert not into just black and white, but into black and white and all shades of gray in between at some particular resolution, and most scanners nowadays are either eight or 16 bit gray scale scanners. But we anticipate continued improvement and one obvious thing that has not been exploited so far is gray scale scanning.

There are several more things, doing research in this field, we are very aware of that are -- that we'll contribute we think yet more potential gains in accuracy in the next two or three or four years. And we -- our experience has been that even though the devices improve, our voting engine improves just as fast. We have been able to sustain about a 50 percent improvement over the best

participating device for three years in a row on DOE data by using a simple voting scheme.

So we expect that the costs are going to go down even more and we are very much hoping to be able to provide you, unless the situation changes, we think we have the best OCR device for the LSS. We've built it already. We are going to demonstrate it for you tomorrow. Okay. We think we have it and in some sense we have it by default, but at least as long as voting works, we think that we have the best OCR device that can be used by the LSS and we think we have in Manicure a post-processing editing system that we think will be more efficient than anything on the market.

So tomorrow the two things you will see are ISRI voting machine and Manicure.

MR. BALCOM: By the "best device," you mean the voting machine set up, the five scanners?

DR. NARTKER: Please?

MR. BALCOM: When you say the "best device," do you mean the voting machine process, the five?

DR. NARTKER: I don't know how many participants Steve will have tomorrow. We could demonstrate voting of, you know, anything from two to 10 devices at the same time.

MR. BALCOM: I mean, you said, you think we think we have the best device. Do you mean the voting machine, the process?

DR. NARTKER: Yeah, the voting machine process, yes. Yeah.

Yeah. Another question?

MR. METTAM: How does Manicure handle numbers, numerical data? You can't really spell check it.

DR. NARTKER: There is a component that -- special components for numbers, special components for acronyms, there is an acronym finder. There is a special component to recognize dates, to recognize -- it a huge expert system with lots and lots of rules of thumb built in about how to recognize things and --

MR. METTAM: I'm thinking more of --

DR. NARTKER: Special emphasis on identifiers, special identifiers, proper nouns and other kinds of special identifiers.

MR. METTAM: Within a document if they'd use a particular number, 4.5 billion, let us say, now obviously your system will recognize that as being a number --

DR. NARTKER: Sure.

MR. METTAM: -- but will it recognize it as being an internally consistent number? I don't know if I'm expressing that question very well, but --

DR. NARTKER: Well, would it recognize that there is an alphabetic "O" in that string of zeros and it really should be a numeric zero?

MR. LEVIN: No, that isn't the question. The question is you've got a string of numbers, 123456789, that's the way it's interpreted, but maybe it should have been 123456786; there's no way, there's no context to put it into to know that you have the right number.

MR. METTAM: It's a misspelled number.

MR. LEVIN: Yeah, it's a misspelled number with no context, no database to --

DR. NARTKER: Certainly where there is no context, that we don't know any magic answers, in terms of exactly how the systems works --

MR. BALCOM: It's probably not going to search for the numbers.

DR. NARTKER: At that level --

MR. BALCOM: Some numbers you will and --

MR. DAVENPORT: It's going to rely on the accuracy of what's in the document which is more important than the searching really.

MR. LEVIN: Then you might have to go back to the image.

MR. SILBERG: I mean, if you have a voting system, at least you can tell if the error is in the scanning process at some level of accuracy.

MR. NARTKER: We in general advocate that the LSS should -- you know, we think the LSS should provide images

to you; we think it should.

I should tell you, and I didn't give enough emphasis to the point, but the rules that the expert system uses are not only rules such as I have described, but it uses knowledge from the document level and from the collection levels. So, in fact, it keeps a dynamic list of every time it finds a new acronym, okay, the acronym finder finds a new one, it enters it into the library. Every time there is a new proper noun that's been found on a previous document, it adds it into the library. So, in fact, part of Manicure is a dynamic system that can be sort of tuned to the documents that are loaded into the LSS dynamically as they're loaded and the information that's collected is then fed back and used to correct OCR error on future documents.

MR. SILBERG: Based on your work and your inclusion that character errors don't matter, errors that's based on the -- are you looking at recommending a change in the 99.8 percent figure? Because that drives a lot of this up and if that's just a number that came out of thin air --

DR. NARTKER: We actually think that with Manicure you will not need to worry about the 99.8 percent number. We think you'll be able to -- I think you'll want to display images. But, in fact, we've done enough experiments with noisy -- when we say "noisy data," we mean a copy of these documents that have come straight out of OCR devices that

haven't been corrected, so there's incorrect characters there. We've done enough experiments with noisy data at different levels. We've taken a prototype, a set of about 600 documents out of the LSS prototype database and we've taken the ground truth characters, the ground truth of those documents and then we've passed those pages through a fairly good OCR engine and then we've passed and created another database and we passed those pages through one of the worst OCR engines we had; it had a lot -- created yet a second database. Then we picked an intermediate one and created yet a third.

So we, in fact, have done experiments with three different databases with different levels of noise in each one; one created by a good OCR system, one by an intermediate, and one by a bad. And we've issued the same set of queries to all four databases and then measured and averaged what the response has been to these queries and we have written publications on those, we have done that. That has been one of the central points our group has focused on and Doctor Taghva will talk about this tomorrow and show you some of his papers.

We have done what I've just described not just with the Basis-Plus text retrieval system which is a Bulian (phonetic) text retrieval system, but we've also done it with two others, with a Vector space model and a

Probablistic (phonetic) model with Inquiry and Smart. So we've done it with three different classes of text retrieval systems.

And I must tell you that one of the things we find is that that precision recall graph doesn't change very much; it doesn't change very much under any of those circumstances.

But, in fact, what we have found -- and I don't want to steal any of Doctor Taghva's thunder, I couldn't even -- I couldn't steal it at all because he's done the work. But what we've found is that by focusing on word correction at the document level we can, in fact, minimize any error that you see at all.

Yeah?

MR. BALCOM: I would assume that there is very little labor associated with Manicure, that it's just compute intensive?

DR. NARTKER: Okay. Manicure does sort of three kinds of corrections. It does as much automatically as it can first. It has a list of rules and juristics -- it tries to recognize dates, it tries to recognize acronyms and keeps track of acronyms and does spell checking at the document level and at the collection level. It does as much automatic as it can.

And when it can't make a decision, but it thinks

there is a problem, it will present at your option, of course, this is probably the way you would use it, it will present an operator a list of alternatives and show the picture of the word that it's having trouble with and the operator can correct it. If one of the alternatives that is listed is correct, the operator can make the correction with a mouse click very quickly.

So it does as much as it can automatically, tries to present the operator the fastest possible user interface to make further corrections, namely a mouse click.

But if neither of those is correct and the operator says, ah, no, none of the things you've got there is correct and we've got to, you know, that's not the correct word, the operator can then manually type in the correct word.

MR. BALCOM: But there will be a threshold there that Manicure will make some assumptions that would exclude most, I mean, you know, 80 percent of the words and just go ahead and do without any manual intervention or something like that.

DR. NARTKER: We're hoping to -- we have not finished our set of tests on it. We have a version of Manicure running. We can demonstrate it and we think it's fruitfully usable today. And we'd be delighted to have TRW or DOE begin to test it. We will make a copy available to

them in the next month I think. But we don't have numbers yet on how good we're doing.

MR. DAVENPORT: So you're not suggesting no human correction, you're talking about human correction with better software.

DR. NARTKER: That's correct. That's correct.

MR. DAVENPORT: And that human correction with better software you think will save DOE how much?

DR. NARTKER: Forty-three million dollars -- well, the sum of everything I've talked about, I mean, 43 or 44 million dollars over original estimates in 1990.

MR. DAVENPORT: It will be -- it could be --

DR. NARTKER: Now, to be honest with you, George Hallnor has read both of these reports and so I think the savings is cranked into much of what George has said. For example, I don't believe you have included any zoning costs at any rate. So George knows about this already. So the savings are --

MR. DAVENPORT: Already taken --

DR. NARTKER: -- already cranked in, yeah, at least most of them, maybe not all of them.

MR. DAVENPORT: Now, if the 17 million pages is reduced by pages with have been character recognized by -- into the electronic image before, like he's proposing to do now, what's your cost savings that way?

DR. NARTKER: I don't understand the question.

MR. BALCOM: If there's only 30 million pages total and it would look from here like most of that data is coming after --

MR. HALLNOR: Let me explain where 17 million comes from.

MR. BALCOM: Okay.

MR. HALLNOR: This is George Hallnor. In the model we had we run out whatever, 107-some-odd million pages, but that is not -- that's a mixture of paper pages --

MR. BALCOM: Right.

MR. HALLNOR: -- and electronic pages.

MR. BALCOM: Right.

MR. HALLNOR: And 17 million represent the paper pages, so that I gave that figure to Doctor Nartker.

MR. BALCOM: So that's already been counted in.

MR. HALLNOR: Yeah, it's all been counted in.

MR. LEVIN: I guess another thing, it's safe to assume that the more you use the system in the production mode, the better your accuracy because it's building this database; is that correct? So when you get large volumes, it gets better and better.

DR. NARTKER: We hope so, yes. Clearly at some point most of this -- you know, most of the key words that people want to use in documents will have been built up and

they'll be in the dictionary and we will spell check everything and it will -- spell checking is a big part of this correction mechanism.

Yes?

MR. METTAM: I'm looking at your precision recall graph. If what you say is true and, you know, the hit percentage is like 15 to 25 percent on recall, well, who's working on developing the query engines so that we can get a better -- you know, I mean, obviously if you have 100 percent accuracy and I'm still only getting 15 to 25 percent hits --

DR. NARTKER: Precision.

MR. METTAM: -- on my query --

DR. NARTKER: That's correct.

MR. METTAM: -- then --

DR. NARTKER: That's the state of the art and there is well over 20 year -- there's a 30 year history of research in information retrieval now and, if you could improve on that very significantly at all, you would have a very hit paper at the next SIGIR meeting.

The tests that the National Institute of Standards and Technology have run in the last three years, the TREC I and TREC II tests have documented over and over again with five and 10 and 15 different commercial text retrieval systems compared side by side on the exact same database

using the exact same set of queries and their precision recall graphs look surprisingly identical. And there are little differences, but not 15 percent difference, very small differences.

MR. ECHOLS: I want to verify one thing.

DR. NARTKER: Yeah, Stan.

MR. ECHOLS: Stan Echols for verification. One thing I'm used to, I guess, is a NEXIS/LEXIS kind of system.

DR. NARTKER: Yeah.

MR. ECHOLS: As I understand it, this differs in that there's a split screen such that if there was a number I was curious about and it was an important number, it would tell the sub parameter ground water travel time or something else, and I wanted to make sure that that was a valid number and it wasn't created by error, I could call up the original side by side --

DR. NARTKER: Sure.

MR. ECHOLS: -- to find that number --

DR. NARTKER: Sure.

MR. ECHOLS: -- and do a check.

DR. NARTKER: In fact, we are prepared not to split the screen at all, but to show you absolutely nothing but the image so all you see is the image, but we'll highlight the words on the image that we found for you.

MR. ECHOLS: Which is -- and now that's something

new in itself, isn't it, highlighting on images?

DR. NARTKER: That's correct. That's what we're going to show you tomorrow; it's part of Manicure.

MR. ECHOLS: Okay. I don't want to get ahead of the kick.

The other question I had dealt with, when we were talking about that precision accuracy curve, again, in LEXIS I have 100 percent accurate documents and still 30 percent is the best I can hope to retrieve or thereabouts if I have absolute perfect data?

DR. NARTKER: Yes.

MR. ECHOLS: So, and that is because it's the skill of the person --

DR. NARTKER: Yes is the answer and there is a reason.

MR. ECHOLS: And the reason is?

DR. NARTKER: The reason is the word "pertinent." Okay. What is considered to be pertinent, relevance. What's relevant? Relevant judgments are subjective things and people who have done these experiments for years have tried to go through fairly large databases and tried to understand, based on a given set of queries, what documents in this database are relevant to that query. Relevance is subjective.

But the experiments that have been done have time

and time again shown that, in fact, if you make some literal interpretation of quote "relevance" and accept what, you know, the way people have done this, the fact that you get less than 30 percent of the relevant documents are on almost every query.

MR. SILBERG: But if you ask for a simple minded query like, you know, identify all documents with the word --

DR. NARTKER: If you say ask all documents that have the name "Stan Echols" in it --

MR. SILBERG: Right.

DR. NARTKER: -- you'll get them all --

MR. SILBERG: Right.

DR. NARTKER: -- of course.

MR. SILBERG: So that's really --

DR. NARTKER: Of course.

MR. SILBERG: -- a question of the skill of the questioner.

DR. NARTKER: Absolutely.

MR. ECHOLS: That's the key. It's not the hardware or the software. The 30 percent you're fighting is the skill of the person working the machine.

DR. NARTKER: More or less. That's one way to look at it, yes.

MR. ECHOLS: The last question that was on gray

scale, can you give us some examples of how that's used. Is that for the photos that you would normally square out and eliminate?

DR. NARTKER: Well, if you just imagine a photograph, you know, a photograph that has the shades of gray versus a photograph where everything that was gray is converted to solid black and everything else is white. I mean, you either have all black and all white or you have some points in that photograph, some pix holes, if you please, that are some level of gray in between white and black and the images we look at have, you know, have shades of gray. So gray scale images retain more information.

Whenever you have to set a threshold in a scanner and you have to say, well, anything that's not black above this amount of blackness, call it white, okay, and you have to binderize at some level, you have to draw some threshold, then you start to throw some information away. If you set the threshold too low, then you get more black and you start to get a lot of touchy characters because, you know. So where you set thresholds on scanners turns out to be a very important part of -- it's something we're doing research on, too.

Now, there are a lot of projects that we have worked on and some of them are longer range. What we're going to show you tomorrow is what we think the LSS could

use today, today. And we will show you working software systems. They are not things that UNLV has immediate plans to commercialize; maybe some day, you know, some of these things will lead to commercial product. Right now we're interested in contributing this to the DOE. Maybe some day they'll make commercial products.

One reason why even though we're convinced that these things might successfully be commercialized today, especially for example, Manicure -- well, most systems could be commercialized today and I'll talk about them both. We see the improvement coming along too fast. We see a one percent improvement a month or 10 percent a year and better. We can see so many improvements coming that we can make improvements faster than we could possibly commercialize this, and we are interested for the time being in trying to produce better technology for the DOE.

Now, in terms of commercializing Manicure, there is no competition for that at the moment.

In terms of commercializing the voting machine, there are, in fact, three different companies who produce voting engines. SAIC makes a voting engine; they sell it for 250,000 dollars a copy, as a matter of fact. I am told they have sold 10 or 12 copies so far. I haven't bought one myself, so I don't know. In fact, TRW makes a prototype voting engine. There is one, at least one other company

that makes voting engines. So there are a number of different companies who already are commercializing the idea of doing voting. Thanks.

MR. HOYLE: Tom, before you leave, could you give us some general directions on how to get to UNLV.

DR. NARTKER: I will have the map for you tomorrow and parking stickers so that you can put a parking sticker in your car. So when we adjourn tomorrow morning at 11:30 or whatever it is, I'll have the parking stickers for those of you who drive and I'll have a map.

MR. HOYLE: Great. Thank you very much.

Let me ask. Dennis, are you going to be able to stay a little longer or are you about ready to get --

MR. BECHTEL: Maybe a few minutes. I need to --

MR. HOYLE: Do you want to take the topical guidelines item, if that's agreeable with others. Is that okay with you?

Okay. Why don't we do the topical guidelines.

MR. CAMERON: Okay. As you all know, we issued a draft reg guide, topical guidelines for public comment in July of 1993. We received comments from mostly everybody around the table and we're still pulling together the final draft of the reg guide on the topical guidelines and it's scheduled -- the staff is scheduled to have it to the Commission for review in February of next year.

And although it's still in process, I think I can summarize some of our existing staff positions relative to most of your more significant comments, and a lot of these related to some of the ambiguity of certain terms that were in the latest revision.

As you know, documents relevant to environmental issues are now within the scope of the topical guidelines. Now, that will specifically -- the guide will specifically identify two sub sets of that type of information; one is socioeconomic, one is transportation -- or the other is transportation.

In the section on information relating to compliance with various statutes, we didn't have NEPA listed; we will specifically identify that.

And in terms of a comment that's a little outside this particular aspect, if there are prelicense application depositions which contain relevant material, those will go into the LSS, they're within the LSS. Now, that doesn't get to the question of when those depositions will be taken but, if there are depositions, they'll go in.

MR. DAVENPORT: We -- I guess we had sort of presumed that, but that's good.

MR. CAMERON: Okay. And basically I think that pretty much covers most of the concerns as I understood them and that is in the existing staff draft. I don't anticipate

that the Commission is going to change that in any way.

MR. BECHTEL: Just out of curiosity, why is socioeconomics and transportation under environmental? Is that just --

MR. CAMERON: Because that's where we anticipate that that information will be most relevant, in other words, to the DOE environmental impact statement and the NRC adoption of that. And I think that's the impression, too, that I had from your -- your previous problems on it, but those are going to be the sub sets of environmental information.

And just to clarify one thing we talked about last time, which I think Jay gave a particularly good clarification on, is that we're talking about environmental information that doesn't -- that's outside the scope of information that would relate to protection of public health and safety, in other words, radiological protection. Those will be in because of Hertz-60, but what we're talking about are issues that are outside of that radiological protection universe. And I suppose, as also Jay pointed out, those could also bring in socioeconomic issues such as effects on population and things like that.

MR. BECHTEL: Will the board have an opportunity to review the document before it goes to the NRC?

MR. CAMERON: If the --

MR. MURPHY: We're already done that.

MR. BECHTEL: Well, I mean --

MR. MURPHY: Didn't we?

MR. HOYLE: We looked at the draft.

MR. CAMERON: Yeah, we looked at the draft.

UNIDENTIFIED SPEAKER: Looked at the draft, comments have gone back and --

MR. BECHTEL: So that what you're describing, the change is just a change in the --

MR. CAMERON: You know, I -- well, there is a number of -- I mean, there's other comments, but they're more clarifications rather than substantive, if I can characterize those as substantive. If the panel did not feel comfortable with -- if they would like to see it another time, I don't see why it would necessarily need to be harmful to the staff's schedule to get it to the Commission.

MR. MURPHY: It's been so long now. What's another year or two?

MR. CAMERON: I left myself wide open for that one.

MR. MURPHY: You sure did.

MR. SILBERG: I don't know. Does anyone feel like they want to see it before it goes up?

MR. BECHTEL: I'd like see a copy.

MR. SILBERG: Is it a problem to send it to us simultaneously when it's going to the Commission? It will take awhile for the Commission to get around to it, is that what you're saying? If there were any problems that we saw or additional corrections, we could still get them into the staff and the staff could still implement it, just so we don't hold it up yet some more.

MR. HOYLE: I think that would be appropriate.

MR. CAMERON: We can either do that or, if the schedule permits, send it out beforehand so that we can get all that up to the Commission at one time. Okay.

MR. DAVENPORT: Chip, to drop back to Dennis' question about transportation and socioeconomics, on socioeconomics I certainly understand the distinction you're making. On transportation that is an issue of concern to the Commission, is it not, within its radiological health and safety concerns as well as the non nuclear aspects of transportation, isn't it? I mean, you care about that outside of NEPA, don't you, transportation?

MR. CAMERON: Well, that's an interesting issue. Transportation has traditionally been excluded as an issue except for as its encompassed by --

MR. DAVENPORT: Yeah, but you've also not --

MR. CAMERON: -- NEPA with the licensing of specific facilities.

MR. DAVENPORT: You've also failed to consider nuclear waste disposal in your licensing and look where we are on that. I mean we presume and will continue to presume that transportation is a citing issue, is relevant to reasonable assurance in the Yucca Mountain case. Now, is that a reasonable expectation?

MR. CAMERON: I think that the jury is still out on that. The Commission has noted many times that the topical guidelines don't have anything to do necessarily, in other words, they don't prejudge what's going to be admissible as a contention or not --

MR. DAVENPORT: Yes, I understand that.

MR. CAMERON: -- to proceeding. But still it's possible, I suppose, under the existing case law for a contention to come in on transportation that would not be already encompassed within the scope of the EIS, but that particular point --

MR. DAVENPORT: Well, I don't mean to suggest that the EIS might be the appropriate vehicle to collect the information and bring it to your attention; that's just a matter of form. But it would seem to me that it is substantively an issue of merit in considering whether this license should be granted. Yes?

MR. SILBERG: Well, I think what Chip said is that's an issue which will undoubtedly be determined, but

not now and the reg guide really is there for a totally different purpose. I think the information will all be there because it's all going to be picked up.

MR. DAVENPORT: But Dennis raises sort of an interesting question. If you're going to say that's just -- just an EIS issue, just a NEPA issue, which is a way of saying that it's not an issue of concern to the Commission because of radiologic health and safety, that begins to concern me that you're going to find a way to categorize that out of the proceeding and --

MR. CAMERON: It does nothing in the topical guidelines, categorizes that out of the proceeding at this point. I guess my first concern would be can you think of information that would not be relevant to the EIS on transportation that might be relevant to transportation being considered as an Atomic Energy Act --

MR. DAVENPORT: No. Because I think that NEPA --

MR. CAMERON: -- issue.

MR. DAVENPORT: -- is more encompassing.

MR. CAMERON: So at least we'll have the information in there if the licensing board makes a determination then, a contention on Atomic Energy Act.

MR. DAVENPORT: I would answer your question the same way with respect to the question of whether there's a solution to the nuclear waste problem before you cite a

reactor, but you have found a convenient way to set that -- categorize it out your decisions.

MR. SILBERG: Well, except they did that by --

MR. CAMERON: Yeah.

MR. SILBERG: -- generic rule making.

MR. DAVENPORT: I know. I know.

MR. CAMERON: But, Jim, I guess the bottom line is is that the Commission has not affirmatively said that transportation as an Atomic Energy Act issue --

MR. DAVENPORT: Uh-huh.

MR. CAMERON: -- as opposed to a NEPA issue --

MR. DAVENPORT: Yes.

MR. CAMERON: -- is within the scope of this licensing hearing.

MR. DAVENPORT: Uh-huh, and I guess that's what concerns me. I would hope that they would see that it is an Atomic Energy Act issue as well as a NEPA issue and that we not create the assumption otherwise in our early documents that aren't binding.

MR. CAMERON: At least not create the assumption otherwise.

MR. DAVENPORT: Right.

MS. NEWBURY: We've got a comment back here?

MR. FLUM: A question. My name is Chris Flum from part of the MNOT. I'm curious as to why you've decided to

put NEPA-type information in the LSS. I was going back and I read through the Commission's amendments to part 51, which is your environmental regulation, and I had the quote in my suitcase. It says, "The Commission firmly believes that the Nuclear Waste Policy Act was intended to have all matters associated with environmental impacts that were plausitory development considered and decided to the fullest extent possible apart from NRC licensing proceedings."

If NEPA issues are supposed to be decided apart from NRC licensing proceedings, why would you want NEPA-type information in the LSS?

MR SILBERG: Well, there's that phrase that's in there, "to the fullest extend possible." To the extent that it isn't possible --

MR. MURPHY: There's also a requirement in the Nuclear Waste Policy Act that the NRC adopt DOE's EIS to the extent practicable, and that phrase, "to the extent practicable," will be the subject of the hearing during the licensing process.

MR. CAMERON: And whether significant or new information.

MR. MURPHY: That same part 51 assigns to the licensing panel, the repository licensing panel the question of determining whether or not it is practicable for NRC to adopt DOE's EIS. And the only way they can do that, the

only way they can have the information before them and necessary to allow the parties to make whatever contentions they have with respect to practicability is to have the NEPA information in the LSS. Otherwise, you've got a three-year license application time line for a repository license and the NRC can say, okay, we will grant you a repository license if 17 years from now when we conclude our EIS hearing we haven't found new information, because the EIS hearing is going to be conducted by hard copy and discovery. That's the alternative.

MR. FLUM: There's going to be an EIS proceeding to decide whether or not the NRC --

MR. SILBERG: It's all part of the same proceeding.

MR. CAMERON: It all is part of the same proceeding, but the --

MR. MURPHY: It's part of the same proceeding.

MR. CAMERON: The licensing board panel has to make a decision on whether it's proper for NRC to adopt the DOE EIS. There's provisions.

MR. SILBERG: And to the extent that NRC doesn't adopt it and does things differently, you know, has a supplement which it includes as its take on some issues where it feels it can't adopt, those all become issues to litigate. All that stuff has to be in the LSS.

MR. DAVENPORT: I think the shortcoming of your assumption is that NEPA doesn't have decisions. NEPA is a process, not a decision. And the process is used in the NRC's ultimate decision or in DOE's decision, but there are not separate NEPA decisions.

MR. FLUM: Forgive me; I'm not an attorney. But when I read part 51, it clearly seems to me that the NRC wanted to keep these environmental issues out of the licensing proceeding and stick to radiological health and safety.

MR. MURPHY: There were certain members of the NRC staff who tried that ploy and we bloodied their noses in a big meeting in which there was also just shouting and screaming over that issue in October 1990 and we won and they lost and let's not reopen it now.

MR. CAMERON: And they're all working out at the Yucca Mountain site.

MR. MURPHY: Right.

MR. HOYLE: Any further discussion of the topical guidelines?

Okay. Let's then turn to the last item for the day. On the agenda it's listed as "Operation of the LSS," and we would like an update on discussions DOE and NRC have been having about the COGR proposal, and Claudia has volunteered to lead this discussion.

MS. NEWBURY: I notice that it's an hour worth of discussion, too. I don't know if we're going to really go that far.

MR. HOYLE: We haven't been precise today on our timing.

MS. NEWBURY: Well, I feel a lot more intelligent than I did in September; let me start off with that.

We've been discussing the COTR option for the operation and maintenance phase of the LSS. As I understood it, that option came about because of a problem in terms of who was going to fund the system and how it's going to be operated. And I took the option of DOE letting a contract and NRC acting as the COR -- there is no such thing as a COTR now, I guess -- to my procurement officer and to our lawyer in the office and we talked about it and they were not at all comfortable with that in terms of DOE not having control -- having the responsibility of the contract and not really having control and if NRC was the representative.

And I came back with a second option which was a cooperative agreement where again DOE would be letting a contract and NRC would be overseeing the cooperative agreement. And again they said, no, that's not a really good idea because again DOE is responsible as the contracting officer and we don't want to put the responsibility for evaluating the contract with another

agency.

So they suggested an interagency agreement or a memorandum of understanding whereby DOE would appropriate the money from the waste fund and transfer it to NRC who could then contract or cooperative agreement or do whatever they choose. We explored that between the procurement officer here at Yucca Mountain and NRC and the two lawyers have spoken and everyone seems to think that this is the best alternative. I'm told actually this is an alternative that was discussed earlier as a potential option and I don't know why it disappeared, but it seems to be the best and I guess you guys can all explain to me why it's not going to work.

MR. CAMERON: If I can -- can I add something when you've done?

MS. NEWBURY: Sure.

MR. CAMERON: Claudia, just to amplify on that, it would -- using the interagency agreement model would be what we contemplated all along, which is the mechanism for -- and it would serve purely as a mechanism for funds to be available to NRC to operate and maintain the system. So this really goes back to the model that we had way back when on this and there would be no -- it would be fairly simple and there wouldn't be any doubt about the fact that the LSS administrator was operating and maintaining the system.

And I think some of the DOE concerns about the COR were not concerns that we absolutely can't do this, although if the interagency agreement wasn't available, we might have to go back and test that. But they were related to the fact that you are going to inevitably get into complications about what the scope of authority of the COR, i.e. mode 11, and the DOE contracting officer authority. So why run the risk of having potential complications that could really fowl up implementation or operation of the system. So where we are now is getting in the point of trying to initiate an MOU to put this IAG in operation.

MS. NEWBURY: We're not quite there yet. We have to make sure that it's amenable to the people who are in charge and actually are writing the checks, but --

MR. CAMERON: Yes. That's an important point, too, is that all of these mechanisms, the COR, the IAG, whatever, are all dependent, of course, on DOE agreeing to provide the funding mechanism. Okay.

MR. DAVENPORT: May I?

MS. NEWBURY: Please.

MR. DAVENPORT: John, remember the last meeting when I said I thought you guys were going to build it, operate it, it was going to be an NRC system and you said I was wrong, and I'm prepared to be wrong. But aren't you just now saying that if we go back to square one that NRC --

that DOE is going to pay for it, NRC is going to at least operate it.

MR. HOYLE: Talking about operation --

MR. DAVENPORT: So the --

MR. HOYLE: I'm not sure --

MR. DAVENPORT: So the only aspect of my memory was the building part that was wrong.

MR. CAMERON: That's right.

MR. DAVENPORT: Okay. So if we're that far, if we're to the point that you're prepared to acquiesce the management, the administration in NRC with DOE picking up the tab for it, why not hand the building over to them as well? You're prepared to pay for it.

MS. NEWBURY: Well --

MR. DAVENPORT: Why can't these guys make the decisions on the building system?

MR. CAMERON: Yeah, the rule specifies --

MS. NEWBURY: The way the rule is --

MR. CAMERON: We as design and development and --

MR. DAVENPORT: Well, that's because you guys thought that you might have to pay for it at that time.

MR. CAMERON: No. It was based on the fact that DOE had already spent a lot of money on design of the system.

MR. DAVENPORT: True.

MR. CAMERON: Now, as we've seen the SAIC design may have had some warts on it, but it still was a fairly good document. But I think that the decision was made that since they were already starting to do this just let them continue to do it, particularly as it relates to some of the DOE design and development of their own records management system as a building block.

MS. NEWBURY: Now, if you bear in mind that the LSS cost, as we shown all morning, as an incremental cost on our records system that we're developing anyway, it makes sense for us for us to develop the system and then --

MR. DAVENPORT: I can see that.

MS. NEWBURY: -- turn it over to the NRC to operate and maintain.

MR. SILBERG: Are the systems really that discreet that you could turn over the LSS part of the RMS, or are they going to be so wrapped up with each other that --

MR. GANDI: That all depends on the February decision whether to take our architectural design and move it over to the NRC as the LSS. We are hoping that that's the case rather than go into a long procurement process to do a turnkey operation, turn over the NRC and then hope for compatibility between the two systems.

MR. SILBERG: No, no. But given the current system design, is it feasible to split out the operation of

the LSS portion of it while the DOE RMS remains on DOE's side.

MR. GANDI: Well --

MR. HALLNOR: As --

MR. GANDI: Go ahead.

MR. HALLNOR: Yeah. As the thing was presented this morning, we stated that DOE would maintain the records responsibility for any documents that's in the RMS, we would transfer a copy to a separate and independent system that is the LSS and, yes, indeed, it will be totally independent.

MR. DAVENPORT: Yeah, that's what I thought we had talked about to start with.

MR. GANDI: It's required by the rule and also we don't particularly want a lot of access under our corporate systems, security reason purposes. We want to be able to complete that link to the LSS system and have that basically fire walled out of our systems.

MR. DAVENPORT: Is there a way that you can guarantee an advance adequate funding then?

MR. CAMERON: Well, I think that that's, you know, I mean, guaranteeing adequate funding is difficult on anything. I mean, ultimately guaranteeing it is tough, but I think that the MOU has to provide some pretty clear directions on that.

MR. DAVENPORT: Do you suggest as a remedy in your

MOU that if the funding isn't there, that the application doesn't go forward?

MR. MURPHY: Yeah, sure. I mean, they've got the ultimate way to guarantee it.

MR. CAMERON: Because they're always full of good suggestions.

Okay. I mean, I heard you. But that's, you know, I mean, that's something that has to be --

MS. NEWBURY: We have to work out details assuming that we're agreeable to doing all this --

MR. CAMERON: Yeah. And, you know, I mean --

MS. NEWBURY: -- taken forward yet.

MR. CAMERON: Keep in mind also that, as I mentioned this morning, we did go back to the Commission at one point awhile back. It was another Commission than the Commission we have now, but they didn't want the fact that the LSS wasn't available, either because it hadn't been basically designed and developed or because the documents weren't there, they didn't want to say that you can't file the license application. Okay. So they didn't want to change the rule.

Now, as I also pointed out, this other Commission paper we did left open that option. I don't know whether the Commission would want to do through the back door the MOU what it didn't want to do directly. But I think that

what we need to do when we present the Commission with this particular problem is to note what the comments were, at least the comment of one participant, or two, whatever, from the advisory review panel. I mean, I can't say it's an advisory review panel position at this point, but we can note that that comment was made.

MR. METTAM: Certainly a dysfunctional LSS due to funding constraints is no better than none.

MR. CAMERON: Oh, absolutely. I think that's what -- I mean, I think that that's what you were suggesting is that if we don't get the --

MR. DAVENPORT: No. I'm suggesting that DOE could save it's money and hide the ball.

MS. NEWBURY: Well, money is money whether we're appropriating it for ourselves or appropriating it and handing it over to the NRC.

MR. DAVENPORT: Right.

MS. NEWBURY: And one of my problems with the CO business was that I still couldn't guarantee that we were going to get out of the funding. It's one of those issues that is going to come up every year and we will fight real hard to make sure we have adequate funding, but nobody is going to make any guarantees unless you have a line item meant from Congress that says we'll spend X number of dollars.

MR. DAVENPORT: Right. But you can't set up a system that has penalties if you don't do it.

MR. CAMERON: Yeah. I mean you can't -- if I mean if -- in theory you can say that if the money isn't there, one of the things that, I mean, the licensing process sort of grinds to a halt. Under your suggestion that's the way it would work.

MR. MURPHY: But on the basic message that you've just said, Claudia, if you can do that, assuming you can persuade the decision makers to sign off on that, that would be --

MS. NEWBURY: It would be a very --

MR. MURPHY: -- much -- it would pretty much solve our problem with management system.

MS. NEWBURY: I would hope so. It seems the cleanest way to do it.

MR. MURPHY: Sure.

MS. NEWBURY: All we do is establish a check in --

MR. MURPHY: It's the way we've always wanted it. It's the way we asked for it to be done in 1987 when we were negotiating the rules, so, yeah.

MS. NEWBURY: So where did we veer off away from that, because I don't have --

MR. MURPHY: You did. The NRC did.

MS. NEWBURY: Me?

MR. MURPHY: You guys have always been --

MR. CAMERON: Oh, I'm sorry. Not you.

MR. MURPHY: No, it's never been your fault. It's these guys at the NRC.

MR. CAMERON: He's pointing his finger.

MR. MURPHY: It's Chip.

MR. CAMERON: Yeah, it's my fault.

MS. NEWBURY: Okay.

MR. CAMERON: It isn't, but at least we can end this discussion.

MR. HOYLE: We thought that we could do the job, the job could get done without a transfer of dollars, that DOE could do the funding, could do the operating and we could oversee their operation of it and satisfy everyone that it was being done properly. And I think upon discussion and review, we find that you're not willing up front to say, yeah, that's the way to do it.

MR. DAVENPORT: Their lawyers.

MR. HOYLE: They have their own, too.

MR. DAVENPORT: Yeah, they have --

MR. CAMERON: Your lawyers are satisfied with this --

MS. NEWBURY: With the interagency.

MR. CAMERON: With the interagency agreement. It really comes down to policy budget, a budgetary issue now on

DOE's part.

MR. METTAM: This discussion is sort of waning down. A related question to Claudia. In '95 you show under the cost analysis now maybe six-odd million for the RMS, a million or so for the LSS --

MS. NEWBURY: I believe it's .8 million that's in the '95 budget.

ME. METTAM: I'm sorry?

MS. NEWBURY: .8 million is in the budget for '95. If you look at the long -- the five-year plan which is not that.

MR. METTAM: For the LSS?

MS. NEWBURY: For the LSS.

MR. METTAM: Okay. I'm just looking at bar charts, though. I mean, I could be wildly inaccurate.

And I assume then that you've got some 15 million in your records management '96 odd request and --

MR. GANDI: In our request, yes, we do.

MR. METTAM: -- three something, huh?

MR. GANDI: That's in the OMB request.

MR. METTAM: Well, I realize that you're not supposed to -- I'm just saying are these -- it's back to the budget issue. Is what you've been showing us as an option here supported by what you're requesting, in sort of general terms, not causing numbers --

MS. NEWBURY: Yes.

MR. METTAM: -- just sort of general terms.

MR. GANDI: As a request and it could change between now and July of next year.

MR. METTAM: Right. Okay.

MR. HOYLE: Any further discussion?

I see Dennis did have to leave again. So we will start tomorrow morning with discussion of the charter for the LSSARP technical working group which Dennis will lead for us. So we'll start at 8:30 tomorrow morning. Thank you very much.

(Whereupon, at 3:40 p.m., the hearing recessed to reconvene December 13th, 1994, at 8:30 a.m.)

CERTIFICATE OF REPORTER

This is to certify that the attached proceedings before: NUCLEAR REGULATORY COMMISSION

In the Matter of:

9th MEETING LSS ADVISORY REVIEW PANEL (ARP)

Case No. N/A

Location: Las Vegas, Nevada

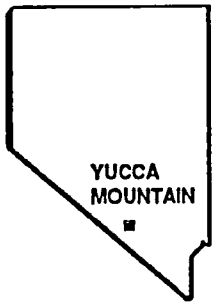
Date: December 12, 1994

were held as herein appears, and that this is the ORIGINAL transcript thereof for the files of the Department or Commission:

Dennis Davis
Reporter

U. S. DEPARTMENT OF ENERGY

OCRWM



YUCCA MOUNTAIN
SITE CHARACTERIZATION
PROJECT

UPDATE OF LSS WORKING GROUP ACTIVITIES

PRESENTED TO

LICENSING SUPPORT SYSTEM ADVISORY REVIEW PANEL

PRESENTED BY

CLAUDIA NEWBURY
ACTING TEAM LEADER, TECHICAL SYSTHESIS

DECEMBER 12, 1994

Introduction

- DOE presented an overview of the Working Group efforts at the September 19, 1994, LSSARP meeting
- DOE committed to providing a more detailed presentation of Working Group activities at the December 12-13, 1994, meeting
- The DOE has not yet endorsed the recommendations of the Working Group

Working Group - Brief History

- Responsibility for LSS was transferred to the YMSCO in fiscal year 1994
- YMSCO chartered a Working Group to examine the LSS and develop an implementation strategy that is consistent with the program approach
- Working Group submitted a draft report to the DOE in October, 1994
- The final report should be submitted to DOE by mid-January, 1995

Working Group Charter

- The Working Group was tasked with performing a comprehensive review of the LSS that included:
  - Development history
  - Identification of commitments and expectations
  - Statutory requirements
  - Evaluation of implementation options
  - Life-cycle costs
  - Expected data volume

Working Group Update - Agenda

- Overview of NRC Hearing Process - *Stan Echols*
- OCRWM Records Management System - *Marty Cummings*
- LSS Requirements - *Fielden Dickerson*
- LSS Functional Models and Cost Evaluation - *George Hallnor*

Working Group Update - Agenda (cont.)

- Schedules and Cost Profiles - *Fielden Dickerson*
- Systems-Level Requirements Document Review - *Camille Kerrigan*
- Data Volume Projections - *Jim Boone*
- Conclusions and Recommendations - *Mike Cline*

1

NRC HEARING PROCEDURES

FOR REPOSITORY LICENSING

Stan Echols
December 12,1994

INITIAL LICENSING PROCESS

* License Application (Part 60)
* Discovery/Motions (Part 2)
* Trial-Type Hearing (Part 2)
* Construction Authorization (Part 60)

LICENSE APPLICATION (LA)

* LSS Administrator Certification

* Docketing - NRC Review

  (note: §60.18, footnote 2, and §60.18(g)

* Contents of LA

  -- General information, SAR, EIS

* General Information

  -- schedules, safeguards certification, physical security plan

* Safety Analysis Report (SAR)

  -- technical information/criteria, performance assessment, models, QA program

LICENSE APPLICATION (LA)

- -- materials, design, alternate major design features, confirmatory design program
- -- program to resolve open safety questions

* Environmental Impact Statement (EIS)

- -- DOE site recommendation EIS adopted by the NRC
- -- NRC NEPA requirements for DOE EIS found in Part 51
- -- Portions of DOE EIS that are adopted by the NRC are excluded from the hearing (and discovery) process

* Filing of Documents Electronically (LSS)

DISCOVERY/MOTIONS

* Document Discovery (Part 2, Subpart J)
* Other Discovery (Part 2, Subpart G)
  - Interrogatories
  - Depositions
  - Requests to Admit
* Motions
  - Summary Jugment
  - Strike
  - Protective Order
* Electronic Filing of Documents (Part 2, Subpart J)

DEFINING THE ISSUES/PARTIES

* First Prehearing Conference
* Discovery Begins
  - Broad Scope
* NRC Issues SER
* Second Prehearing Conference
  - Finalize issues
  - Identify witnesses - testimony
  - Hearing schedule

DEFINING THE ISSUES/PARTIES

* **Motions**
  - **Summary judgment**
  - **Stipulations**
  - **Based on discovery record**
* **Reality Check**
  - **3-year schedule**

THE HEARING

* Adjudicatory or Trial-Type Hearing
* Safety (Part 60) and Environmental (Part 51)
* Consolidated Parties/Issues
* Testimony and Cross Examination
  - Individual experts
  - Use of panels
  - Expert elicitation
* NRC 90 Day Estimate for Hearing
* Findings of Fact and Conclusions of Law
  - Reasonable assurance finding
  - Part 60, Subpart E criteria

THE HEARING

- QA
- Emergency Plan
- Operating Procedures
- Safeguards (certification)
- Environmental

CONSTRUCTION AUTHORIZATION

* After Affirmative Consideration of Previously-Referenced Factors, the NRC Issues a Construction Authorization

* Conditions

  - Based on the record at the end of the hearing

  - Health/safety/environment

ATTACHMENT

SCHEDULE FOR NRC PROCEEDING ON APPLICATION FOR A LICENSE TO RECEIVE AND POSSESS HLW AT A GEOLOGIC REPOSITORY

TIME ELAPSED (in days)*****/	ACTION
Approximately 365 days	Parties and potential parties receive electronic access to the LSS. (§2.1003).
180 days	LSS Administrator certifies that DOE has complied with its obligations concerning the LSS (§2.1003(h)(1)).
0	DOE submits to the NRC its license application, Safety Analysis Report and EIS. The NRC determines that the application is complete and acceptable, NRC publishes a Federal Register notice announcing the docketing of the application and the availability of a public hearing on the application. Members of the presiding NRC Licensing Board are named. (§2.101(f)(8); (§2.105(a)(5)).
30	Petitions to intervene/requests for hearing, along with proposed contentions, are filed with NRC. (§2.1014(a)(1)). Petitions for status as interested government participant & interested government participant petitions are filed. (§2.715(c)).
50	Answers to intervention & interested government participant petition are filed. (§2.1014(b)).
70	The Board conducts 1st Prehearing Conference. (§2.1021).
100	The Board issues 1st Prehearing Conference Order which identifies participants in hearing, admits contentions, and sets discovery and other schedules. (§2.1018(b)(1)). Deposition discovery begins. (§2.1019).
110	Parties may file appeals from 1st Prehearing Conference Order, w/briefs. (§2.1015(b)).

*****/ Day Zero (0) represents the publication of the Federal Register notice announcing the NRC's docketing of the license application.

120	Parties may file briefs in opposition to appeals. (§2.1015(b)).
150	The Commission issues order ruling on appeals from 1st Prehearing Conference Order.
548	The NRC Staff issues **SER** on license application.
578	The Board holds 2nd Prehearing Conference. (§2.1022).
608	The Board issues 2nd Prehearing Conference Order finalizing contentions for hearing and setting schedule for pre-filed testimony and hearing.
618	Parties may file appeals from 2nd Prehearing Conference Order, w/briefs. (§2.1015(b)).
628	Parties may file briefs in opposition to appeals. (§2.1015(b)).
658	The Commission issues order ruling on appeals from 2nd Prehearing Conference Order.
660	Last practicable date for filing motions for summary disposition.
680	Replies to last practicable motions for summary disposition.
690	**Discovery complete.** (Supp. Info)
700	**Board issues order on last practicable motions for summary disposition.**
710	Parties may file appeals from last practicable summary disposition order w/briefs. (§2.1015(b)).
720	**Evidentiary hearing begins.** Parties may file briefs in opposition to appeals from last practicable summary disposition orders. (§2.1015(b)).
810	**Evidentiary hearing ends.**
840	Applicant files proposed findings. (§2.754(a)(1)).
850	Intervenor files proposed findings. (§2.754(a)(2)).
860	NRC Staff files proposed findings. (§2.754(a)(2)).

865	Applicant files reply to proposed findings. (§2.754(a)(3)).
955	**Board issues initial decision.** (§2.760).
965	Parties may file any stay motions responding to Commission notices of appeal. (§2.788(a), §2.762(a), §2.1015(c)).
975	Parties may file replies to stay motions. (§2.788(d)).
995	Commission issues ruling on stay motion. Appellant files brief before Commission. (§2.786).
1005	Stay motions to Commission. (§2.788(a)).
1015	Replies to stay motions. (§2.788(d)).
1025	Appellee files brief before Commission. (§2.786).
1035	NRC Staff brief before Commission.
1055	Completion of NMSS and Commission supervisory review; Commission ruling on any stay motions; issuance of construction authorization; NWPA 3-year period tolled. (§2.1023 Supp. Info).
1065	Commission hears oral argument on appeal are heard by Commission. (§2.763).
1125	**Commission issues decision on construction authorization.** (§2.1023).

Civilian Radioactive Waste Management System

Management & Operating Contractor

TRW Environmental Safety Systems, Inc.

OCRWM Records Management Overview

Presented to the LSSARP

Martin Cummings

December 12, 1994

B&W Fuel Company
Duke Engineering & Services, Inc.
Fluor Daniel, Inc.
INTERA, Inc.
JK Research Associates, Inc.
E. R. Johnson Associates, Inc.
Logicon RDA
Morrison Knudsen Corporation
SAIC, Inc.
TRW Environmental Safety Systems, Inc.
Winston & Strawn
Woodward-Clyde Federal Services

Justification for Existing RMS

- Management of QA Records
- Support of various litigation efforts
- Prudent business practice

Relationship Between OCRWM RMS & the LSS

- **Records Management System (RMS) Summary**
  - Source of all DOE data submitted to LSS
  - 85% of LSS data will come from RMS
  - LSS & RMS capture requirements must be consistent
- **Critical Overlaps Between RMS & LSS**
  - Header fields
  - Data formats
  - Inclusion Requirements

Description of Current RMS

- System Requirements
  - Capture & Manage all Program Records
  - Comply with all OCRWM Records Policies
  - Support all OCRWM Records sites
- System Architecture
  - VAX/Ingres indexing & retrieval system
  - Records Information System (RIS) database
  - Micrographics technology for image capture, distribution & storage

Description of Current RMS (cont'd)

- **Operational Features**
  - Managed & staffed by M&O contractor
  - Utilizes record inclusion screening criteria based on:
    - » 10 CFR 2, Subpart J
    - » OCRWM QARD
    - » Site Characterization Plan
    - » Waste Acceptance Requirements
    - » Storage & Transportation
  - Excluded records types include
    - » Administrative/Personnel
    - » Financial

Description of Future RMS Enhancements

- Record Information System (RIS)in conversion to Interim-RIS
  - Improves integrity & accuracy of data
  - Improves retrieval
  - Supports reprocessing of 500,000 records
- Micrographics to be replaced by Electronic Imaging
  - Micrographics equipment outdated
  - Electronic imaging simplifies storage & retrieval
  - Electronic imaging facilitates text conversion

Description of Future RMS Enhancements (cont'd)

- **Anticipated Benefits from Planned Enhancements**
  - Broader & more effective retrieval capabilities
  - More efficient indexing practices
  - Opportunities to reengineer records processing workflows
  - More efficient transfer of data from RMS to LSS

Civilian Radioactive Waste
Management System
Management & Operating
Contractor

TRW Environmental Safety
Systems Inc.

LSS Working Group

LSS REQUIREMENTS

DEFINED IN 10 CFR 2 SUBPART J

Fielden Dickerson
December 12, 1994

B&W Fuel Company
Duke Engineering & Services, Inc.
Fluor Daniel, Inc.
INTERA Inc.

JK Research Associates, Inc.
E. R. Johnson Associates, Inc.
Logicon RDA

Morrison Knudsen Corporation
TRW Environmental Safety Systems Inc.
Winston & Strawn
Woodward-Clyde Federal Services

Section 2.1002, HIGH-LEVEL WASTE LICENSING SUPPORT

- The LSS is an electronic information management system containing the documentary material of the DOE and its contractors, and the documentary material of all other parties, interested governmental participants and potential parties and their contractors.

- Access to the LSS by the parties, interested governmental participants, and potential parties provide the document discovery in the proceeding.

- The LSS provides for the electronic transmission of filing by the parties during the high-level waste proceeding, and orders and decisions of the Commission and Commission adjudicatory boards related to the proceeding.

Section 2.1003, SUBMISSION OF MATERIAL TO THE LSS

- Submission of material to the LSS shall be accomplished by submitting an ASCII file, an image, and bibliographic header for all material to be included in the LSS.

ASCII File means a computerized text file conforming to the American Standard Code for Information Interchange which represent characters and symbols.

Image means a visual likeness of a document, presented on a paper copy, microform, or a bit-map on optical or magnetic media.

Bibliographic Header means the minimum series of descriptive fields that a potential party, interested governmental participant, or party must submit with a document or other material.

Section 2.1007, ACCESS

Public Access

During the pre-license application phase terminals for access to full headers and images shall be provided at DOE Headquarters, NRC Headquarters, and at all NRC or DOE public reading rooms in the vicinity of the candidate site for a geologic repository. Additionally, terminals will be provided at the Uranium Recovery Field Office in Denver, Colorado, and at Las Vegas, Nevada; Reno, Nevada; Carson City, Nevada; and Lincoln County, Nevada.

After the license application is docketed, the public access is to include searchable full text at the identified sites.

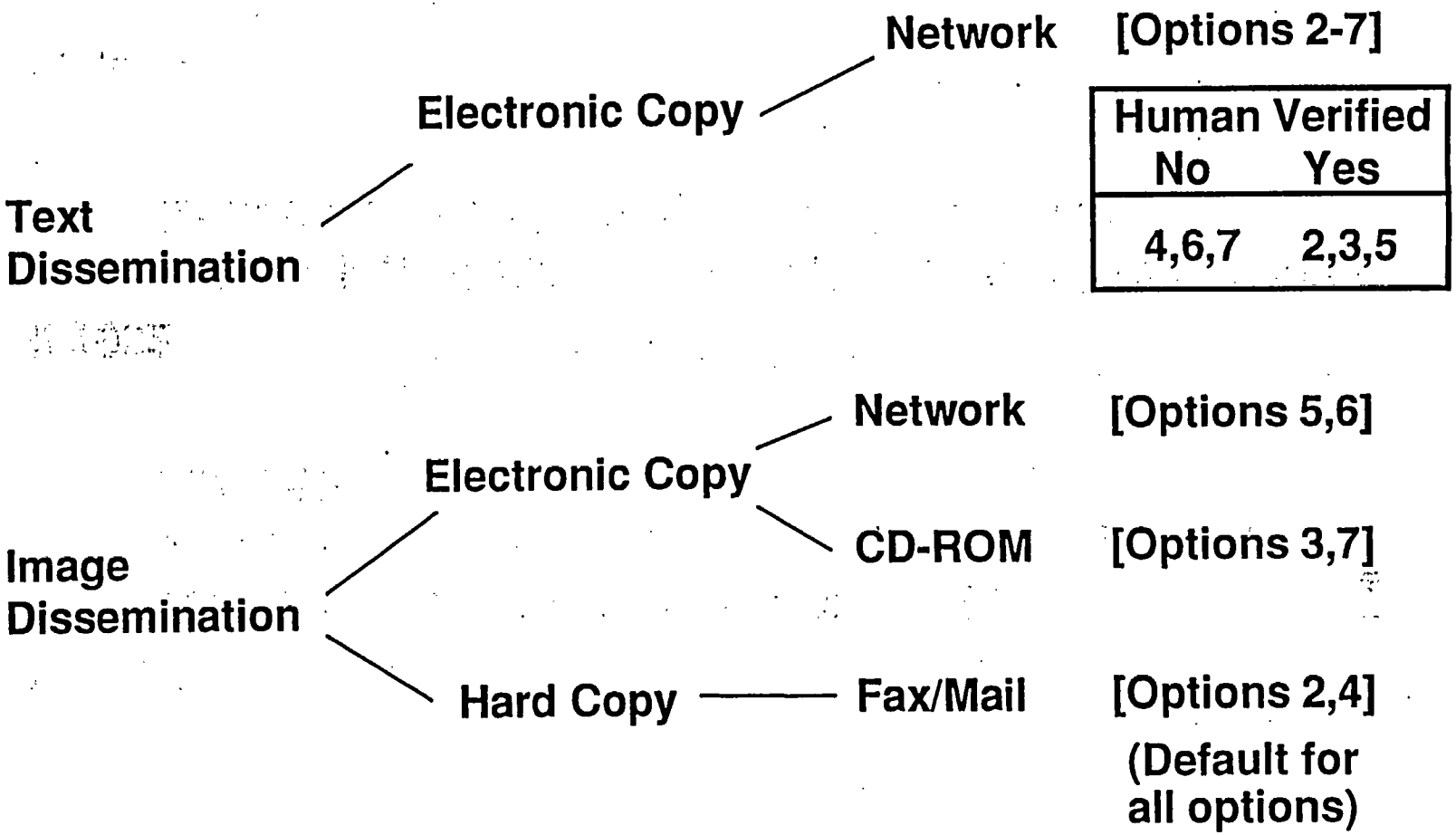
Access for Potential Parties, Interested Governmental Participants, and Parties

Access will be provided in the following manner:

- Full text search capability dial up access from remote locations.
- Image access at remote locations
- Capability to electronically request a paper copy of a document at the time of search.

These requirements were the basis for the examination of possible LSS options and the identification of costs associated with each option.

LSS Options Considered



COSTING OF LSS

I 1989

- 1989 and earlier computer technology
- No separation of costs between Records Management and LSS

II 1992

- Calculation directed at specific issue of using Record Management technologies to support the LSS.

III 1994

- Cost sensitivity

Civilian Radioactive Waste
Management System

Management & Operating
Contractor

TRW

TRW Environmental Safety
Systems Inc.

LSS Working Group

Cost Analysis

George Hallnor

December 12, 1994

B&W Fuel Company
Duke Engineering & Services, Inc.
Fluor Daniel, Inc.
INTERA Inc.

JK Research Associates, Inc.
E. R. Johnson Associates, Inc.
Logicon RDA

Morrison Knudsen Corporation
TRW Environmental Safety Systems Inc.
Winston & Strawn
Woodward-Clyde Federal Services

Cost Model Assumptions

- OCRWM Records System (RMS) will capture the majority of the LSS records
  - Model assumes that 85% of the LSS holdings come from the RMS
  - The remaining 15% is submitted directly to the LSS (NRC)

- The LSS is a separate system, operated independently of the RMS
  - Copies of records are transferred from the RMS to the LSS, the official records are kept by DOE
  - All dissemination of documents and information is done from the LSS
  - DOE records not included in the LSS are distributed by DOE as required.

Facilities

- The cost model assumes DOE and NRC operates server and capture facilities

- The location of the facilities are

  - Server Facilities:
    - DOE: Dunn Loring (VA) Las Vegas (NV)
    - LSS: Las Vegas (NV)

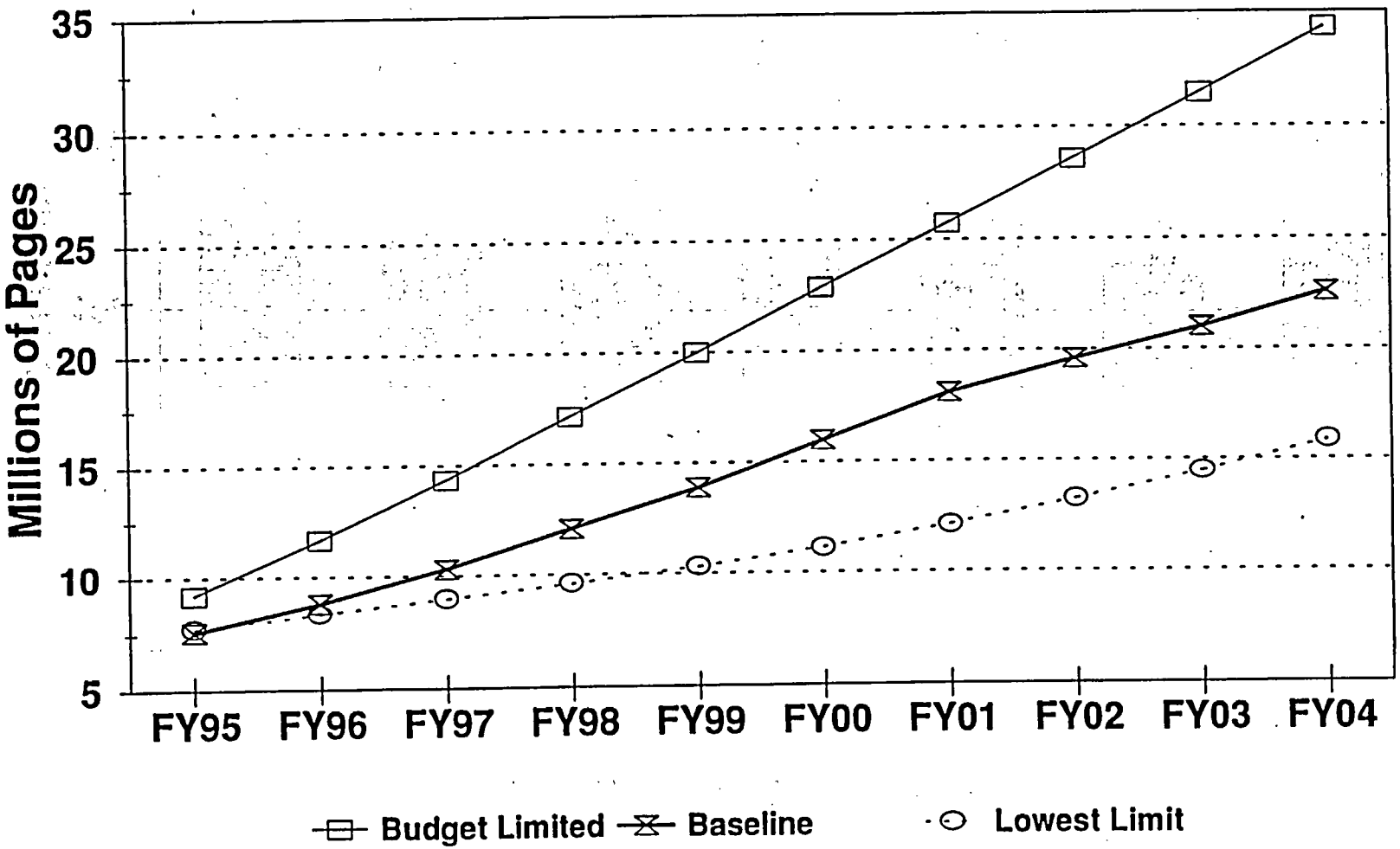
  - Capture Facilities:
    - DOE: DOE Headquarters (D.C.) Dunn Loring (VA)
      Las Vegas (NV)
    - LSS: Las Vegas (NV)

- A total of 9 public LSS access facilities will be located in Washington D.C. and the State of Nevada (Las Vegas)

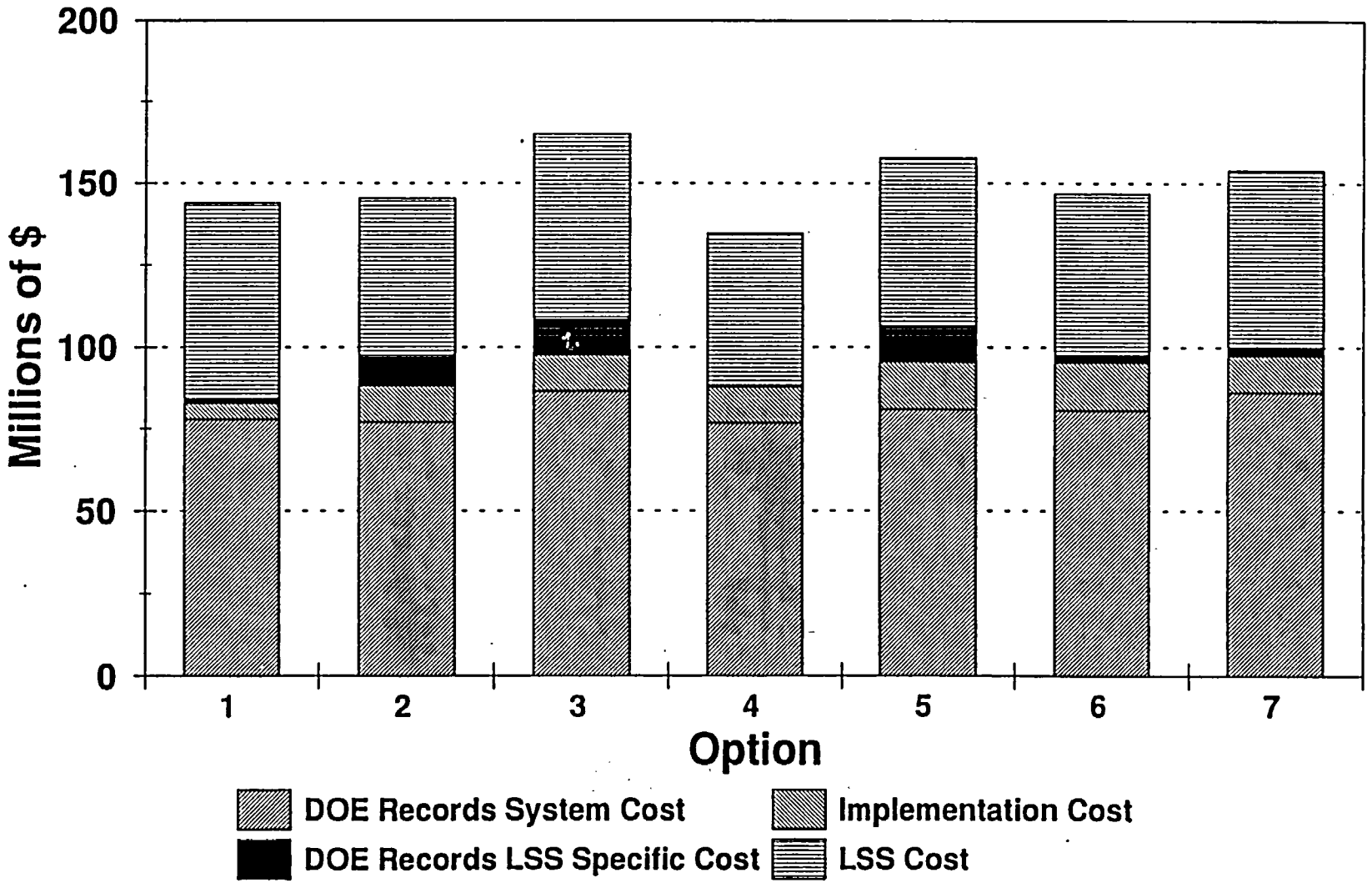
Key DOE Cost Drivers

- Data Volume, i.e. the number of pages that has to be processed through the RMS
  - Sensitivity analysis done using three intake volumes
- Human corrected text, and OCR accuracy
  - Options defined with and without text correction
  - Analysis of sensitivity to OCR accuracy performed
- Dissemination cost (an LSS cost) for varying page volumes
  - Sensitivity analysis done on electronic dissemination of a mix consisting of 80% text 20% image for multiple dissemination volumes

LSS Relevant Pages



Total Option Cost



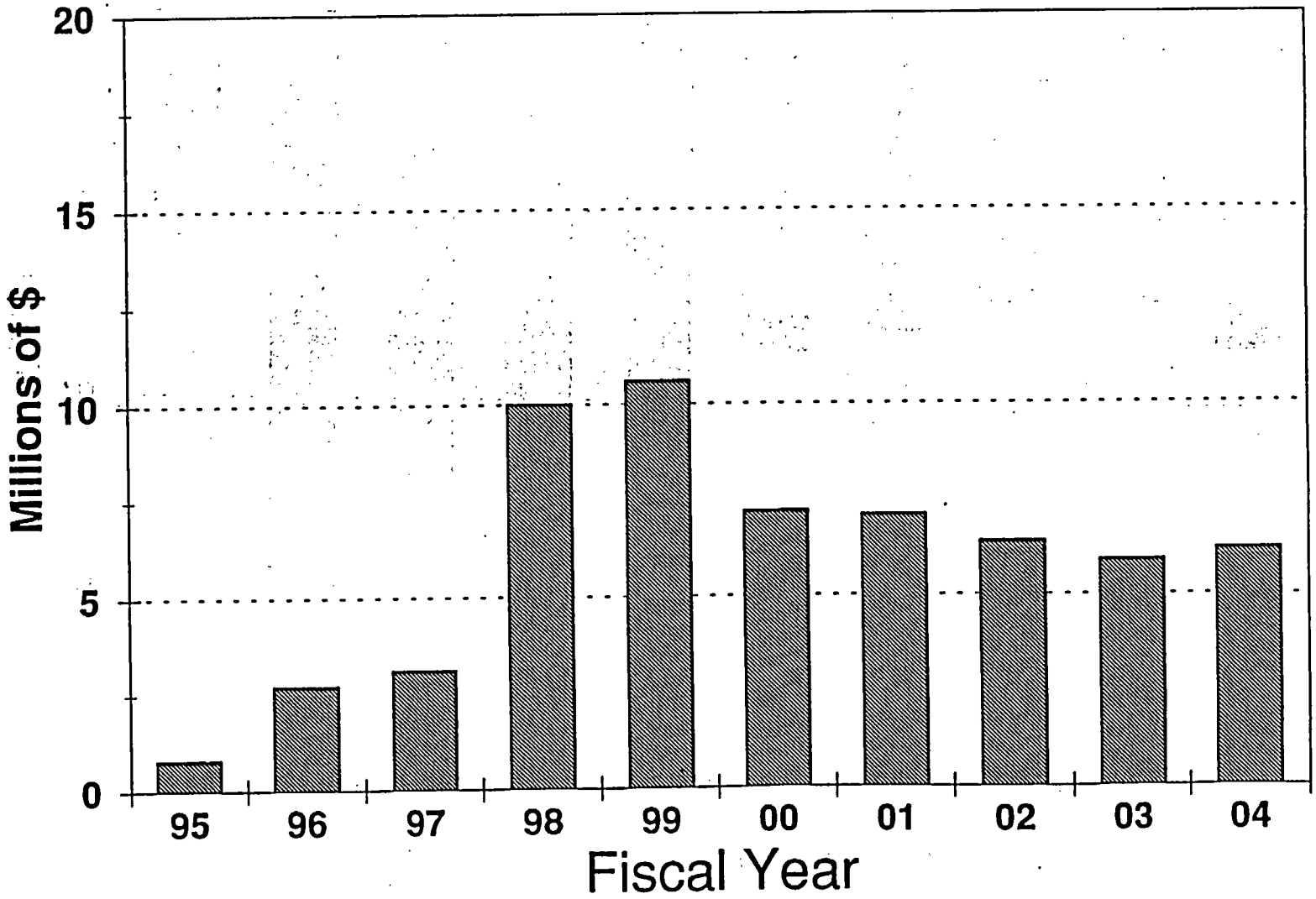
Summary

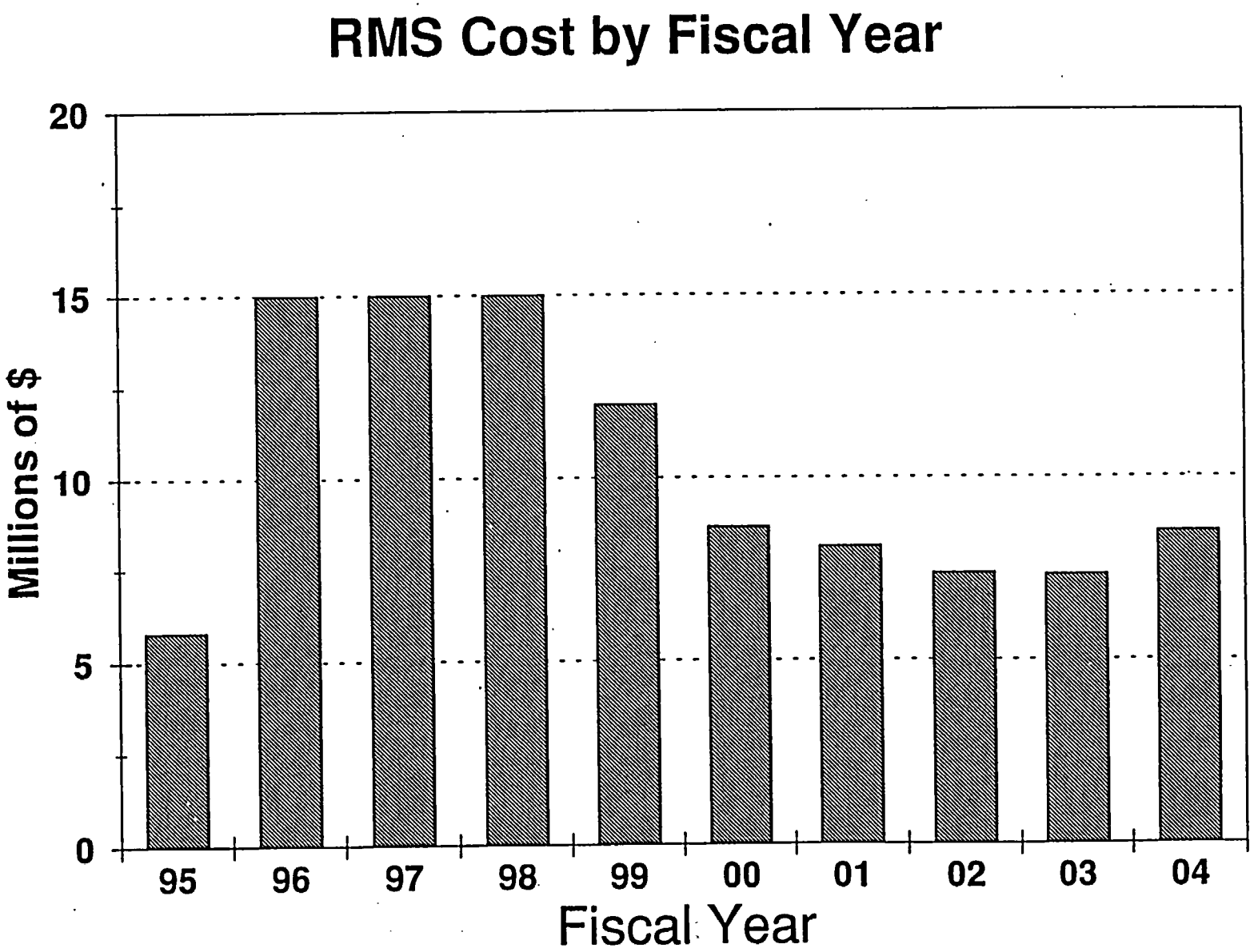
- The LSS Related Cost for the options are:

Option 2:	$57M
Option 3:	$64M
Option 4:	$54M
Option 5:	$62M
Option 6:	$59M
Option 7:	$61M

Supporting Information

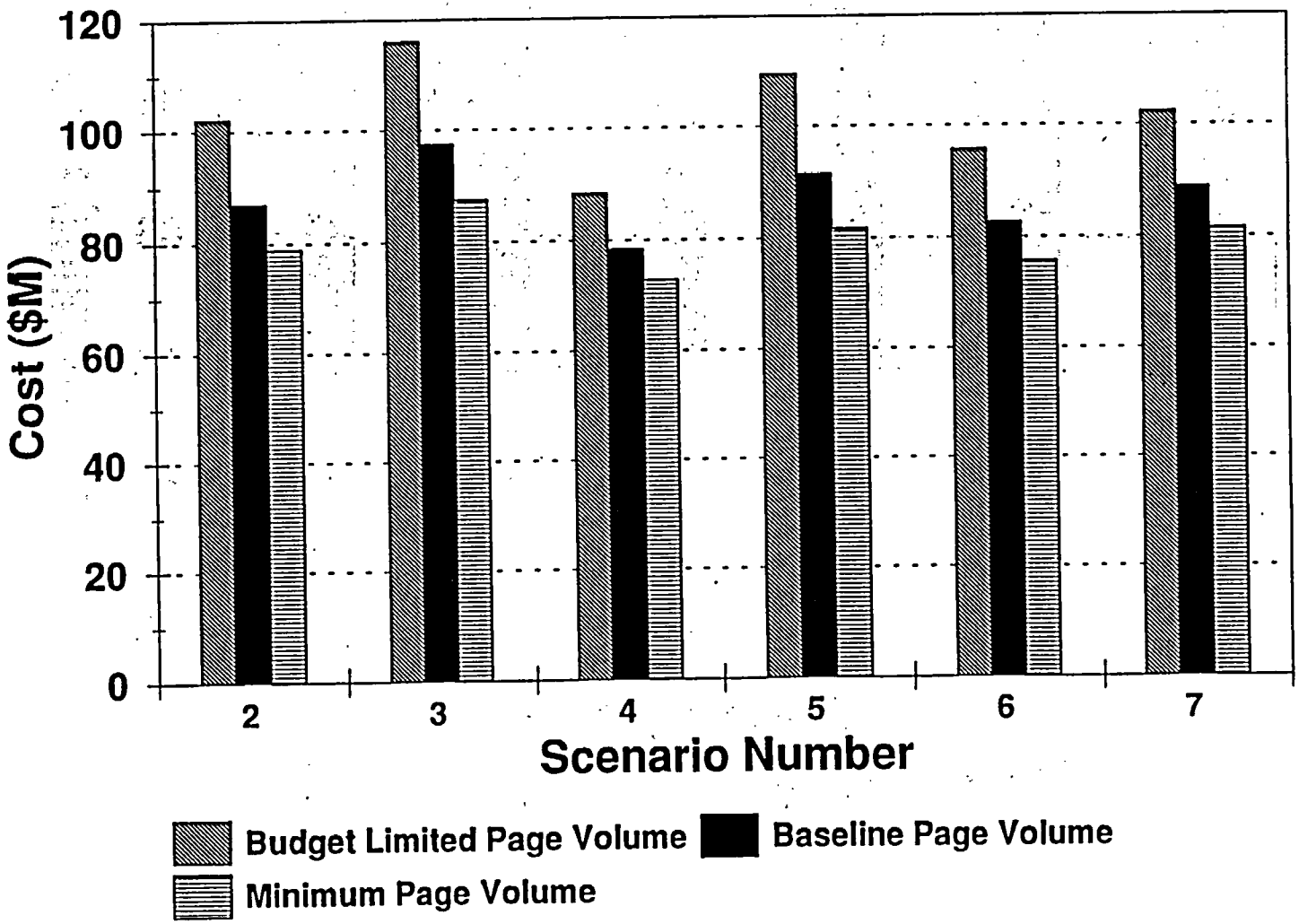
LSS Cost by Fiscal Year




RMS Cost by Fiscal Year
Millions of $
20
15
10
5
0
95
96
97
98
99
00
01
02
03
04
Fiscal Year

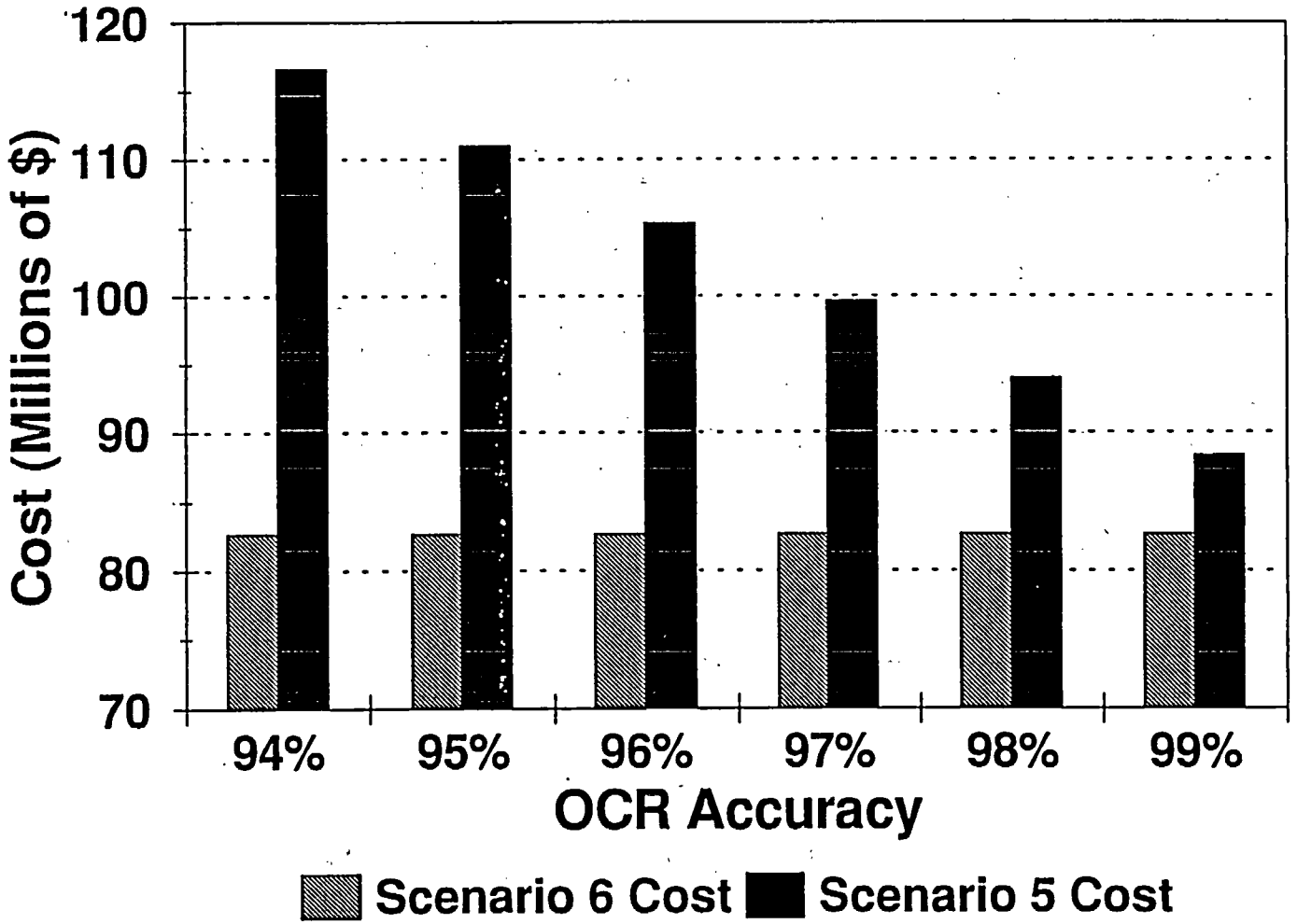
Sensitivity to Page Volume

RMS Cost Only



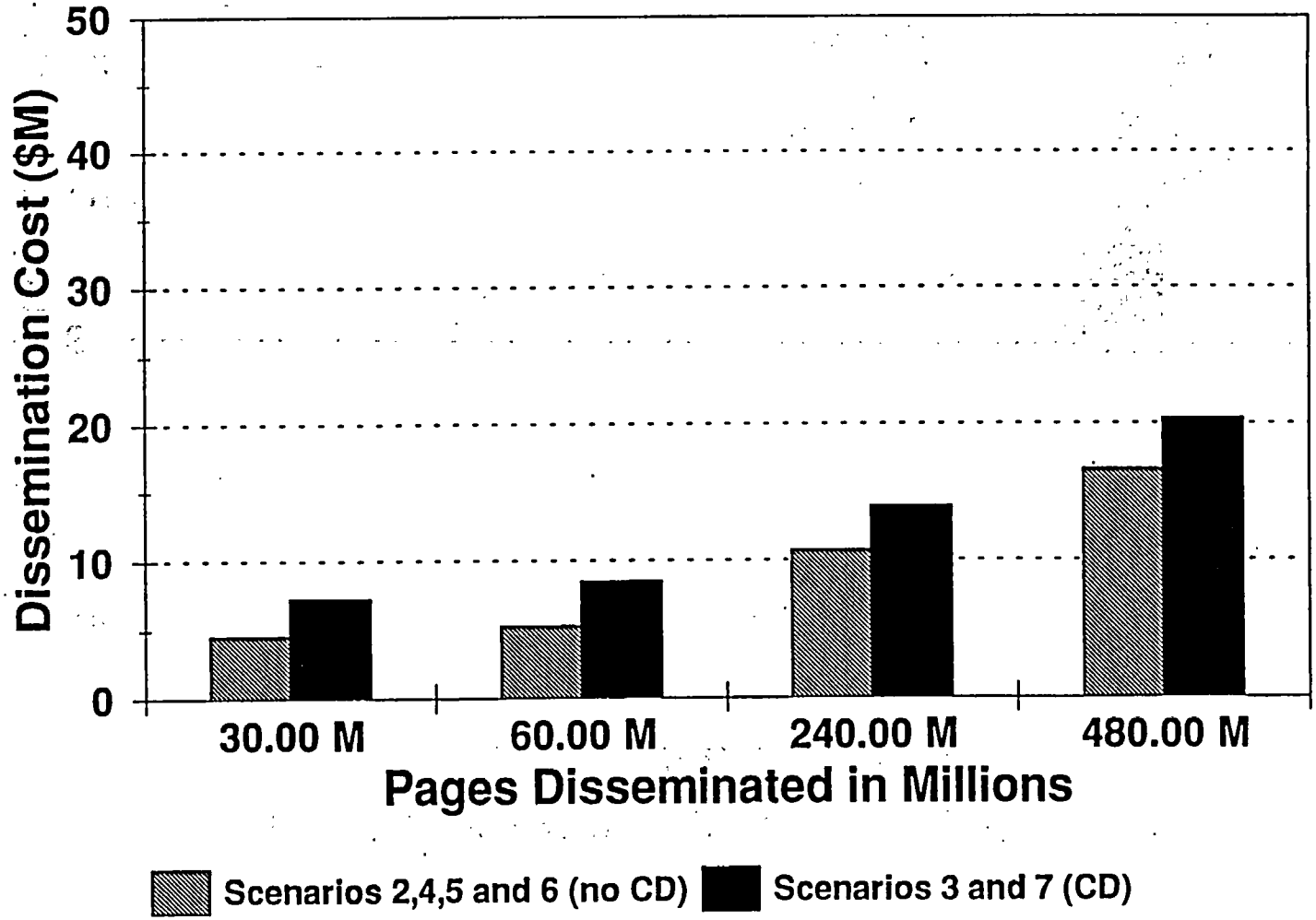
Sensitivity to OCR Accuracy

RMS Cost Only



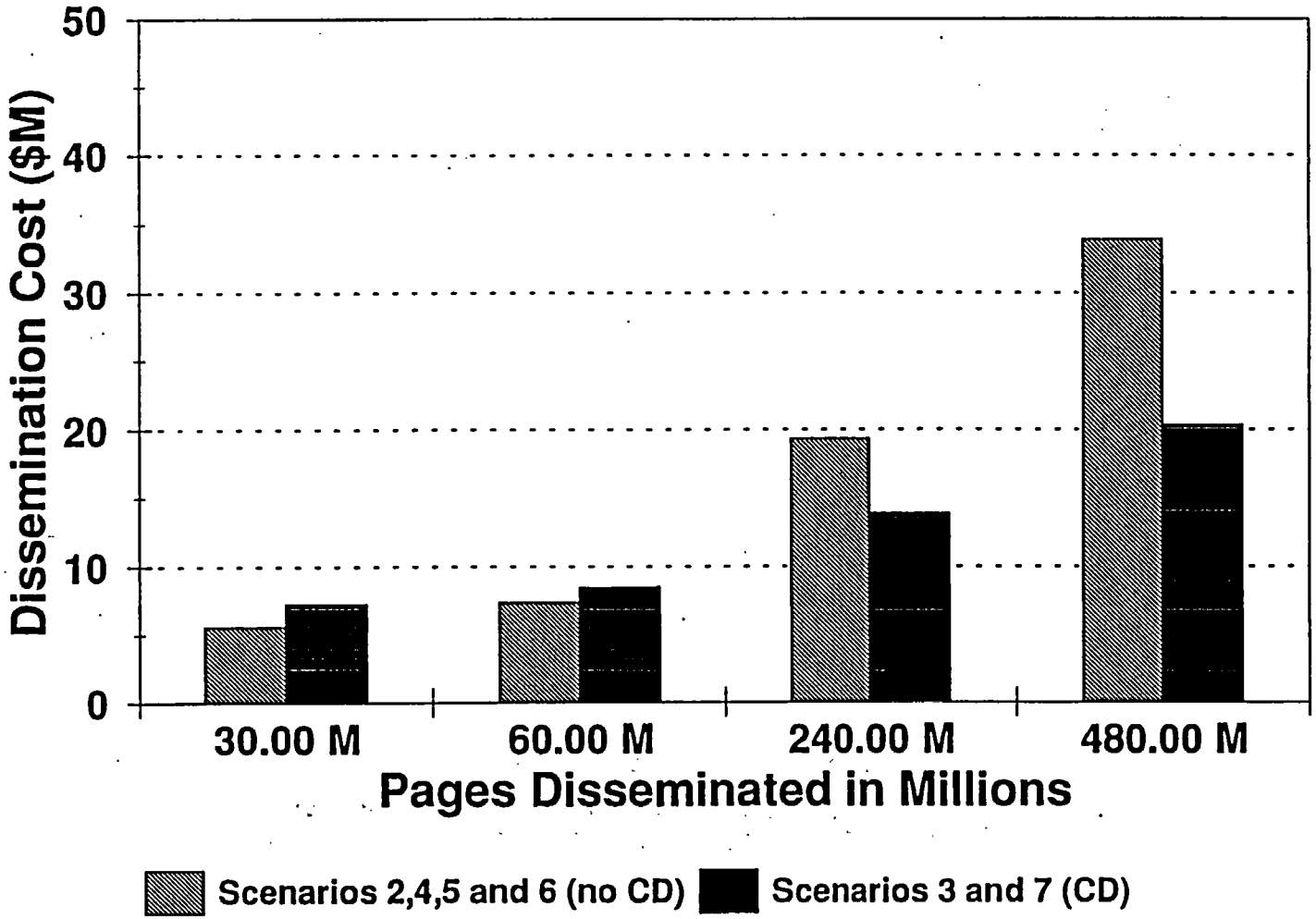
LSS Dissemination Sensitivity

Dissemination 20% Image



LSS Dissemination Sensitivity

Dissemination 80% Image



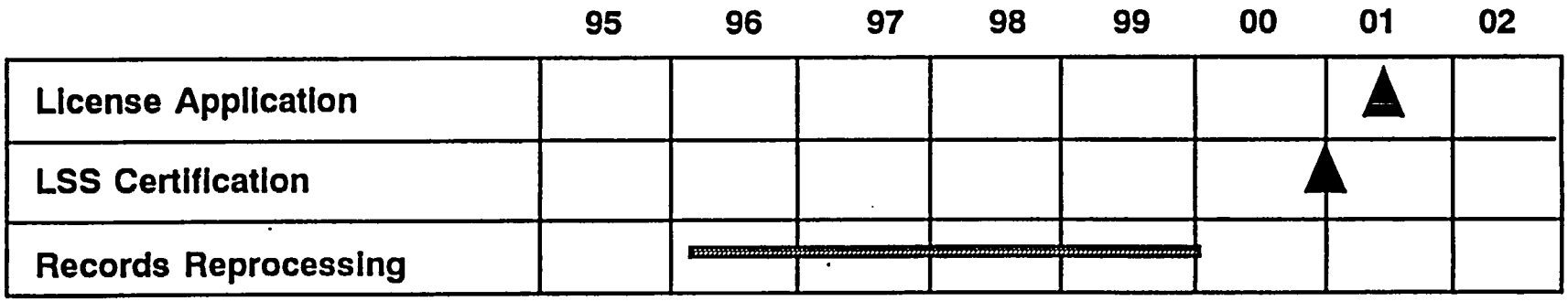
LSS ARP MEETING

Projected LSS Schedule

Fielden Dickerson
December 12, 1994

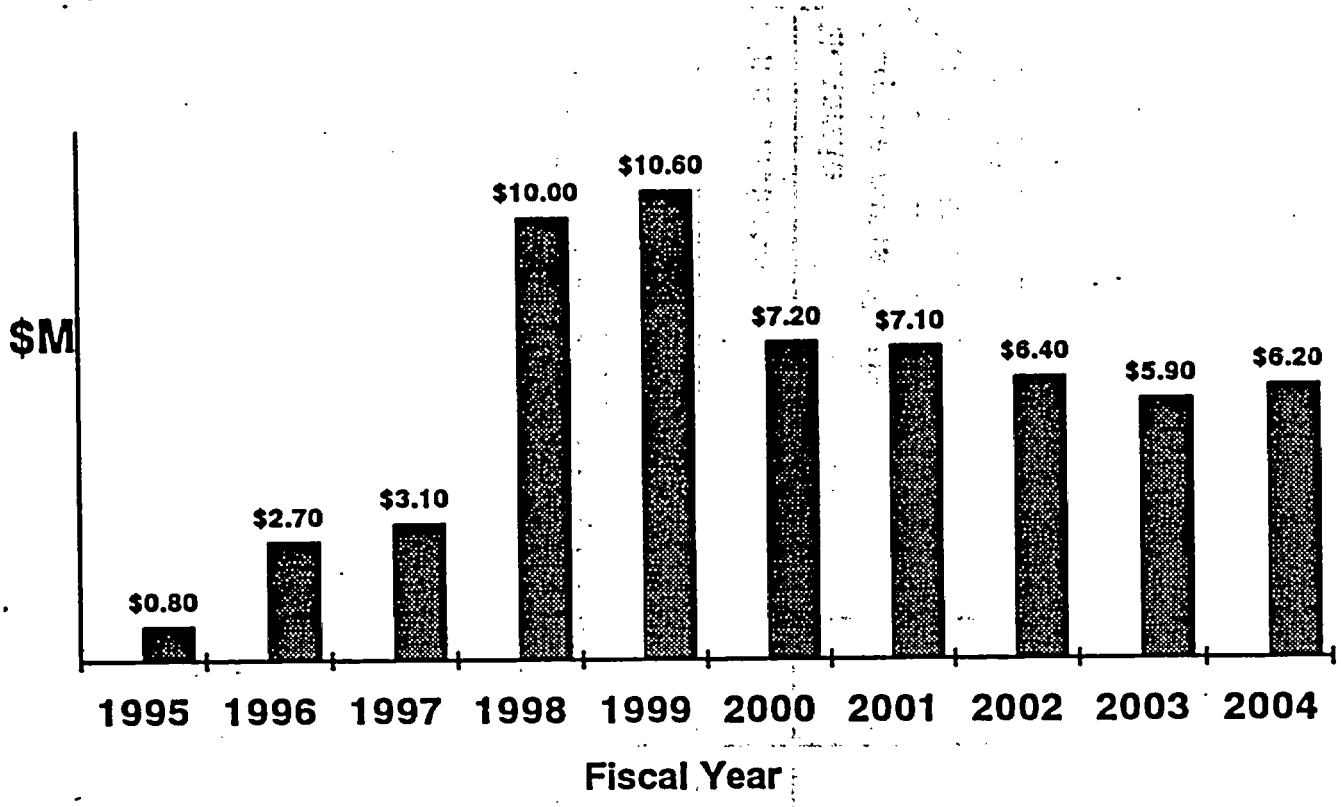
Time Constraints for the LSS

Fiscal Year

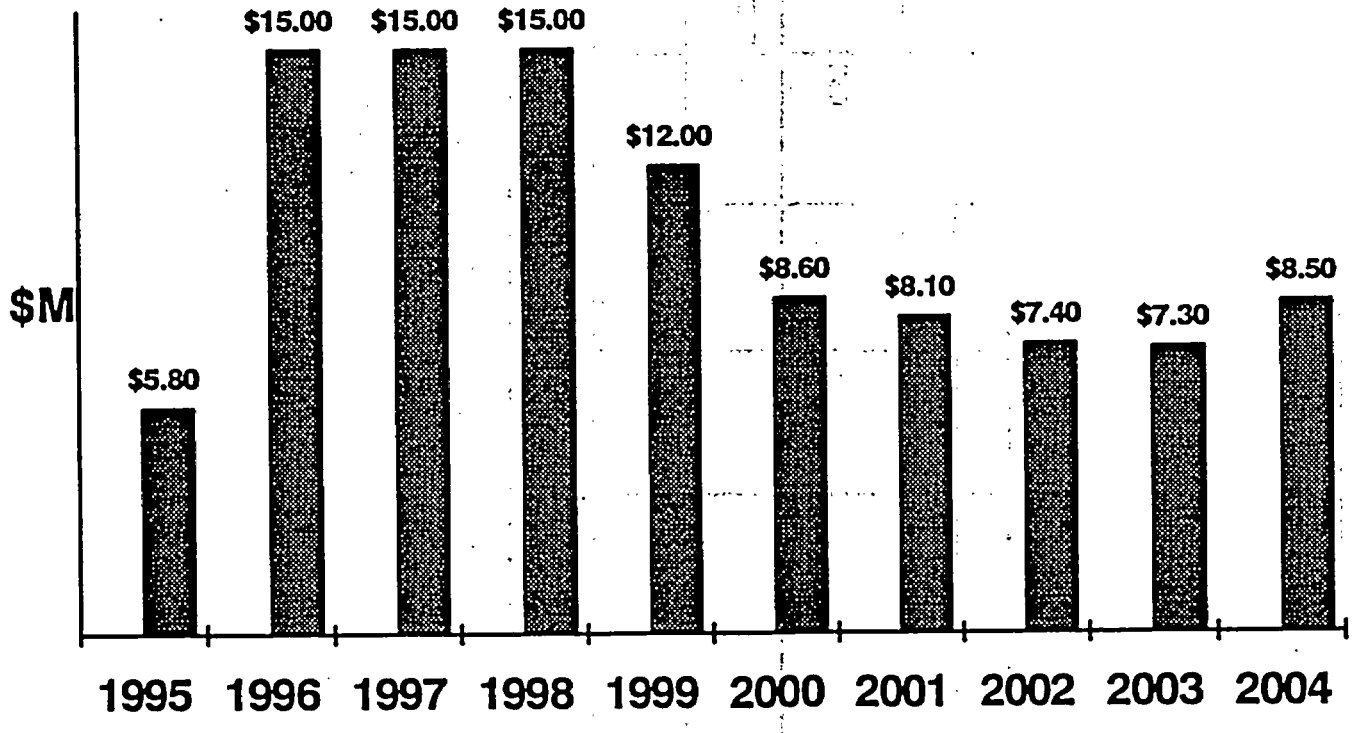


------ Milestones

------- Funding Profile (required---feasible)

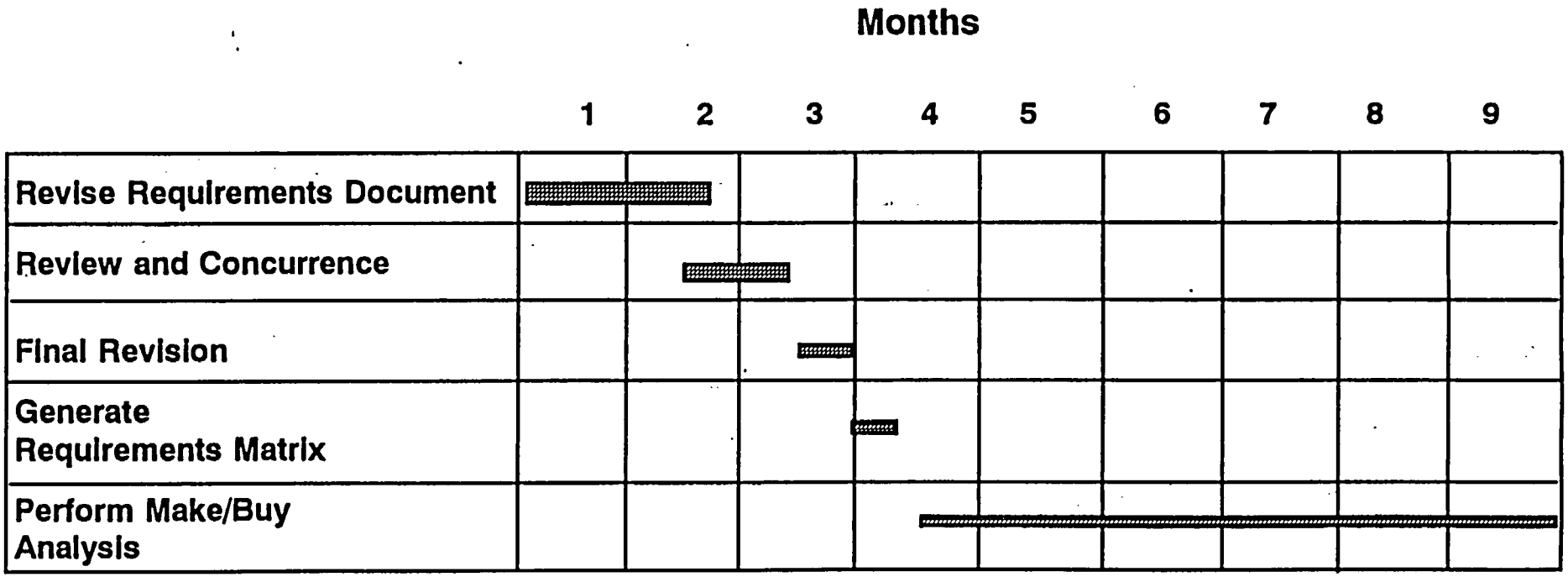


Annual LSS Cost

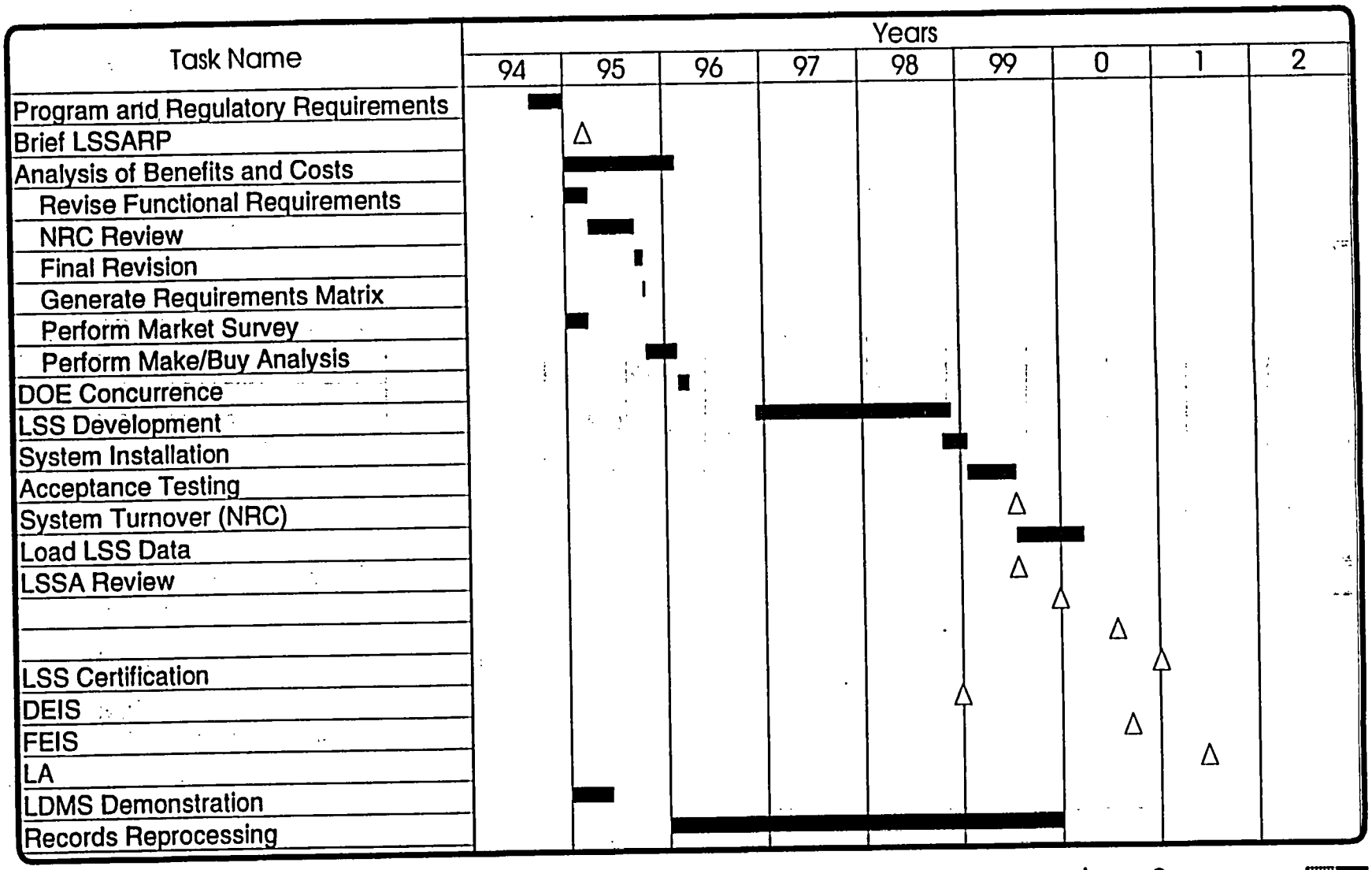


Annual Record Management System Cost

Make/Buy Analysis

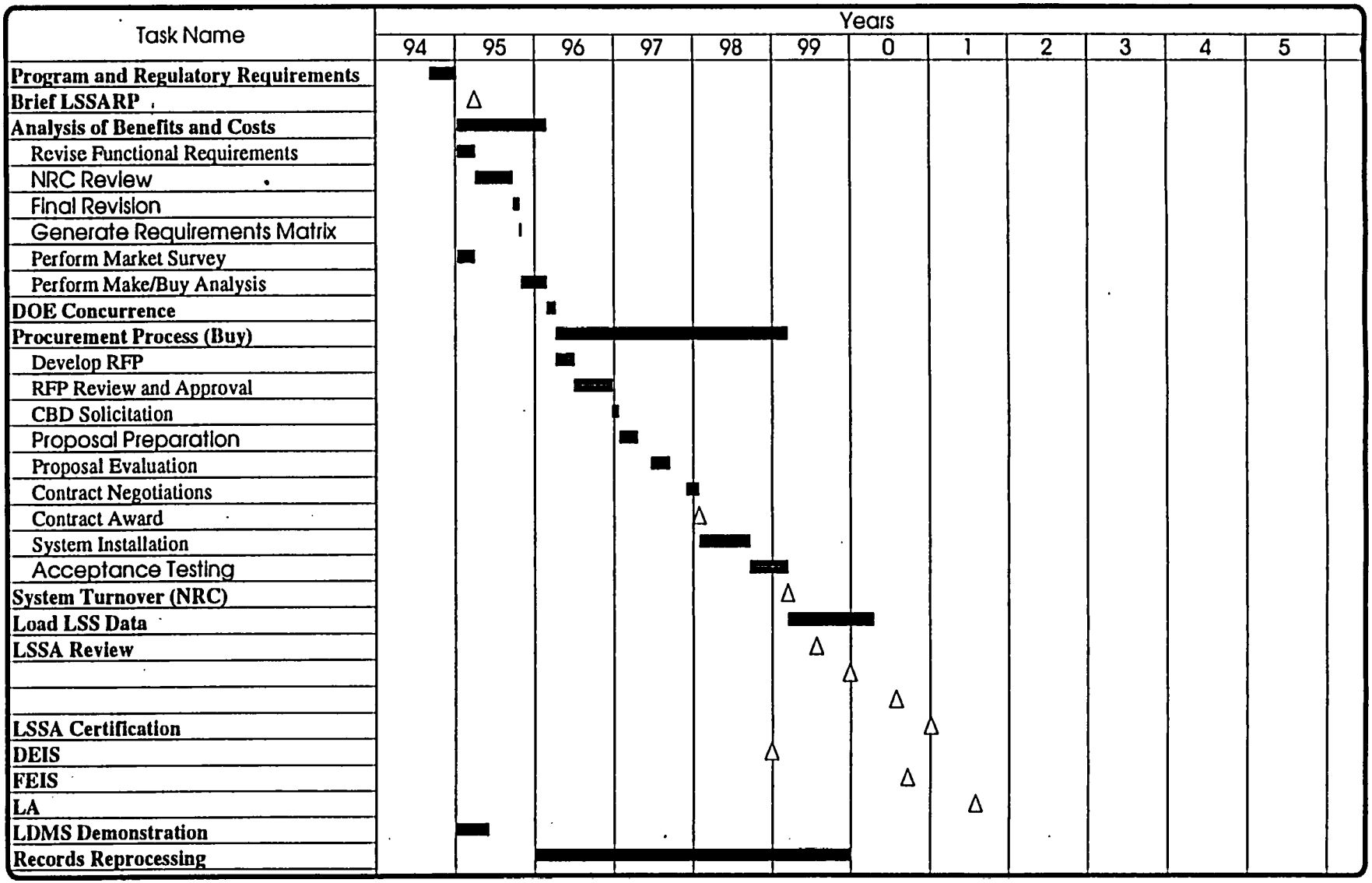


LSS Build Schedule



Printed: Dec/07/94

LSS Buy Schedule



Printed: Dec/07/94
Page 1

Milestone Δ Summary ▨
Fixed Delay - - -

REVIEW OF THE LICENSING SUPPORT SYSTEM SYSTEM-LEVEL REQUIREMENTS DOCUMENT

Camille Kerrigan
December 12,1994

BACKGROUND

- Document was developed in 1988 by SAIC and last revised in 1990
- No formal documentation of LSSARP approval of the document has been identified
- Document was prepared subsequent to a prototype proof-of-concept performed by SAIC
- Document is predicated on a pre-conceived system design

GENERAL COMMENTS

- Document specifies design and implementation (not just requirements)
- Document is a mixture of system and procedural requirements
- Difficult to determine hard requirements due to the mixed use of "shalls" and "shoulds" throughout the document
- Document contains extensive specification of non-quantifiable (non-measurable) requirements which are untestable
- Requirements are restated for different system functions and are often conflicting

EXAMPLES OF REQUIREMENTS WHICH SPECIFY DESIGN

- "This section lists the system-level functional requirements that have been identified for each LSS component
  - Capture System
  - Image System
  - Search System
  - Communication System"
- Comment: The document discusses interfaces between these "pre-conceived" subsystems

EXAMPLE OF CONFLICTING REQUIREMENTS

COMMUNICATION SYSTEM

- **"Hard copy of large outputs (greater than 100 pages) should be available overnight"**

IMAGE SYSTEM

- **"Large volumes of hard copy (thousands of pages) should be available, via remote print and distribution, within 5 working days after request"**

EXAMPLES OF AMBIGUOUS AND CONFLICTING REQUIREMENTS

- "Average query response time shall be less than 10 seconds"
- "Terminal response time 'to begin to respond' shall be no more than a few seconds"
- "Search time to identify all documents ... [search phrases]... shall be completed within a few minutes"

EXAMPLES OF UNTESTABLE REQUIREMENTS

- "The user interface must be interactive and intuitive"
- "The user interface shall be consistent"
- "The user interface shall be unambiguous"
- "The user interface shall be flexible and convenient"

EXAMPLES OF PROCEDURAL REQUIREMENTS

- **"All documentary materials shall be submitted to the LSS Capture System in accordance with procedures established by the LSS Administrator"**
- **"For non-English documents, headers, including optional fields shall be completed in English by the submitter in accordance with procedures established by the LSS Administrator"**
- **"There should be a procedure to identify and minimize or avoid duplicate records"**

CONCLUSIONS

- Document cannot be used to do a "make/buy" analysis due to:
  - » ambiguity of requirements
  - » conflicting requirements
  - » untestable requirements
  - » specification of system design
- Document cannot be used as a specification for an RFP
- Document needs to be revised to correct deficiencies, remove ambiguities and conflicts and to remove references to design

OBSERVATIONS

- **Revision of the document is on the critical path for the implementation of the LSS**

- **Recent discussion with the NRC indicates they are amenable to document revision**

Civilian Radioactive Waste
Management System
Management & Operating
Contractor

TRW Environmental Safety
Systems Inc.

LSS Working Group

Data Volume Projections

Jim Boone
December 12, 1994

B&W Fuel Company
Duke Engineering & Services, Inc.
Fluor Daniel, Inc.
INTERA Inc.

JK Research Associates, Inc.
E. R. Johnson Associates, Inc.
Logicon RDA

Morrison Knudsen Corporation
TRW Environmental Safety Systems Inc.
Winston & Strawn
Woodward-Clyde Federal Services

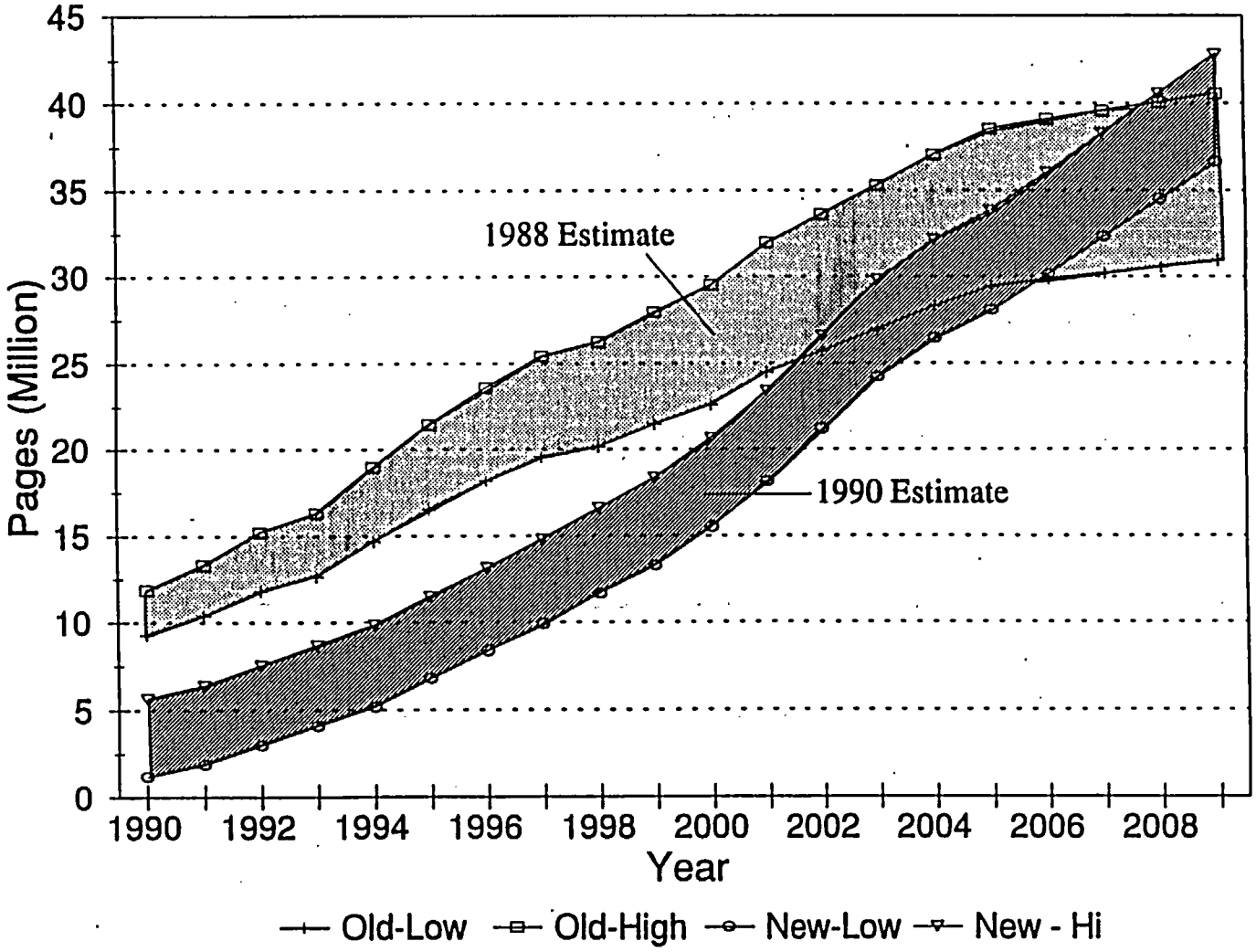
LSS Data Volume Projections

- An estimate of LSS data volume is needed for:
  - total life-cycle cost projections
  - system design considerations

SAIC Data Volume Projections

- Estimates were based upon:
  - historical records growth rates
  - nominal number of pages per document
  - percentage of documents judged to be licensing relevant
  - percentage of duplicate documents in the system
- Estimates were presented as a range (high and low)
- Low and high estimates were extrapolated through 2009 by assuming compounded annual growth rates of 10% and 20%, respectively

SAIC Data Volume Projections



Working Group Volume Projections

- Updated projections were needed to:
  - reflect current program milestones and deliverables
  - reflect refinements in records system screening practices introduced since SAIC studies were performed

Volume Projections - DOE Contributions

- DOE Headquarters and YMSCO records system contributions are lumped together
- Document licensing relevancy is not considered explicitly
- Volume projections are assumed coupled to the current program approach

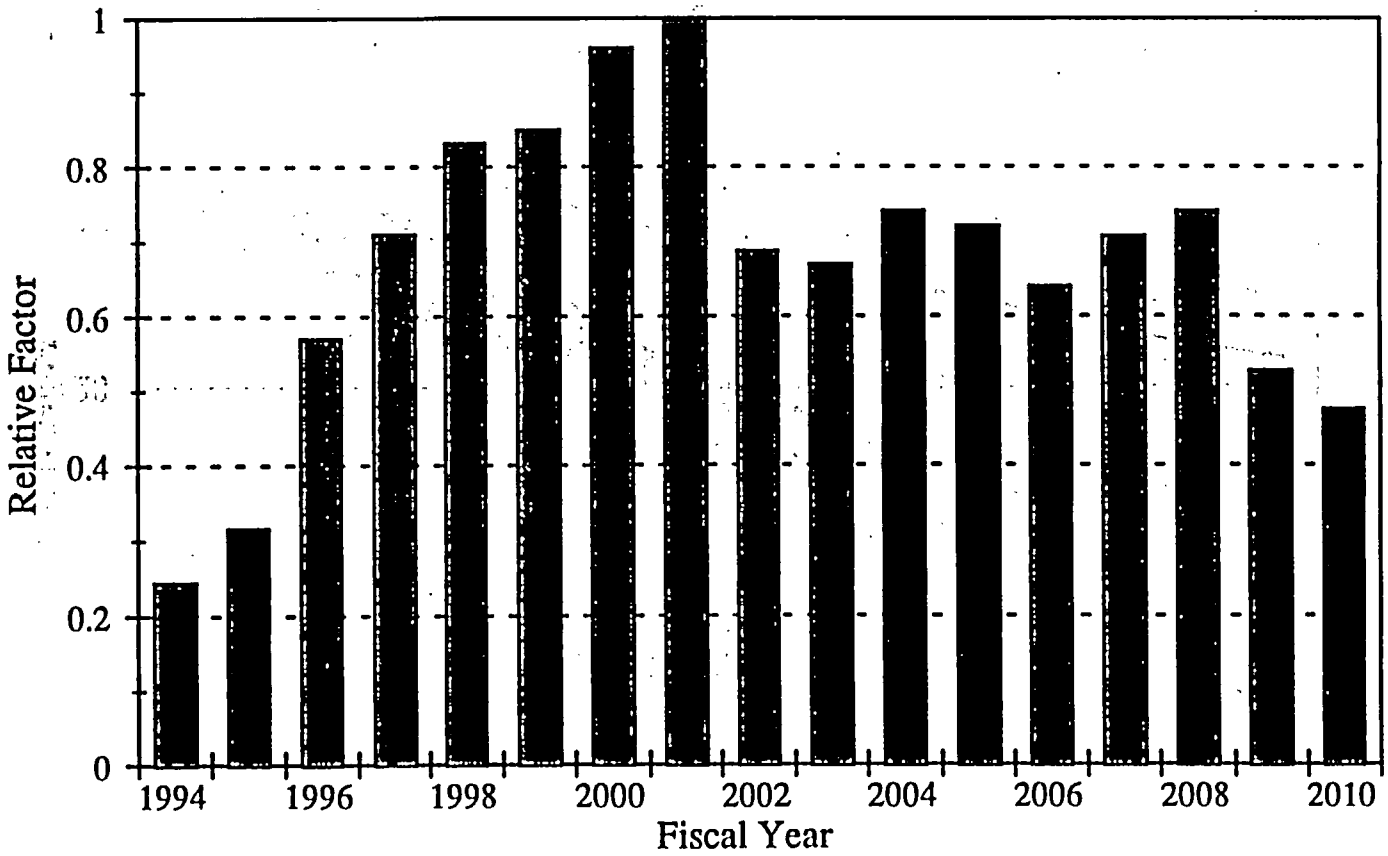
Volume Projections - NRC and Stakeholder

- NRC Contributions are assumed to grow linearly based upon 1990 and 1994 historical data obtained from NRC
- All NRC contributions are considered licensing relevant
- Stakeholder contributions are assumed to represent 10% of the NRC total

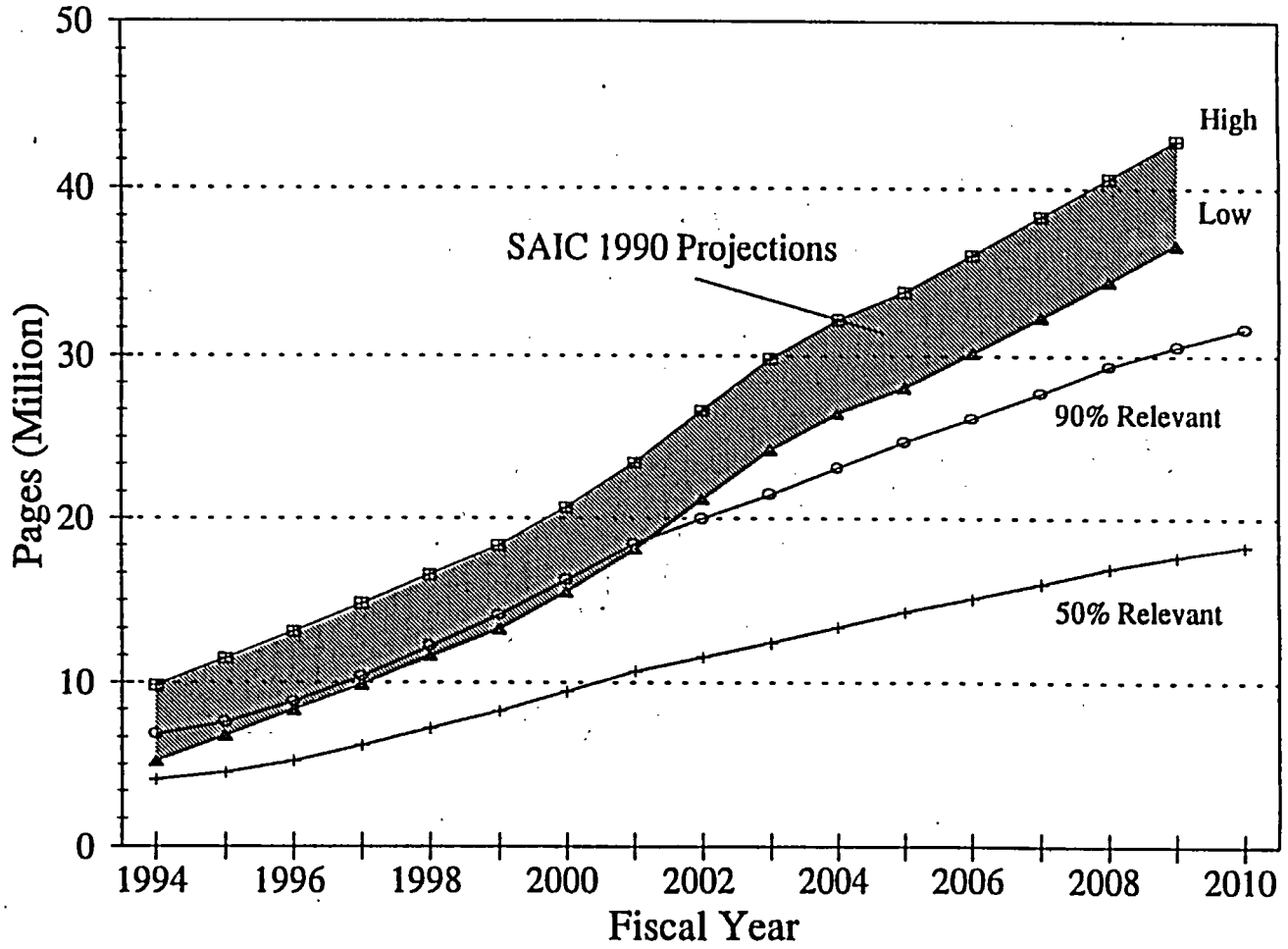
Volume Projections - Relative Volume

- A select group of project employees were asked to estimate the relative yearly records system activity. Group selection was based upon project experience and involvement with project planning
- The year with the peak activity is labeled as 1.0. All other activities are normalized to peak year
- Results were statistically combined to yield an average relative distribution

Relative LSS Volume



Comparison to SAIC Projections



Working Group Projection of LSS Data Volume

Year	OCRWM Pages/Year	OCRWM Cumulative	NRC Pages/Year	NRC Cumulative	Others Cumulative	Total Pages Added Yearly	90% Relevant Cumulative	50% Relevant Cumulative
1994	580,000	6,905,000	59,000	550,000	55,000	645,000	6,819,000	4,057,000
1995	750,000	7,655,000	59,000	609,000	61,000	814,000	7,559,000	4,497,000
1996	1,351,000	9,005,000	59,000	668,000	67,000	1,415,000	8,839,000	5,237,000
1997	1,682,000	10,687,000	59,000	726,000	73,000	1,746,000	10,417,000	6,142,000
1998	1,970,000	12,657,000	59,000	785,000	79,000	2,035,000	12,255,000	7,192,000
1999	2,013,000	14,670,000	59,000	844,000	84,000	2,077,000	14,131,000	8,263,000
2000	2,276,000	16,946,000	59,000	903,000	90,000	2,340,000	16,244,000	9,466,000
2001	2,371,000	19,317,000	59,000	961,000	96,000	2,436,000	18,442,000	10,716,000
2002	1,628,000	20,945,000	59,000	1,020,000	102,000	1,693,000	19,972,000	11,594,000
2003	1,584,000	22,529,000	59,000	1,079,000	108,000	1,649,000	21,462,000	12,451,000
2004	1,756,000	24,285,000	59,000	1,138,000	114,000	1,821,000	23,108,000	13,294,000
2005	1,708,000	25,993,000	59,000	1,196,000	120,000	1,772,000	24,709,000	14,312,000
2006	1,514,000	27,506,000	59,000	1,255,000	126,000	1,578,000	26,136,000	15,134,000
2007	1,674,000	29,181,000	59,000	1,314,000	131,000	1,739,000	27,708,000	16,036,000
2008	1,756,000	30,937,000	59,000	1,373,000	137,000	1,821,000	29,353,000	16,978,000
2009	1,247,000	32,184,000	59,000	1,431,000	143,000	1,312,000	30,540,000	17,666,000
2010	1,124,000	33,308,000	59,000	1,490,000	149,000	1,188,000	31,636,000	18,293,000

Volume Projections - Conclusions/Observations

- "Order-of-magnitude" projections are sufficient to support the LSS system design
- The uncertainty associated with the Working Group estimate is large
- Working Group will attempt to incorporate stakeholder volume estimates into final report

Civilian Radioactive Waste
Management System
Management & Operating
Contractor

TRW Environmental Safety
Systems Inc.

SUMMARY OF WORKING GROUPS CONCLUSIONS AND RECOMMENDATIONS

K. Michael Cline
December 12, 1994

B&W Fuel Company
Duke Engineering & Services, Inc.
Fluor Daniel, Inc.
INTERA Inc.

JK Research Associates, Inc.
E. R. Johnson Associates, Inc.
Logicon RDA

Morrison Knudsen Corporation
TRW Environmental Safety Systems Inc.
Winston & Strawn
Woodward-Clyde Federal Services

LSS WG Conclusions

- A certified LSS is very important for a timely LA submission. A timely LSS is part of DOEs licensing strategy.
- The LSS must be made operational and then certified six months prior to LA. The NRC and LSSARP expect an operational LSS well in advance of what the regulation calls for.
- The DOE's records management systems (RMS) is separate and distinct from the LSS but provides the majority of data for the LSS.

LSS WG Conclusions

(Cont'd)

- The total number of pages entered into the LSS through 2004 is estimated to be significantly less than previous studies suggest. The total number of pages entered into the system and pages disseminated are key cost divers.

- Care must be exercised in using the historical LSS work as a basis for decisions. The documentation is somewhat incomplete.

LSS WG Conclusion

(Cont'd)

- The WG identified a smaller set of LSS design requirements than previously documented.
- There exists a need to monitor technical issues to maintain user satisfaction and to minimize cost over the life of the LSS.
  - 50+ years of continuous and effective operation
  - rapid improvement in performance and reduction in cost of all LSS related technologies

LSS WG Conclusions

(Cont'd)

- A reevaluation of the LSS cost structure has identified a significant cost reduction from previous estimates. This reduction is due to:
  - identification of selected costs related to the RMS
  - improvements in technology; and
  - reduction in total labor costs due to improved OCR systems and other LSS components

LSS WG Conclusions

(Cont'd)

- The expected 10-year cost for the LSS is projected to be less than $80 million.
- A schedule for developing and implementing the LSS has been prepared that is consistent with the Program Approach.

LSS WG Conclusions

(Cont'd)

- A set of six LSS options were defined and evaluated.
  - All options satisfy Subpart J
  - The full range of options is represented in the analyses
  - A set of evaluation criteria were selected to discriminate among the options

LSS WG Conclusions

(Cont'd)

- Of the six options evaluated, options 5, 6, and 7 clearly rank higher.
- Options 2, 3, and 5 utilize human verified text, a significant cost driver.
- Options 5 and 6 offer significantly greater value to the user.
- Of the two, option 6 has a lower cost.

Features	Scenario					
	2	3	4	5	6	7
Relative Cost (10)	80	69	94	71	82	80
Operational Cost Risk (7)	14	42	28	35	63	70
Flexibility (7)	21	42	14	70	63	35
Text Accuracy (6)	60	60	48	60	48	48
Search Accuracy (6)	60	60	54	60	54	54
Image Display (6)	12	42	12	60	60	42
Responsiveness (6)	12	42	12	60	60	42
Image Print (3)	6	30	6	18	18	30
Total Score	**265**	**387**	**268**	**434**	**448**	**401**

Totals with Cost Values deleted	121	276	146	328	303	251

RECOMMENDATIONS

- The DOE should proceed immediately toward:
  - revising the System Requirements document for the LSS; and,
  - conducting the Analysis of Benefits and Costs (make versus buy analysis)

RECOMMENDATIONS
(Cont'd)

- Options 5 and 6 should be further evaluated for development. Both have:
  - greatest benefits to the user; and,
  - considerable flexibility.
- The DOE should proceed expeditiously with converting the current micro film based RMS to a image based RMS.

BACK UP MATERIAL

Evaluation Factors

- **Evaluation factors deemed to have sufficient measurement capability to differentiate among LSS options:**
  - *Text Accuracy* - **Number of errors in captured text relative to the original text**
  - *Search Accuracy* - **Percentage representative of the documents actually identified compared to the number that should have been identified**
  - *Response Time* - **Interval from the user input until response is complete (image response, printing response)**

Evaluation Factors

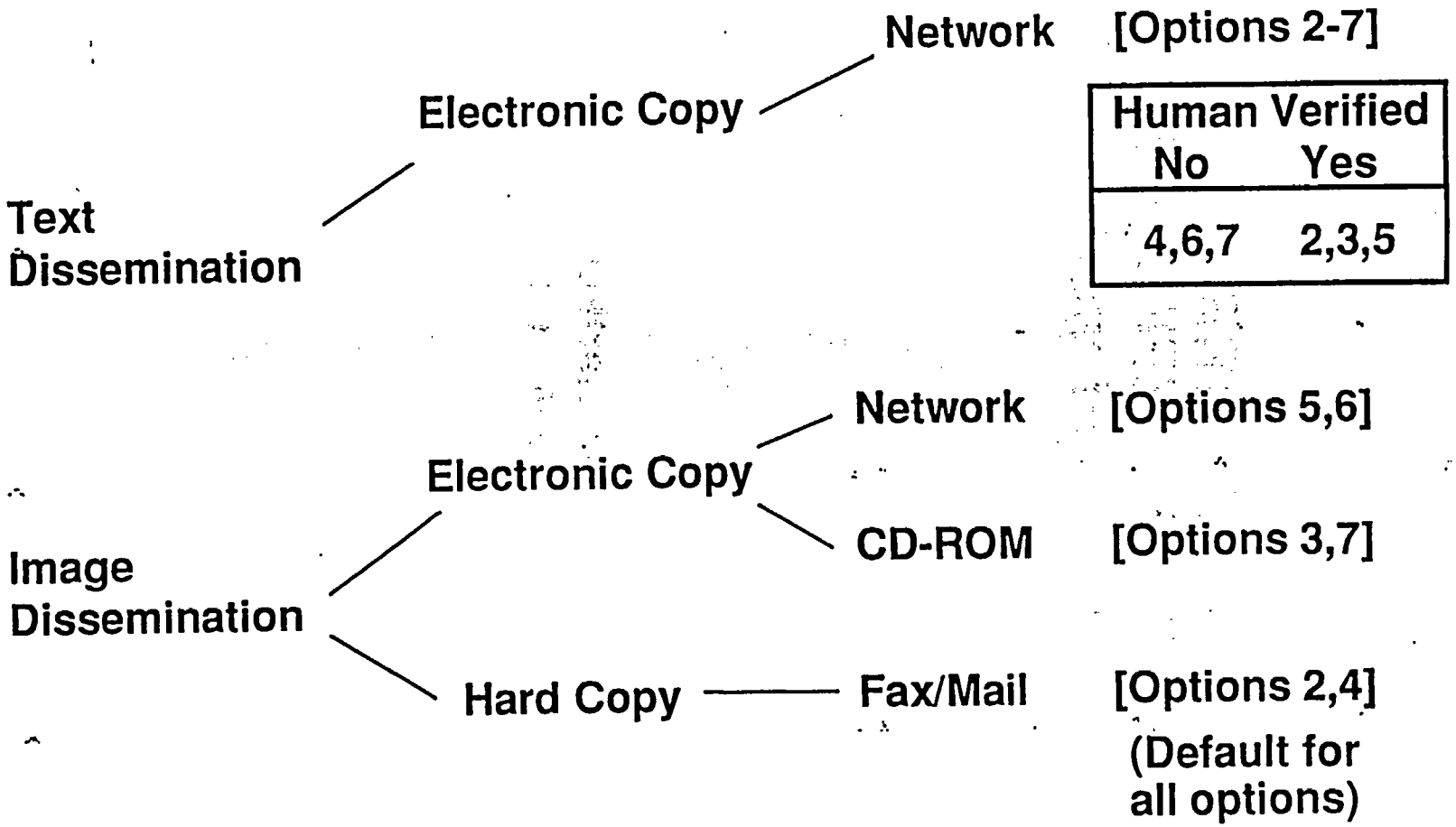
- *Operational Cost Risk* - Risk that actual costs will be higher than original estimates and whether that impact might be a high, medium, or low delta.

- *Cost* - Life cycle cost; 10 years

- *Flexibility* - Preservation through design of options until latest possible time

- *Responsiveness* - Rapid access to information on the image but not in the text

LSS Options Considered



PROGRESS TOWARD AN EFFICIENT LSS: An overview of work at UNLV

Tom Nartker

December 1994

UNLV/Information Science Research Institute

1990 → 1994

HOW HAVE LSS CRITICAL TECHNOLOGIES EVOLVED?

HOW WILL THESE CHANGES EFFECT LSS COST?

UNLV/Information Science Research Institute

WHAT ARE THE CRITICAL LSS TECHNOLOGIES?

1. OPTICAL CHARACTER RECOGNITION (OCR)

2. INFORMATION RETRIEVAL (TEXT RETRIEVAL)

WHAT IS MEANT BY CRITICAL?

OCR ACCURACY:
(What accuracy is achievable?)

TEXT RETRIEVAL EFFECTIVENESS:
(What accuracy is needed?)

If **ACCURACY NEEDED**

Is-Greater-Than

ACCURACY ACHIEVABLE

Then **MANUAL CORRECTION WILL BE REQUIRED.**

QUESTIONS?

ACHIEVABLE 1990? - NEEDED 1990? - COST 1990?

ACHIEVABLE 1994? - NEEDED 1994? - COST 1994?

Year:	**1990**
Company:	**CALERA RECOGNITION SYSTEMS**
OCR Product:	**RS9000**
Cost:	**$22,500.00**
Accuracy:	**97.4%**

NEEDED 1990 (Assumed) = 99.8%

ACHIEVABLE 1990 = 97.4%

COST 1990 ?

Cost of Labor for Converting a Typical LSS Page Assuming a 99.8% Accuracy Requirement

(Lois Dickey - SAIC - 1991)

Capture Step	Cost
Preprocess page & Scan Image	0.33
Zone Image	0.62
Correct OCR Output	2.56
Quality Control	0.28
TOTAL	$3.79

FOR 17,000,000 HARD COPY PAGES:
$2.56 X 17,000,000 = $43,520,000.00

NEEDED 1994 = ????%

ACHIEVABLE 1994 = ????%

COST 1994 ?

Year:	**1994**
Company:	**CALERA RECOGNITION SYSTEMS**
OCR Product:	**WordScan Plus 3.0**
Cost:	**$595.00**
Accuracy:	**98.5%**

Character Accuracy of Contemporary OCR Systems for a set of test Pages from the LSS-Prototype Database

(460 pages - 817,946 characters)

Year	OCR System	# Errors	% Accuracy
1992	Calera RS9000	21,351	97.4
1993	Calera WordScan	16,013	98.0
1994	Calera WordScan	12,459	98.5
1994	ISRI Voting	7,204	99.1

66% FEWER ERRORS TO CORRECT IMPLIES:
SAVINGS PROPORTIONAL TO
66% OF $43,520,000.00

NEEDED 1994 = ????%

ACHIEVABLE 1994 = 99.1%

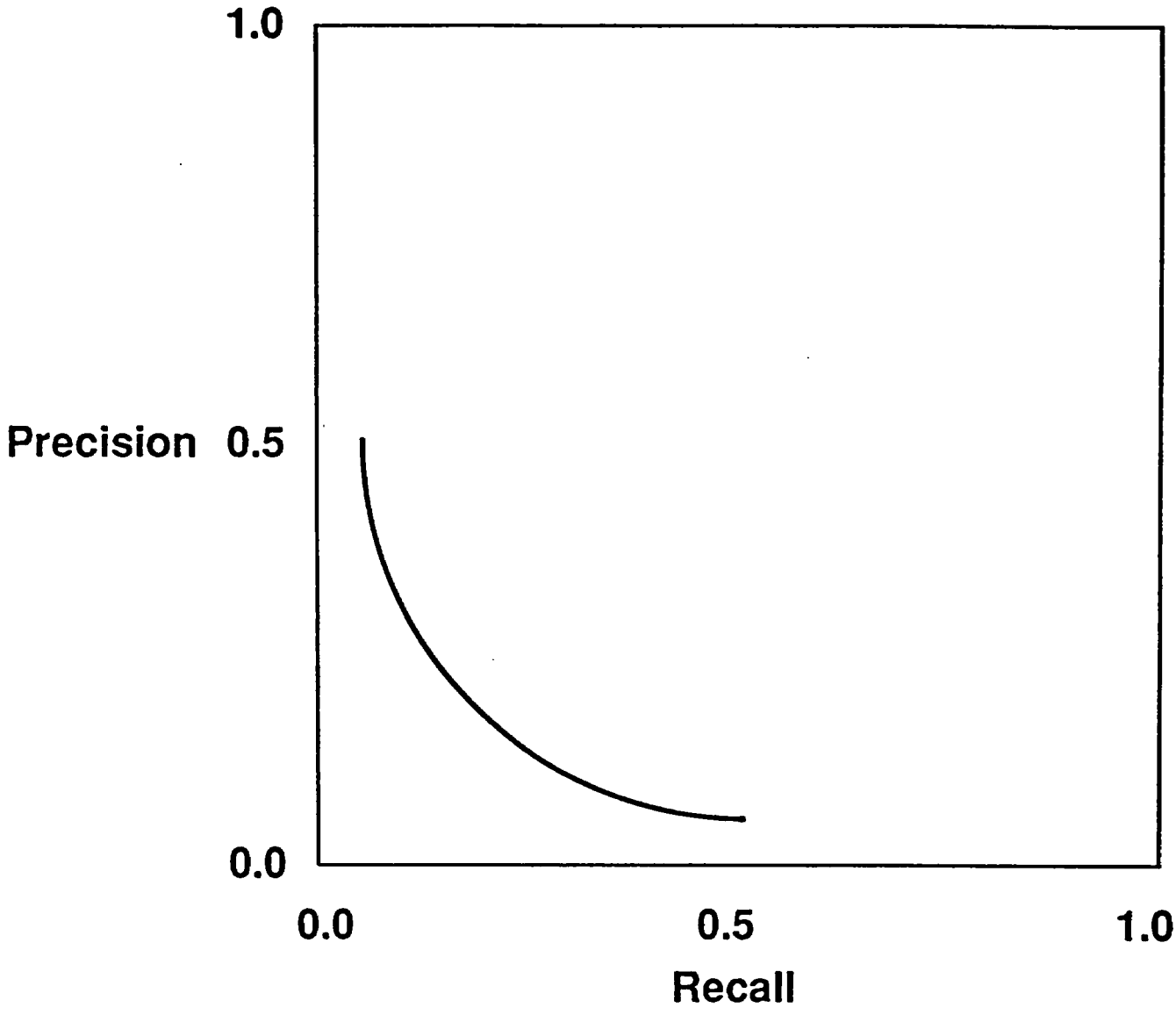
COST 1994 ?

RELEVANCE JUDGEMENTS

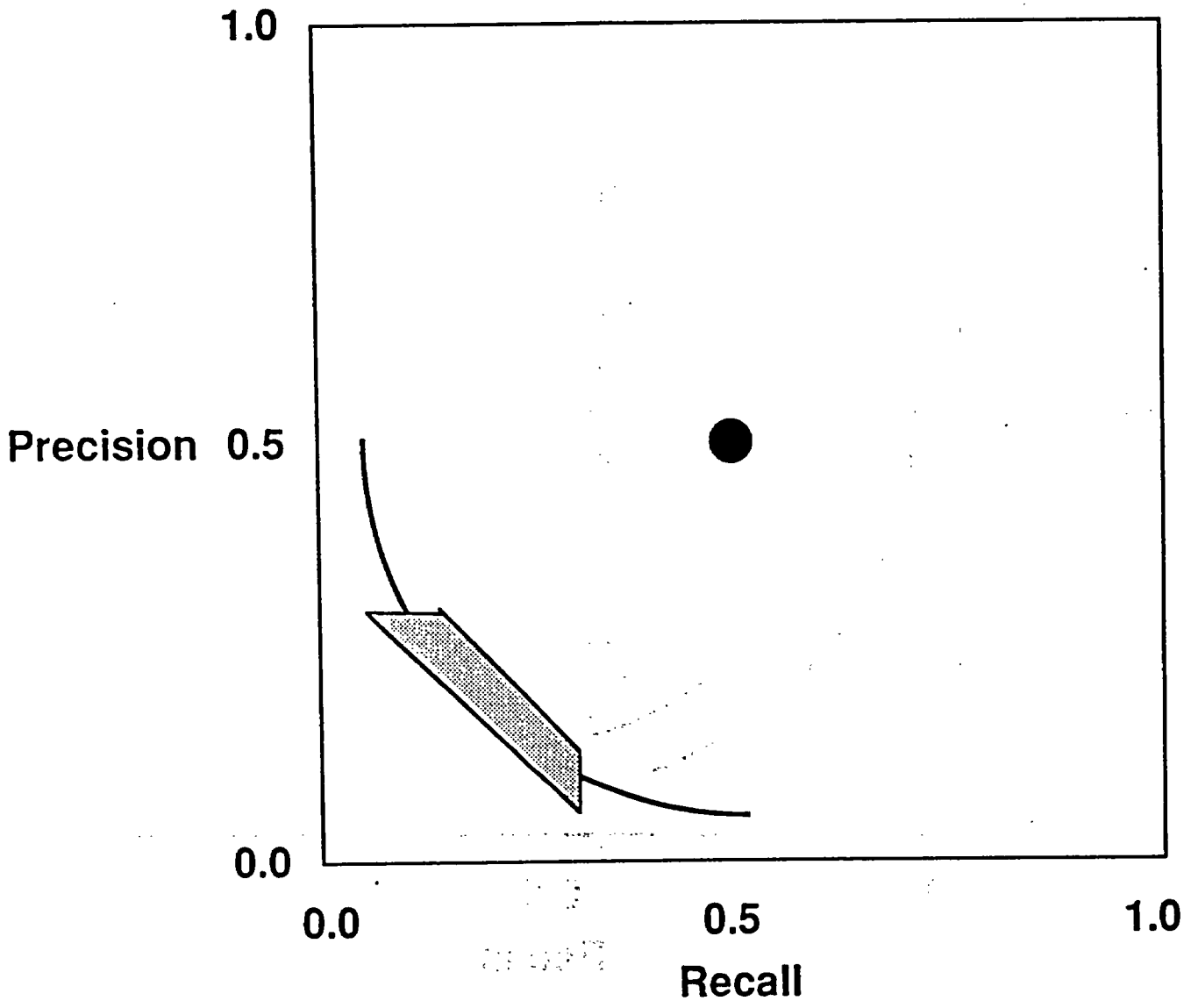
RECALL: the ratio of relevant documents retrieved for a given query over the number of relevant documents for that query present in the database.

PRECISION: the ratio of the number of relevant documents retrieved over the total number of documents retrieved.

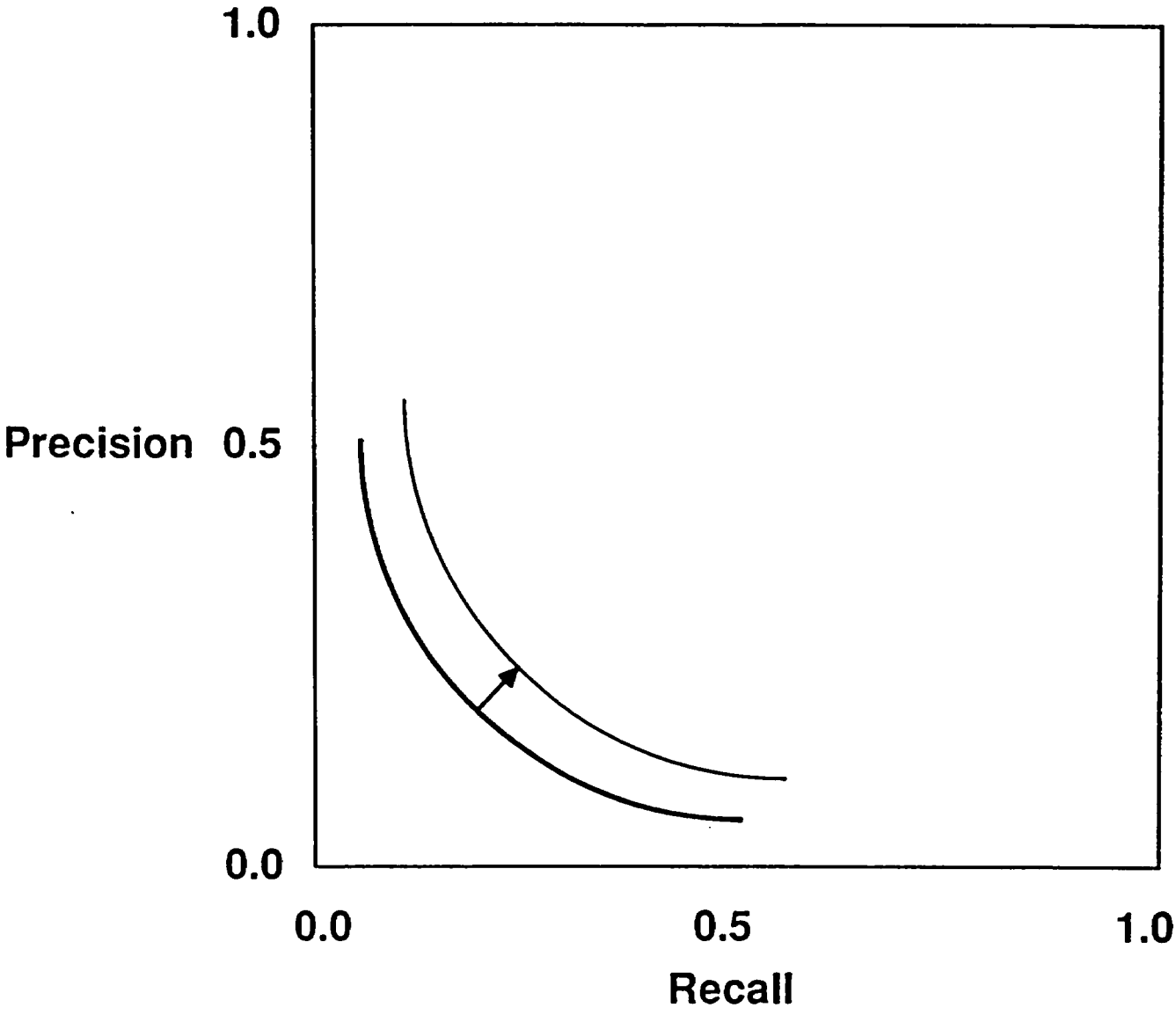
Both recall and precision take on values between 0 and 1.



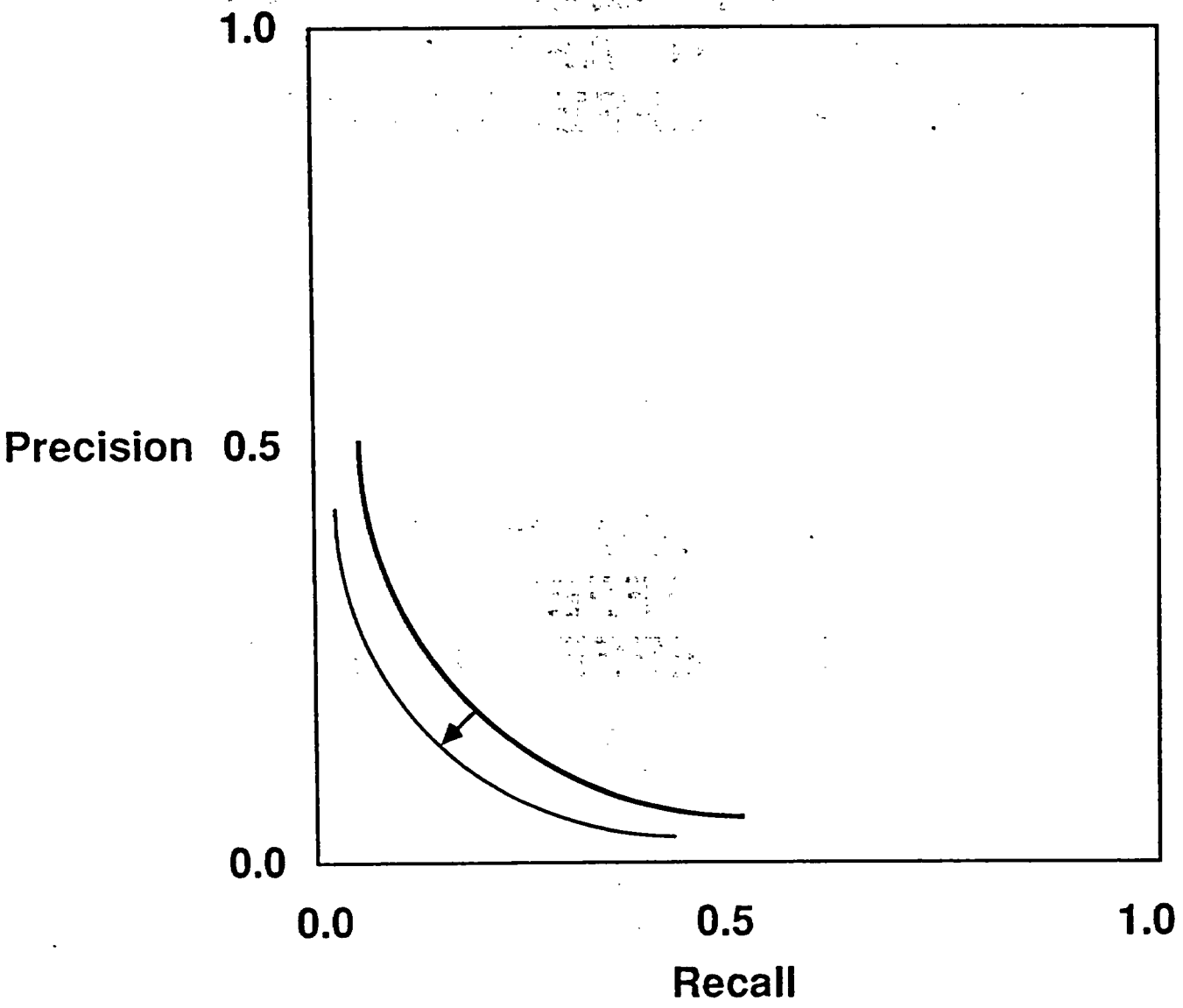
PRECISION/RECALL GRAPH
CHARACTERIZING A
TYPICAL TEXT-RETRIEVAL SYSTEM



PRECISION/RECALL GRAPH
SHOWING AREA
OF NORMAL USAGE



PRECISION/RECALL GRAPH
CHARACTERIZING
HIGHER RETRIEVAL EFFECTIVENESS



PRECISION/RECALL GRAPH
CHARACTERIZING
LOWER RETRIEVAL EFFECTIVENESS

INCORRECT CHARACTERS DO NOT AFFECT RETRIEVAL EFFECTIVENESS.

MISSPELLED WORDS DO AFFECT RETRIEVAL EFFECTIVENESS.

UNLV/Information Science Research Institute

MANICURE:

A SEMI-AUTOMATIC/ EXPERT SYSTEM BASED DOCUMENT EDITOR WHICH CORRECTS "WORDS AT THE DOCUMENT LEVEL" INSTEAD OF "CHARACTERS AT THE PAGE LEVEL"

MANICURE WILL ACCEPT
AUTOMATICALLY ZONED PAGES

SAVINGS = $0.62/PAGE IN ZONING COST
AND
AT LEAST $2.00/PAGE IN ERROR CORRECTION

$2.62 X 17,000,000 = $44,540,000.00

UNLV/Information Science Research Institute

SUMMARY

1990

ACCURACY ACHIEVABLE = 97.4%

ACCURACY NEEDED = 99.8%

CORRECTION (&MANUAL ZONING)
COST > $54,000,000.00

1994

ACCURACY ACHIEVABLE = 99.1%

ACCURACY NEEDED = (Word based accuracy)

REDUCED COST > $44,000,000.00

BACK UP MATERIAL

PURPOSE OF ~~SCENARIOS~~ OPTIONS

SEVEN SCENARIOS WERE DEVELOPED TO:

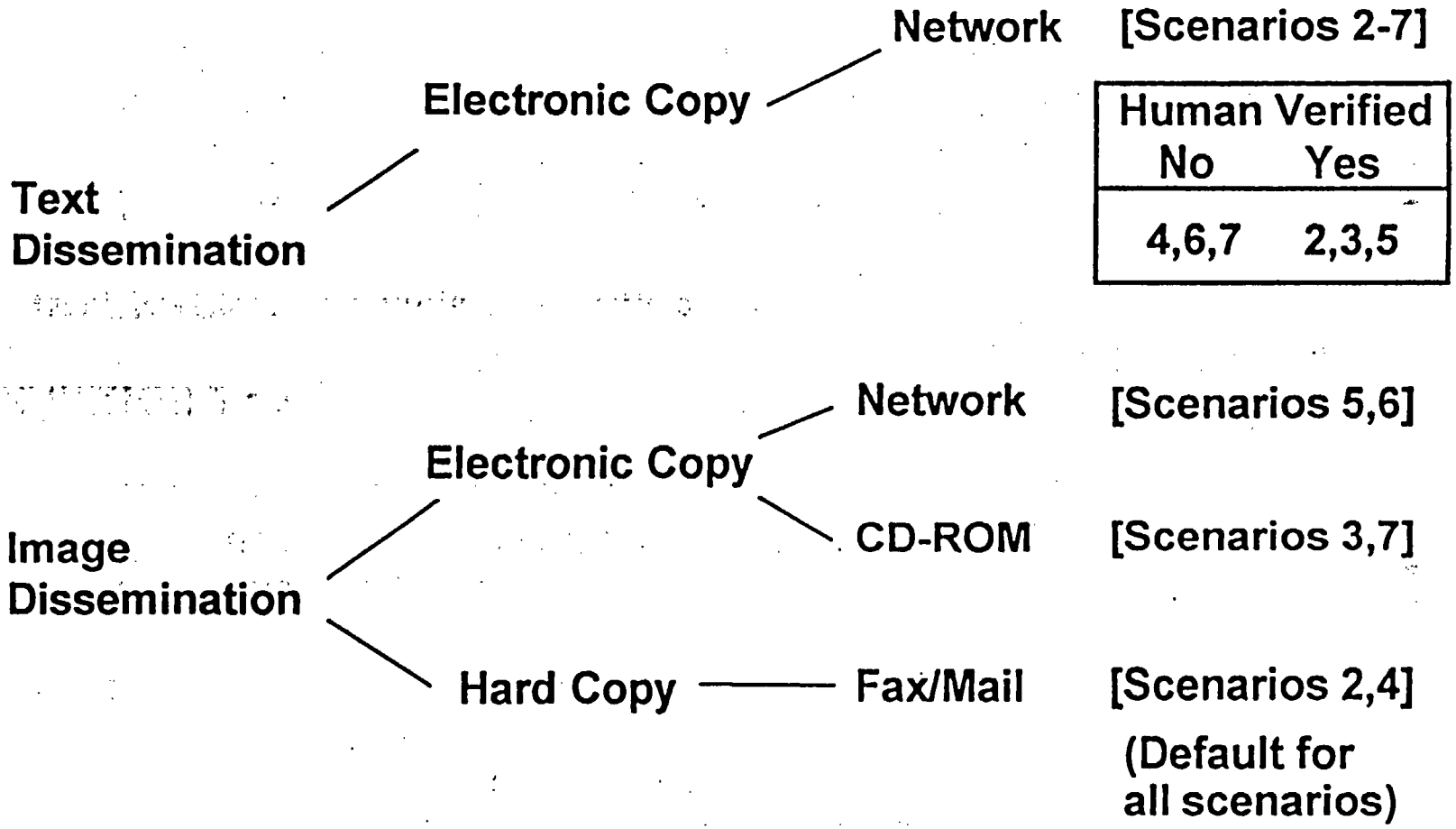
- **Identify principal operational concepts and the evaluation criteria (ECs) they may affect**
  - on-line transmission of bit-mapped image vs. mail transmittal of hard copy or CD-ROM
    - image availability response time (yes)
  - human-verified vs. machine corrected (only) Optical Character Reader (OCR'd) full text files
    - enhanced user confidence (maybe)
    - improved retrieval accuracy (probably not)
    - regulatory requirement for originator to verify files loaded into LSS (probably not)
    - purported user need to download full-text files for subsequent use (maybe; not critical)

- **Identify cost drivers, such as:**
  - human verification of machine-corrected OCR'd full text files
  - dissemination mechanisms:
    - implementation of an on-line bit-mapped image transmission system
    - use of wide-band communication lines for image transmission
    - handling mail transmittal of hard copy or CD-ROM

- **Cost a full array of potential LSS operational concepts**

LSS ~~Scenarios~~ Considered

OPTIONS



SUMMARY OF DIFFERENCES BETWEEN ~~SCENARIOS~~ OPTIONS

SCENARIO 1: (Current OCRWM records managment system, included as base-case for comparison only)

DIFFERENCES between Scenario 1 and other Scenarios:

- NOT Subpart J compliant
- Records stored on microfilm, no bit-mapped images
- Headers but no full text available for on-line search

SCENARIOS 2 - 7:

DIFFERENCES among Scenarios 2 - 7:

- Image available *only* by mail or fax (Scenarios 2, 3, 4, 7)
- On-line transmission of electronic images (Scenarios 5, 6)
- Human-corrected OCR'd text of documents (Scenarios 2, 3, 5)
- Machine-corrected (only) OCR'd text of documents (Scenarios 4, 6, 7)

DESCRIPTION OF ~~SCENARIO~~S OPTIONS

SCENARIO 1: (Current OCRWM records managment system as base-case for comparison)

- ***NOT*** **Subpart J compliant**
- **On-line searchable headers to support retrieval of documents**
- **No on-line searchable full text to support retrieval of documents**
- **Transmittal of hard copy blow-backs from microfilm per DOE and Participant user request**

DESCRIPTION OF ~~SCENARIOS~~ OPTIONS (Continued)

For Scenarios 2 - 7, the following features common to all are included in individual descriptions:

- **Are Subpart J compliant**
- **Use bit-mapped (scanned) images for records storage and to produce OCR'd full text for some documents**
- **Have on-line searchable headers to support retrieval of documents**
- **Have on-line searchable full text to support retrieval of documents**
- **Provide for transmittal of hard copy image by mail or fax per LSS user request**

DESCRIPTION OF ~~SCENARIOS~~ OPTIONS (Continued)

SCENARIOS 2, 3, 4 and 7: no on-line transmission of bit-mapped images

SCENARIO 2:

- On-line headers and human-corrected full text searched to identify documents for retrieval
- Hard copy (only) library available for image retrieval at required LSS stations
- No CD-ROMs available for transmittal by mail per LSS user request

SCENARIO 3:

- On-line headers and human-corrected full text searched to identify documents for retrieval
- CD-ROM library available for image retrieval at required LSS stations
- CD-ROM sent by mail per LSS user request

SCENARIO 4:

- On-line headers and machine-corrected (only) full text searched to identify documents for retrieval
- Hard copy (only) library available for image retrieval at required LSS stations
- No CD-ROMs available for transmittal by mail per LSS user request

SCENARIO 7:

- On-line headers and machine-corrected (only) full text searched to identify documents for retrieval
- CD-ROM library available for image retrieval at required LSS stations
- CD-ROM sent by mail per LSS user request

DESCRIPTION OF ~~SCENARIOS~~ OPTIONS (Continued)

SCENARIOS 5 and 6: <u>on-line transmission of bit-mapped images</u>

SCENARIO 5:

- **On-line headers and <u>human-corrected full text</u> searched to identify document**
- **On-line access to image means <u>no separate image library</u> needed at required LSS stations**
- **<u>CD-ROMs available</u> for transmittal by mail per LSS user request**

SCENARIO 6:

- **On-line headers and <u>machine-corrected (only) full text</u> searched to identify document**
- **On-line access to image means <u>no separate image library</u> needed at required LSS stations**
- **<u>CD-ROMs available</u> for transmittal by mail per LSS user request**

Evaluation Criteria Used to Descriminate Among ~~Scenarios~~ OPTIONS

CRITERION	DEFINITION	WEIGHT
Cost	Total cost to OCRWM Records System + LSS life-cycle cost	10
Response Time (Image)	Responsiveness in receiving image at LSS stations	6
Response Time (Printing)	Responsiveness in printing text files at LSS stations	3
Operational Cost Risk	Low risk of cost increase for erroneous assumptions	7
Flexibility	Ease in expanding LSS capabilities in response to user needs	7
Text Accuracy	Increase in user confidence w/human verified text	6
Search Accuracy	Increase in search success w/human verified text	6
Responsiveness	Capability to provide most up-to-date text/image/motions	6