Official Transcript of Proceedings

NUCLEAR REGULATORY COMMISSION

Title: Advisory Committee on Reactor Safeguards
Joint Digital I&C and Plant Operations

Docket Number: (n/a)

Location: teleconference

Date: Wednesday, November 16, 2022

Work Order No.: NRC-2171 Pages 1-166

NEAL R. GROSS AND CO., INC.
Court Reporters and Transcribers
1716 14th Street, N.W.
Washington, D.C. 20009
(202) 234-4433

1
2
3
4 DISCLAIMER
5
6
7 UNITED STATES NUCLEAR REGULATORY COMMISSION'S
8 ADVISORY COMMITTEE ON REACTOR SAFEGUARDS
9
10
11 The contents of this transcript of the
12 proceeding of the United States Nuclear Regulatory
13 Commission Advisory Committee on Reactor Safeguards,
14 as reported herein, is a record of the discussions
15 recorded at the meeting.
16
17 This transcript has not been reviewed,
18 corrected, and edited, and it may contain
19 inaccuracies.
20
21
22
23

UNITED STATES OF AMERICA

NUCLEAR REGULATORY COMMISSION

+ + + + +

ADVISORY COMMITTEE ON REACTOR SAFEGUARDS

(ACRS)

+ + + + +

JOINT DIGITAL I&C AND PLANT OPERATIONS SUBCOMMITTEES

+ + + + +

WEDNESDAY

NOVEMBER 16, 2022

+ + + + +

The Subcommittee met via hybrid in-person and Video-Teleconference, at 8:30 a.m. EST, Vicki Bier, Chairman, presiding.

COMMITTEE MEMBERS:

VICKI BIER, Chair

RONALD G. BALLINGER, Member

CHARLES H. BROWN, JR., Member

VESNA DIMITRIJEVIC, Member

GREGORY HALNON, Member

WALT KIRCHNER, Member

JOSE MARCH-LEUBA, Member

DAVID PETTI, Member

JOY L. REMPE, Member

MATTHEW SUNSERI, Member

ACRS CONSULTANTS:

DENNIS BLEY

MYRON HECHT

STEPHEN SCHULTZ

DESIGNATED FEDERAL OFFICIAL:

CHRISTINA ANTONESCU

ALSO PRESENT:

REED ANZALONE, NRR

LUIS BETANCOURT, RES

MATT DENNIS, RES

TREY HATHAWAY, RES

BRIAN GREEN, NRR

JOSHUA KAIZER, NRR

GENE KELLY, Constellation Energy

THERESA LALAIN, RES

THIAGO SEUACIUC-OSORIO, EPRI

JAMES SLIDER, NEI

RICHARD SZOCH, Constellation Energy

CONTENTS

Call to Order 4
Opening Remarks 6
Introductory Remarks 9
Discussion of Draft AI Strategic Plan 12
Discussion of EPRI AI activities 73
Corrective Action Program Automation and
Informed Inspection 128
Public Comments 166
Closing 166

P R O C E E D I N G S

8:30 a.m.

CHAIR BIER: The meeting will now come to order. Good morning, everybody. Before we get too far downstream, can somebody who's online let me know if they can hear me? Vesna?

MS. ANTONESCU: Yes, we can hear you, Vicki.

MEMBER DIMITRIJEVIC: We can hear you.

CHAIR BIER: Okay. So, this is a meeting of the Joint Digital I&C and Plant Operations Subcommittee. I am Chair of this --

MEMBER DIMITRIJEVIC: Now we cannot hear you.

MS. WALTON: Vicki, we can no longer hear you.

MEMBER DIMITRIJEVIC: Yes, we cannot hear anything.

CHAIR BIER: All right. Let's try again. Now do you hear me?

MEMBER DIMITRIJEVIC: Yes.

CHAIR BIER: Okay, great. All right. I will go ahead and start the roll again.

So, ACRS members in attendance are: Charlie Brown; Greg Halnon; Joy Rempe; myself, Vicki

Bier; Dave Petti; Jose March-Leuba; Ron Ballinger. I believe Matt Sunseri is online, is that correct?

MEMBER SUNSERI: Yes. Yes, I'm here, Vicki.

CHAIR BIER: Okay. Vesna I know is online. Is Walt Kirchner calling in online?

MEMBER KIRCHNER: I'm here. Thank you.

CHAIR BIER: Awesome. Great. And we have some consultants with us. Steve Schultz in person. I believe Dennis Bley is available online, correct?

MR. BLEY: Yes. I am.

CHAIR BIER: And is Myron Hecht available?

(No response.)

CHAIR BIER: Maybe not. I know he viewed the slides ahead of time and sent some comments.

Christina Antonescu of the ACRS staff is the Designated Federal Official for this meeting.

The court reporter, are you on? Can you hear correctly?

MS. ANTONESCU: Yes, he is.

CHAIR BIER: Okay. Great. Thank you, you. The purpose of this meeting is for the staff to brief the Subcommittee on the NRC's Draft Artificial Intelligence Strategic Plan for Fiscal Years 2023 to 2027, and we will also hear about the EPRI AI

activities and Constellation Energy's Corrective Action Program for Automation and Informed Inspection.

For background, the ACRS was established by statute and is governed by the Federal Advisory Committee Act, FACA. That means that the Committee can speak only through its published letter reports. We hold meetings to gather information to support our deliberations.

Interested parties who wish to provide comments can contact our office requesting time. That said, we also set aside time for comments from members of the public who are attending or listening to our meetings, and written comments are also welcome.

The agenda for today's meeting was published on the NRC's public meeting notice website, as well as the ACRS meeting website.

On the agenda for the meeting and on the ACRS meeting website are instructions for how the public can participate. As of this time, I believe that we have not received a request for making a statement to the Subcommittee, but will still reserve time for such statements, if they arise.

Due to COVID-19, we are connecting today as a hybrid meeting.

A transcript of the meeting is being kept

and will be made available on the Committee website.

In order to support an accurate transcript -- now I have an echo. All right. Thank you.

MEMBER MARCH-LEUBA: I wouldn't admit it.

(Laughter.)

CHAIR BIER: Thank you. In order to support an accurate transcript, we request that participants in this meeting should first identify themselves before they speak and speak with sufficient clarity and volume so that they can be readily heard.

Presenters should also please pause from time to time to allow members to ask questions during the presentations, and indicate the slide number you are on when moving to the next slide, so that people who are online can follow along.

We have the Microsoft Teams phone line established, audio only for the public to listen to the meeting.

Based on our experience from previous virtual and hybrid meetings and from this morning already, I remind the speakers and presenters to speak slowly and keep themselves on mute when they are not presenting.

We will take a short break after each presentation to allow time for screen-sharing, and we

have additional breaks, if needed, during long presentations.

Please do not use any virtual meeting features such as chat to conduct sidebar technical discussions. If you are listening and have any technical questions that you want to raise, you can contact the Designated Federal Official, Christina Antonescu, and she will let us know if anything needs to be addressed.

We'll now proceed with the meeting. As such, I would like to ask Matthew Dennis, Reactor Systems Engineer of the Accident Analysis Branch in the Office of Nuclear Regulatory Research, to share his screen. And I believe we will also have introductory comments from Theresa Lalain, Deputy Director of the Division of Systems Analysis in the Office of Nuclear Regulatory Research.

MR. HECHT: Vicki, this is Myron Hecht. I just wanted to let you know that I had audio problems earlier, but I'm now on the line.

CHAIR BIER: That's fine. Thank you for joining us. Glad to have you.

So, Theresa? Are you ready, Teri?

MS. LALAIN: Thank you. This is Teri Lalain. Good morning, ACRS Members. Thank you for

the opportunity to highlight the agency's proactive effort to ensure readiness to regulate artificial intelligence in NRC-regulated activities.

I also want to thank NEI, EPRI, and Constellation for their participation and presentations to the ACRS today.

As I said, I'm Teri Lalain, the Deputy Director of the Division of Systems Analysis, and serve as the executive sponsor for AI. I joined the NRC in 2021 from the Department of the Army, where during my 18-year tenure there, it included readiness to assess new technologies being deployed into the field.

Our goal for today is to inform the ACRS and participants on the current and future state of practice of AI in the industry and what the agency is doing in parallel to prepare for this future.

AI technology interest is widespread in many commodity areas. In FY21, we hosted a series of data science and AI regulatory application public workshops to better understand the nuclear industry's activities and plans for AI, including potential uses in the future in NRC-regulated activities. Based on the feedback from the public workshops, licensing applications that may include the use of AI

technologies may be submitted to the NRC for review in the near future.

We developed the AI Strategic Plan focused on the external use of AI in NRC-regulated activities. It also facilitates that conversation, so we can work jointly to move forward in this area. The NRC plans to engage ACRS, industry, and other stakeholders early and often throughout the agency's work on AI, particularly when developing guidance.

I look forward to the conversation today, and I'm going to turn over the opening remarks to Jim Slider. Thank you.

MR. SLIDER: Good morning. I'm James Slider. I'm the Technical Advisor for Innovation and Regulatory Affairs at the Nuclear Energy Institute. In this role, I have lead responsibility for helping smooth the way for our members to apply technological innovations in their operating nuclear power plants. This includes identifying and addressing regulatory challenges to the deployment of artificial intelligence and other advanced technologies.

I appreciate this opportunity to speak with you today on the subject of artificial intelligence. We at NEI appreciate your interest in this very important and, for our industry, cutting-

edge topic.

We applaud the NRC staff for publishing their Draft Strategic Plan on Artificial Intelligence this summer and their forward thinking on potential regulatory improvement with the AI in the years ahead.

We see great promise for artificial intelligence to enhance safety and reliability, to improve the quality of decisionmaking, and to streamline business processes. As you will hear later, industry is developing and piloting AI-based tools in several areas. These early applications are giving industry opportunities to learn about AI and gain practical experience with it. This experience is helping industry to quantify the benefits and costs of AI to better inform the business cases for more ambitious applications yet to come.

Thank you for convening this meeting and I look forward to the discussion.

CHAIR BIER: Okay. I think whoever is the next speaker can proceed. If that's Matt?

MR. HATHAWAY: I don't have a share.

CHAIR BIER: Okay.

MR. DENNIS: This is Matt Davis.

We did have a share.

CHAIR BIER: I may be able to share.

MR. HATHAWAY: It should be Trey Hathaway.

CHAIR BIER: All right. Thank you. Okay, we can now proceed.

MR. DENNIS: All right.

MEMBER MARCH-LEUBA: Say your name for the court reporter.

MR. DENNIS: Okay. Good morning, everyone. My name is Matt Dennis in the Office of Research, Accident Analysis Branch. And this morning we'll be speaking to you about an overview of the Draft NRC Artificial Intelligence Strategic Plan for fiscal years 2023 through 2027. In the future, if I just say, "Strategic Plan," that's what I'm referring to, not the agency's Strategic Plan. I'm referring to this Artificial Intelligence Strategic Plan.

This morning we're going to discuss the Draft Plan, provide a background on its origins, and review the progress and path forward for how the NRC must be prepared to review and evaluate AI technologies in NRC-regulated activities.

Next slide, slide No. 2.

As I mentioned, my name is Matt Dennis. I'm a Reactor Systems Engineer/Data Scientist in the Office of Research in the Accident Analysis Branch, and with me here I have Dr. Trey Hathaway.

Trey, if you'd like to introduce yourself?

MR. HATHAWAY: Yes. My name is Trey Hathaway, Office of Research, Accident Analysis Branch.

MR. DENNIS: Thanks, Trey. And you've already heard from Dr. Teri Lalain, our Deputy Division Director, and joining her, also, in the back is Luis Betancourt, the Branch Chief for the Accident Analysis Branch. Teri and Luis are serving as the management AI champions for the agency. So, they're helping direct everything from the management side.

So, a couple things I want to mention, also, before we move on is that the Strategic Plan is being led in development by the Office of Research, and we're leading the development of the Strategic Plan with expertise in AI-related topics to ensure that we're capturing a broad spectrum of technical disciplines and use cases around AI, such that the NRC can be best prepared for future use, or future review and evaluation of AI usage in NRC-regulated activities.

The presenters today come with lessons learned -- and this also goes for the other presentations -- lessons learned from developing program strategy and development and deployment of AI

applications across the industries. Presenters also have been early investigators on the potential uses of AI and evaluating those impacts at the agency.

The next slide, slide No. 3.

So, just to review the agenda for our presentation this morning, we'll be discussing the artificial intelligence landscape outside the agency and at the NRC briefly. We will discuss the Draft Strategic Plan development background, and we'll give you an overview of the Strategic Plan and its goals, and then, discuss where we are currently and what our next steps are with development of the Strategic Plan.

Next slide, slide 4. All right. Trey is going to take it over now.

MR. HATHAWAY: Thank you, Matt. Can everyone hear me okay? Is that better? Okay, thank you. Well, I talk quiet, so I've got to enunciate.

So, my name is Trey Hathaway of the Accident Analysis Branch in the Office of Research. And as Matt said, I'm just going to give you sort of a high-level overview of the contents of the Strategic Plan, and then, pass it back to Matt to go into some of the details. But I'm also going to give you some of the motivation for why the Strategic Plan was developed.

There were numerous drivers for the development of the Strategic Plan. They were external and internal, but, as sort of been said previously, the goal of the Strategic Plan is to focus on the external applications.

And when I say, "external," we recognize that there is interest in the use of artificial intelligence based off efforts within the agency and various sort of AI developers looking at AI solutions, as well as research performed by the National Labs. And the agency desires to be ready for the efficient review of potential applications. Therefore, the Strategic Plan is intended to lay the foundation for what would be required to facilitate efficient reviews.

There are some external drivers that we sort of recognize. One is Executive Order 13960, which is pushing agencies to consider how they implement AI internally. And there's also the Evidence Act, which necessitates the agency to make its data more accessible for constructive purposes, including the potential usage of AI applications to increase effectiveness and efficiency. But these are internal considerations and not the focus of the Strategic Plan.

So, I'm moving to slide 5. The development of the Strategic Plan is for the agency to take a proactive approach in order to be prepared for potential artificial intelligence applications that may require regulatory review in the future.

Prior to the pursuit of the strategy, exploration in artificial intelligence was sort of grassroots efforts spread across the agency, and the strategy was intended to focus some of these efforts.

We had a series of workshops and public meetings in 2021, so we could share our efforts with the industry, as well as here, about the applications from the industry and National Labs.

We surveyed information available from other federal agencies to inform the development of our Strategic Plan and maintained awareness of ongoing activities through participation in workshops and meetings or through Memorandums of Understanding with EPRI and the Department of Energy.

We finalized the --

MR. BLEY: I think we've lost the audio out here.

MR. HATHAWAY: Do I still have an echo? Okay. Sorry.

We finalized the Strategic Plan and made

it available for public comment in the summer of 2022, and we are currently working on addressing public comments that we received.

So, next slide, slide 6. Are there any questions?

MEMBER BROWN: Yes, I'll ask you one. I know you said you had public comments that were addressed. Just a curiosity question, and since you spent this time with the public, as well as, I presume, industry participated in that, did anything pop up that populated into your presentation or --

MR. HATHAWAY: We have a backup slide on comments. We could talk about it now or later.

MEMBER BROWN: Okay. I'll hold it for that until you're done.

MR. HATHAWAY: Okay. So, slide 6. Development of the AI Strategic Plan began in early 2021 and was born out of the need to address the data availability and usage of artificial intelligence applications. That was also the beginning of the recognition that the nuclear industry was looking at using artificial intelligence and machine learning.

As I said previously, I'm just going to give a high-level overview of the Strategic Plan and Matt's going to provide more discussion about the

individual goals.

The vision of the Strategy is to ensure the safe and secure use of AI in NRC-regulated activities. There are two high-level outcomes. Establish a framework to review the use of AI and, also, develop an AI-skilled workforce.

To achieve those high-level outcomes, we identified five strategic goals. These goals are intended to complement each other, to ensure its success, not necessarily we need to complete this goal to complete the next goal. They are really kind of woven.

Goal One is to allow us to establish a regulatory framework as we did.

Goal Two establishes an internal structure to support AI activities within the agency. And this will focus the efforts across the agency to ensure AI activities are explored efficiently and to prevent duplication of effort.

Goal Three is to help is build partnerships to leverage the knowledge and lessons learned from other organizations.

Goal Four allows us to ensure we have a staff with sufficient AI literacy, so they'll be able to ask the right questions when confronting a review

that may contain an artificial intelligence component. And so, you can think about that as: and they ask the right questions if they're confronted with something.

And finally, Goal Five is the pursuit of use cases to build foundational knowledge with the technology.

We also include in Appendix B that it's just sort of to acknowledge that there is the potential to have internal usage, and those efforts, when undertaken in the Strategic Plan, could be leveraged to assist in those internal applications.

So, slide 7. Finally, there are many different definitions of artificial intelligence and a broad array of algorithms and really ways to combine algorithms to create solutions. So, we tried to create a notional table of AI and autonomy for uses of artificial intelligence within commercial nuclear activities. And again, these are notional levels. They're intended to expand and contract for further consideration or examination.

The notional levels are based on the level of human oversight or the human in the loop the system relies upon, when allowing more machine independence would entail the potential for higher regulatory scrutiny. Each level would require an assessment of

the regulatory scrutiny needed, and success would require early engagement by licensees on the potential uses.

MEMBER HALNON: Trey, this is Greg Halnon.

You know, as we get into this -- and maybe Matt will get further into more detail -- but that regulatory scrutiny is obviously the "purple word" right now, right? Because we don't know what we're going to regulate because we don't know what system is out there. If you don't know what system is out there, you can't formulate a regulation and do it.

So, there's a situation here where we're kind of looking at each other saying, who's going to make the first move? And these different levels, and where you're going to insert the regulatory regime to do oversight and maybe review and approve prior to operation -- and I think Teri mentioned, you know, we're looking for future applications, more what regulatory oversight we need.

As you go, are you trying to get out in front of that? Because you're going to be asked to review some conceptual designs, things that haven't had a lot money, but here we've got an idea. Because without that bright line of how the regulatory regime fits into the design process, you're not going to

develop a very good business case to spend a lot of money in designing a system without knowing how that regulatory oversight, inspection, renewal, and approval system is going to work, because that's such a significant portion of -- well, it's a go/no go for one thing. And it's a significant portion to the business case from the dollar and resource perspective.

So, as we get into today, I'd like to explore that a little more and how we're going to try to break that. It's almost like a logjam, you know; it's never going to go anywhere unless we get some inching forward towards a regulatory oversight plan and review and approval-type document.

So, I just wanted to put that out there. Because I saw this and I don't know where that NRC regulatory oversight would fit into what level. And I assume it's, you know, the big-picture process, thinking it's going to get more and more. As you get more and more autonomous, you'll have more and more oversight, more and more inspection, more and more verification and validation, and that sort of thing.

But that's so conceptual at this point. I know that it's just early. So, there's probably no direct answer, but, as we go through this, I'd like to

kind of continue to explore that process.

MEMBER REMPE: So, along those lines, it's not just the more autonomous where you might have more regulatory oversight. It's the importance to the plant safety, right?

MEMBER HALNON: Right.

MEMBER REMPE: And I'm wondering, I mean, you've had all these workshops and you've reached out not only domestically, but internationally. Do you have any good examples where they have developed a process to decide how much they need to oversee various types of applications that give you some good insights?

MR. HATHAWAY: I guess my initial comment is I can provide you some potential insights, but would you want to wait for the industry, who is going to come up soon to discuss some of the applications that they are actually considering at this moment in time?

MEMBER REMPE: Okay, but it's not just their applications and nuclear-related, but what about other organizations within the government that you've tried to interact with? Or have they said, okay, this is just a "no, never mind"; it doesn't impact safety or risk? You can do this, but other things -- well,

if you're dealing with what we call safety-related systems, structures, and components, well, we need a little more oversight. Has anybody kind of developed a risk-based oversight for AI? is what I'm trying to go to.

MR. DENNIS: I can say that we have. So, this is later on I was going to mention this anyway. But we have taken the action early on in this to do a lot of survey and reference across, as you mentioned, internationally and domestically.

And at least one good case that came up was the FDA has put out a risk-based framework for assessing medical devices. So, they have some guidance documents out there that we have looked at, as like an example of how they're regulating the use of AI in medical devices.

MEMBER REMPE: That's good. Thank you.

MR. DENNIS: Yes. So, that's one example in the federal government.

MEMBER PETTI: Is the FAA doing anything? Because I know there's a lot of discussion about how many pilots in the cockpit, you know, all of that.

I mean, I'm trying to figure out, where is NRC on the government side? Are you ahead of everybody else? Kind of middle of the pack? Or

trailing relative to other government agencies?

MR. DENNIS: At this stage, we're somewhere -- the point of the Strategic Plan is to get us to be ahead of it, but we're somewhere slightly behind it. And that's, from our perspective, the industry is not ahead of it, either.

The use cases that we've seen so far, as we'll see later on in the presentations, are business-process-oriented. And so, right now, they're not in that vein.

One example I will mention is the IAC standard. IAC is working on an AI standard for nuclear applications. And in their specific description of the standard, it's assumed that there will not be -- it's for non-safety applications. So, even the AI standards for nuclear right now are very much focused on the non-safety aspect of it.

And to answer the question about FAA, yes, that was another one we surveyed. And so, the FAA is actively working on other applications, you know, their own applications of AI in the aviation industry.

MEMBER HALNON: This is Greg. Do you see any application of the NRC oversight in non-safety applications? I mean, are you thinking that there is going to be some NRC involvement in non-safety, even

if it's a secondary system or a monitoring system?

MR. DENNIS: To go back to the question about the levels of autonomy, we don't know right now. The short answer is we are keeping abreast of what industry has presented as far as applications at this point. And we don't exactly know where in the level 2, level 3, level 4 -- as you mentioned, it's probably apparent that somewhere in there there's going to be a transition point, level 2, level 3, but at this point, no.

MEMBER HALNON: In your workshops, were you able to put out that type of question on these levels with the industry to see? I just feel uneasy being in a reactionary mode only, but I understand where you're at. We're very early.

MR. DENNIS: Well, I'll discuss more if you can at least let me defer to Strategic Goal No. 5.

MEMBER HALNON: Okay, yes.

MR. DENNIS: And we'll talk a little bit more about it --

MEMBER HALNON: Feel free to hold us off.

Dennis, I cut you off. Go ahead.

MR. BLEY: That's all right. I had wanted to ask something completely different.

Your four levels are kind of a structural

look at the use of AI. And I admit I'm more than a decade out of touch with the research in AI, but I suspect some of the things I had seen quite a few years ago are still there.

And what I don't see in a quick look through your slides, although the backup slides might get to this, is a look from the research and science side of what's going on. Machine-learning tools, rule-based systems, all those things that were around, at least in their infancy, two decades ago, show a lot of promise when you look at the positive side, both on safety and on control and operations. And it's easy to demonstrate that.

What I remember, and what I'd really love to hear before the day is out, is how a regulator can look at these systems and make decisions about them. And the reason I put it that way is, on most of them, unlike other kinds of software, there's no real way to verify that the system does exactly what you expect it to do today, and for machine-learning systems in the future as well.

I played with rule-based systems some. And even though they looked very simple, and you can show that they do what you want them to do, you get interactions among the rules that lead to very

surprising responses under certain conditions.

And for the systems that adapt themselves, I have real trouble seeing how we make sure we don't get great surprise sometime in the future as to how they're reacting and how we protect against that.

If you're going to talk about that, let me know about when that's going to happen. And if you're not, well, I think we ought to.

MR. HATHAWAY: I'd like my mute button to work. My fingers are apparently not very effective.

We will talk in the next slide just a little bit maybe touching on that subject about research and applicable standards and guidance. So, I think that's about as far as we intended to discuss today. But I think, to answer Dennis' question, wait for the next slide.

MEMBER MARCH-LEUBA: I kind of support Dennis' point that we eventually want to be using informed regulation on these things, and you need to figure out what the risk is that one of these AIs won't respond the way you expect it.

When you're designing a system to predict what is the next word that I'm going to type on a keyboard, it's really good -- because you get billions of hits per day -- that you can train your next word.

And even then, you consider yourself happy if you get it 99 percent of the time. And if you start moving to safety, 99 percent doesn't do it. You need four or five finites.

So, yes, of course, you will do that. I'm glad the industry is sustaining the organization and at least synthesize where they can produce real money, and the safety, we already have it done.

Let me ask -- you can continue. I see Charlie with his finger on it.

Let me finish with a short question. These four levels, are those the standard in the industry or did you make them up? If I tell you a level 3, does everybody understand what level 3 is?

MR. HATHAWAY: Okay. So, this framework here very much mimics the existing framework in NUREG-0700 for levels of automation and autonomy. So, this is already consistent with what other parts of the NRC are familiar with in human factors for autonomy. So, in fact, there is a level zero; basically, manual operations, so one above level -- so, level zero, 1, 2, 3, 4. The numbering is not common explanation or common knowledge, but more along the lines of the definition.

MEMBER MARCH-LEUBA: Yes. So, Google

wouldn't understand what a level 3 means, but anybody in the NRC and nuclear industry would?

MR. HATHAWAY: It would map to the autonomy table in NUREG-0700.

MEMBER MARCH-LEUBA: Just nuclear, yes.

MS. LALAIN: I'd like to add onto that question. This is Teri Lalain. Thank you.

As we were working through the development and having a lot of our public meetings and our engagements, what we were finding is you might have been talking about something that would be a level 4 and that I'm talking level 1, and we wouldn't realize until about five minutes into the conversation we weren't on the same page.

So, this came out of that, of needing to have at least a starting point; that, when we communicate, we're also communicating not just that it's AI, but what level is human in the loop. Because that's going to shape the conversations we're going to have going forward on what those applications will be; how we would regulate. So, we do envision a lot of this will evolve as we go forward, but we needed a starting point for common understandings.

Over.

MEMBER MARCH-LEUBA: I strongly support

what you just said. I see why you need it and it works. But let's keep it consistent and stick to this one. Okay? Don't change it tomorrow.

CHAIR BIER: Are there other questions right now? I think somebody has a hand up on Teams.

MR. HECHT: Oh, yes, this is Myron.

CHAIR BIER: Great. Thanks.

MR. HECHT: The question that I have is I think this is -- I agree with Jose and others -- this is a very useful framework to start thinking about autonomy and levels of autonomy.

But one thing I didn't see in the plan was -- and Vicki alluded to this earlier -- the fact that, as you increase in levels of autonomy, there are an increasing number of acceptance criteria that you need to consider.

If you're at level 1, if the machine-learning system comes out with the wrong results, it's a nuisance. But if you're at level 3 or level 4, it's not.

So, I would have assumed that in Goal One you would have also included a task to start thinking about what the acceptance criteria were for each of these levels and I didn't see that. Is there a plan to or is that going to be part of your plan going

forward?

MEMBER HALNON: We had a --

CHAIR BIER: Yes, but I think they need to answer the question that was already asked first. So, thank you.

MR. HATHAWAY: Yes, I don't know if it was spelled out explicitly, but I do think that was sort of the point of Goal One, to sort of look at our regulatory processes and see how we do acceptance criteria. And as sort of what's been discussed a lot, it's a very challenging problem because there are lots of different ways to do it. So, giving like a hard-and-fast definition of this is acceptable, it might be acceptable for method A and might be more challenging for method B. Different methods are going to have different levels of explainability, but there are certain criteria that can be common for all applications.

So, I think right now, at least preliminarily, we're thinking like, what are these things an application needs to address that are challenges with AI?

MR. HECHT: Well, do you think that you might want to include that explicitly in the plan, rather than have it implicit, for Goal One?

MR. DENNIS: So, I think that's good feedback, and we have actually already changed it. We did change some terminology recently in our revisions that we're currently working through as part of our activities to address the public comments.

And we recognize that this table only lays out just one facet, but there's many facets of criteria that would be applicable. As Teri mentioned, that there's also -- right now, the Strategic Goal No. 1 will have an implementation through Strategic Goal No. 1 in assessing, or developing the regulatory framework to consider AI applications. And so, I think we can go back and look to see, if we were not clear in any of our paragraphs, to mention how we would go about doing that.

MR. HECHT: Well, thanks.

CHAIR BIER: And any additional public comments or questions right now?

MEMBER KIRCHNER: Vicki, yes. This is Walt.

CHAIR BIER: Yes?

MEMBER KIRCHNER: Just echoing some of my colleagues' comments -- Myron and Dennis and Jose -- as you go through this table and you go down, obviously, algorithms here are somewhat different than

algorithms that we use in making more deterministic calculations and predictions about system performance and such.

But certainly, in that, you know, in the deterministic area of codes and methodologies, there are very well-established techniques for V&V and for uncertainty evaluation, and so on.

I often think about how many embarrassing situations I've put myself in where I'll allow my iPhone to complete a text message, and I don't review it. And the next thing I see is it's chosen a word that might be unfortunate, to say the least, and I hit Send.

But you can see algorithms, as you go through this table, in terms of complexity, and especially when they interface with the safety systems and control, I would expect that RES would have its own program to look at how are we going to V&V these and evaluate in this case not the uncertainty, but the potential adverse impact of these algorithms. And it seems to me that's a research topic in and of itself.

And then, it also brings up the issue of quality assurance. And as we use these kind of algorithms, are they going to meet the Appendix B-like criteria that we apply to other methodologies?

So, I just put that on the table. I think that's actually a very large topic, and I would support the NRC putting its own research team to work on those topics.

Thank you.

MEMBER HALNON: Hey, Walt, this is Greg.

In my mind, you were talking to the level 2 leap to level 3. That's going to be a real big leap, and then, the regulatory regime and how will we permit a system to go from collaboration to operations is going to be crucial. And that could be a real logjam in the regulatory world requiring additional data. Because you never 100 percent know if you've got all the situations taken care of.

I'll give you just a quick story. My grandson is diabetic. So, he had an AI system, an artificial pancreas. And it took two weeks at home, apparently, to learn his system, but he almost died in that two weeks because it was making errors in learning. And making errors in learning in that life safety is really important. So, we had to cut it off and stop it.

So, that leap from the level 2 to level 3 is going to be a really tough leap when you guys are asked to approve a system to go into operation

autonomously, or at least even to make a control decision, and even if it's supervised, where you may not feel like you have 100 percent of all the situations taken care of from a learning perspective.

So, that's another reason why we need to probably have these tabletops to discuss how much do I need; what do I need; what V&V do I need; what quality assurance do I need, along with what Walt was saying: how do you get from level 2 to level 3 without making it very unproductive, right, from a system perspective?

CHAIR BIER: I see there's a hand raised online from Josh Kaizer.

MR. KAIZER: Yes. Sorry, this is Josh Kaizer, NRC staff.

I just wanted to jump in about the VVUQ activities. Very much agree. The NRC is actually actively involved in multiple VVUQ committees that are looking at how do you do VVUQ with machine learning. One is through ASME. Another is an international organization. OECD NEA has an activity and we're, basically, providing data for them to go play with it; try to find all those issues that you're talking about.

So, I did want to come in and say that we

are very much aware of the potential differences in VVUQ that you would need for AIML versus conventional models. One of the things we are also trying to do is find out where there is significant overlap. My favorite area, critical heat flux, looks like a data-driven model to me. And they even have data-driven models for critical heat flux. So, that seems like one of the low-hanging fruits where we can play there.

I've even developed neural networks that do about as good of a job of predicting CHF as these highly empirical correlations. So, where there is that significant overlap, we are taking advantage of it. We're not trying to add additional regulatory burden when it's not needed, but the use of different structures of AI and ML, especially neural networks, do add some potential new things that we haven't considered before, and that's what is being studied in those VVUQ groups.

So, I did want to provide that to kind of say that we are thinking about it. We certainly don't have an answer, but we are moving in that direction.

Thank you.

MEMBER MARCH-LEUBA: Hey, Josh, this is Jose. I'd like to emphasize what you just said, that AI has applications for the low-hanging fruit which

are very easy; they're cheap, and they're implementable, and they can be regulated. There are areas where it would cost so much money to even think about it, that it's not worth it. And I don't think the industry is going to spend that money, anyway.

But if you think about it, we know how to do reactor protection systems, and they're fully autonomous and they're machine-driven. So, you can say that we have implemented level 4 AI systems on the RPS. And RPS is a really good level 4 AI system. It was brain-driven when it was designed. There is no emphasis, no need to reinvent the wheel on reactor protection systems.

On a CHF correlation, if you can get a little better statistics on your prediction with an AI model, and you have sufficient data to do it, it's a low-hanging fruit. That's what we need to concentrate on.

MEMBER BROWN: I'm finally going to speak up because I'm probably the worst of the resident skeptics on here. There's less acceptance of this. We couldn't even get Teams to work right all the time with the computer running around in the background making decisions for us as to what we can see or not see.

With reference to the reactor protection systems, those are, the new ones, computer-driven. And there, you're dependent upon the computer starting with a cycle and it completes it, and it evaluates all the data it goes through. What if it stops? How do you evaluate that? How do you protect yourself in an AI system where you're depending on the continuous input of information in order to learn what's going on?

The other question I was going to ask earlier, and I'm not sure it's been addressed, is, for the fully autonomous, even the level 3, you're looking at systems that replicate human thought. The first two, you can question whether those are really that replicant or not.

But humans are fuzzy logic people, and you input thousands of pieces of information all the time. Like when you're driving a car, if the lines on the road disappear, if there's a fence on the side, you can make sure you stay in your lane. The driver-less cars you have today can't do that without a thousand detectors, and if the lines disappear, you're going to go somewhere you don't want to go.

How do you input that type of thought process into a machinery, you know, a complex plant

like a nuclear power plant, and how it reacts and how it starts or stops things, or what have you?

How do we relate the way humans really think with abstract information, which they discard? We discard it all the time when we're making decisions. And the machines have to have something that tells them what to discard or not discard. How do they figure the way we do that we've learned as we've grown?

I think we've had a lot of comments now that there's going to be a real leap going from any 1- and 2-type levels to the partially autonomous. If you want to ride in a driver-less car, go ahead; I won't be on the road with you. Just let me know when you're doing that.

So, it's not that we shouldn't go look at stuff. We might find, we'll probably find some good things we could use. But I think the Committee, as you have heard, is going to be kind of skeptical in terms of when do we replace the true reactor protection systems, safeguard systems.

What if this nifty human-learning system you've put in place says, oh, all of a sudden, I've lost my external grid and I'm going to start the diesel generators in parallel? But yet, you didn't

really lose it and it tries the parallel and blows up the diesels -- not a good thing to have happen. Whereas, a human would know what's going on just from all the sad stuff that they see.

So, it's just a very difficult problem for you all. We're just making it hard on you in reality.

So, I'll be quiet, and I just thought I'd add my skeptic voice to it.

CHAIR BIER: Yes, I think we should probably go ahead and get further into the presentation, and we may have more comments later.

MR. DENNIS: We did have a colleague, Brian Green, who had his hand up. And he was actually going to address some question earlier. So, I wanted to give him an opportunity, if he's still available.

MR. GREEN: Hi. Thank you, Matt. This is Brian Green, the Human Factors Team Lead.

I just wanted to go back, I forget who the speaker was that was asking about acceptance criteria; you know, do we have those as we move up through the levels? And the answer is, at least within the human factors world, not yet, but we have started those discussions with the Office of Research and some of our research partners, to start to think about what it would take to get us there.

To some extent, we think that some of the automation-related criteria in 0700 may be reasonable, but there is a very high probability we would need additional criteria to support the more sorts of advanced sorts of machine thinking sorts of capabilities.

So, don't have the perfect answer yet, but that is something that's in progress at this time.

MR. DENNIS: All right. Thank you very much. We'll move on to slide No. 8. There we go. Okay.

And again, this is Matt Dennis and I'll be taking over the remainder of the slides here, as we go through discussing the goals.

So, this is the first one. I mentioned, I said I promised I would at least discuss the AI research component of this, because that is very important in Goal No. 1. So, you'll see that on the lefthand box, as I'll talk about that more in a second.

So, keeping the end in mind, Strategic Goal No. 1 focuses on determining what is the depth of regulatory review that will be required to somewhat address that previous slide with the levels of AI automation.

So, one of the first things that we need to do is conduct AI research. So, AI research specific to nuclear activities is needed to develop a regulatory framework to review the use of AI in NRC-regulated activities. In part, this can be done by evaluating the approaches that assess AI characteristics, and I believe that was somewhat brought up earlier. And in the Strategic Plan, we call out some of those things, such as explainability, trustworthiness, reliability. And those all are different characteristics of AI that need to be assessed in order to make a finding.

And assess the existence -- and we need to identify regulatory and research gaps that need to be conducted -- and assess the existence of standards and guidance for identifying where work needs to be done.

Second is developing the framework and tools that are either a product of that research or support that research. And so, we need to clarify what the NRC needs to regulate in those levels; how we need to regulate it, and what the criteria will be -- which we just were speaking about the criteria -- and develop options for future frameworks that may require long-range changes for AI regulatory review and oversight that may require either rulemaking or new

guidance.

And then, lastly, communication is an absolute priority for us. We need to harmonize our own internal AI activities to ensure that we have a cohesive message, and that activities we undertake, we have direct engagement with external stakeholders frequently, which we've tried to do up to this point and plan to continue to do, to advance the development of appropriate regulatory positions.

Some early efforts in the Strategic Goal No. 1 and implementing the Strategic Plan will be to conduct research to perform a regulatory gap analysis, and as Josh mentioned earlier in response to one of the questions, the staff are already very much involved in AI standards in other organizations, such as NIST, ASME, or IAC. And we will be engaging our counterparts in NRR, as you just heard from Brian, about research that needs to be done to expand the research around artificial intelligence.

And as was mentioned earlier, we've already undertaken several activities to be able to communicate with the stakeholders, such as the data science and AI workshops that were conducted in 2021, and then, RIC, the Regulatory Information Conference. In 2021 and 2022, there was also a technical session

on AI, and there will be one planned for 2023.

MEMBER HALNON: Matt, this is Greg.

Do you have a conceptual idea of the timing for each one of these goals? I know the RIC is 2023. Inevitably, it goes on to 2024 and 2025. You know, it goes on for a long time. But any idea of when you might (audio interference) specifics?

MR. DENNIS: At this time, I can't say, but I can say that is -- and that will be on one of the next slides -- setting up the structure internally for the agency. The Steering Committee, the AI Steering Committee, is going to be one of the first things. We've already identified the participants in that and begun to implement that activity, as well as, as I just mentioned, doing a regulatory gap analysis.

So, we have identified some work that we are already going to kick off this fiscal year in order to get the ball rolling on implementing our goals. But, at this point, we don't have an implementation plan set up with a timeline.

MEMBER HALNON: Is the intent to keep this as a living document or to develop a new Strategic Plan? Or is it you're going to just keep this one updated and public?

MR. DENNIS: The intent is for this one to

be a living document and the implementation plans for each of the goals to be also a living document.

MEMBER HALNON: Okay.

MR. DENNIS: Also, we did develop this very quickly in an agile fashion.

MEMBER HALNON: Okay.

MR. DENNIS: So, we were keeping pace with the rapidly-evolving technology space as best we can.

MEMBER HALNON: Do you have any idea when the next revision might come out? Is it a year after the Steering Committee is put in place, or something to that effect?

MR. DENNIS: I don't know.

MEMBER HALNON: Okay.

MR. DENNIS: I think, right now, we're going to -- and Luis just raised his hand. So, he wants to chime in as well on this. But this will be something that will have to be predicated on looking at --

MEMBER HALNON: Okay. And I know it's early.

MR. BETANCOURT: Well, good morning. My name is Luis Betancourt. To your question about the timeline and if this is going to be a living document or not, the strategy-based itself is going to be

updated every four years, and the strategy will try to align with the Strategic Plan. The implementation plans that Matt is talking about, those are going to be more living documents.

And to your question on the timeliness of the goals, that's one of the things that we're going to be trying to do in fiscal year 2023, to provide the transparency of timelines of when and where industry as well as the staff should be working on not only good research, but the guidance development. So, we're working on that right now.

But, like Matt mentioned, one of the goals, which is Goal No. 2, that will be done in fiscal year 2023 because that, basically, establishes the framework, so everybody can work together. But the rest of them are going to be trying to continuously. But we need to provide that transparency to both stakeholders in the industry and the staff of: what are the timelines and what are we trying to accomplish? That will be done in this year.

MEMBER HALNON: Thank you.

MEMBER BROWN: This is Charlie Brown again. You talked about you're establishing a committee, Steering Committee, whatever you want to call it. And they'll sit around and debate and try to

figure out what they're going to do.

And this is based on experience only, my next comment, which will not be well-received probably. Anytime you have a committee that's doing that, you can go on and on and on, because you'll never get total agreement as to how you proceed. It normally takes, if you want to make progress, you need somebody that has the vision; has an idea of how to do it, and has a 1-to-nothing vote on the next step.

If you look at why we have nuclear power plants today, Admiral Rickover went to Congress in 1948 or so. They trusted him to build a non-air-breathing nuclear power plant. Nobody even knew how it would work. They didn't know how the materials would behave. And in six years, he produced a prototype out in Idaho and a nuclear submarine that went to sea in 1954. He had a 1-to-nothing vote on every decision process it went through.

As you understand the way the system works, okay, at some point, if you want to make progress in this area, you really have to have a driver and not have a constant rehashing/reconvening. Somebody has to say, "We're going to stop here and go there." And committees don't do that very well.

It's just a suggestion on future thinking

on how you want to organize within the agency, on how you want to go forward. It doesn't work. I had that personal experience myself in the mid-'80s when my four-star asked me to manage a program I didn't want to manage, because I thought it was too risky and it would cost too much, and it had a high probability of failure. And he said, "Well, we have no choice and I'm asking you to do it," since I worked for him.

But he agreed to I would not have to report to any committee; I would not have to ask for anybody's permission on how I did the technical direction, and the money would be coming as I needed to go out with the industry and build this machine that did not exist in the world at the time.

And we made it and we did it, and we had to take a strategic asset out of service in order to do that, in order to prove it at sea.

It's that type of thought process. I couldn't have done that without my four-star. Okay? He was right there behind me the whole time. So, he made it work by allowing the program to do with it single-point guidance.

But I'm just saying you might want to think about that. This is a critical area. People are going to be looking at it over the next 5 or 10

years. And you really need a single point of direction, not a group, a group grope on it. That's just an observation for you.

Thank you.

MR. BLEY: This is Dennis Bley. Charlie just got me interested in a third point. I want to raise just three points with you.

The first one is, a few years ago, the staff brought to us a vision and strategy document for non-LWR reactor licensing with a set of implementation action plans. And as was said just a few minutes ago, they didn't keep that fully up-to-date, but they kept the implementation action plans up-to-date. And those gave us a place to review and the staff a way to organize and proceed through that rather complex set of demands quite well.

And it looks like you might be setting up that way, but we'd be very interested in seeing the structure of those implementation action plans for each of the goals as soon as that's possible, because that was a real key.

The second is a couple of the people on the Committee commented on things that I just wanted to make a note about.

Jose mentioned that some of the systems we

have are like a level 2 AI system. And they're not AI systems, though. The AI systems that the staff's looking at are things like neural nets and other mission-running algorithms.

And the kind of things that we have out in the field may do what we envision these AI systems doing, but they really aren't. They are things you can go in and V&V in the traditional sense, and if you pass through with a set of data, you can show exactly what's going to happen.

These other systems -- let's take neural nets. You can test them, but you can't V&V them. And they change their connections over time.

And I really think it's a major difference that becomes very hard to ensure, from a regulatory sense, that you won't have unexpected results coming in.

The current slide we're on hasn't gotten to me to the point of understanding even how you're looking at that as yet.

So, those were the things I wanted to offer up and hear what you have to say about them.

MEMBER MARCH-LEUBA: Can I do a followup on Dennis before you go to the up hand?

I'm going to say something very important.

Self-erupting AI systems after you are implementing the plan, is there anything, any thought in the plan about not allowing self-erupting things?

I mean, you review an AI system that has learned based on however much data you have, and that's the one you regulate. Now, if you put one of these self-erupting that's changing, how do we keep track of that? Have you given any thought to that?

MR. ANZALONE: So, actually, this is Reed Anzalone from the staff.

That was the thing that I wanted to try to address, Jose, if I can.

MEMBER MARCH-LEUBA: Please go ahead.

MR. ANZALONE: So, the thing that I wanted to say is that I think we recognize that there's definitely a distinction between adaptive neural networks or things that generate their own rules internally on the fly. And that's something that we've been talking about and exploring with our counterparts in other regulatory agencies, both domestically and internationally, and definitely recognize that there is a distinction there and that those kinds of systems need to be treated differently.

I think we want the Strategic Plan to encompass all of those kinds of technologies,

including both the ones that do adapt themselves and the ones that don't. I don't think we've given any specific thought -- and my colleagues can, of course, correct me -- as to whether we should not allow those, but we definitely recognize that dealing with those kinds of systems is much, much, much more difficult than complicated. So, I just want us to say that we recognize there's a distinction there.

MEMBER KIRCHNER: This is Walt Kirchner. So, just joining with Dennis and Jose, going back to your previous slide -- don't go back -- but there's some keywords in there that, for me, make a distinction. And Jose has hit on one of them. For intelligence, it's adaptive. That's one aspect of intelligence in terms of a machine, I believe, and that's a critical thing.

So, somewhere you're going to have to -- you don't want to freeze progress by defining things, which in a field that's obviously in a state of flux, but there is that subtlety. And there are some keywords buried in level 2 and 3 and 4 that the definition of I think is very important.

Because Jose gave an example of a reactor protection system, but that's not an intelligent system in the sense that the word "intelligent" has

that adaptive kind of characteristic implied for the system.

So, it seems to me that you're going to have to lock into that level 1 through 4 in some way, so that the groundrules aren't in flux. But, then, you're going to have to define what you really mean.

And if a system, an algorithm is learning, as Dennis gave an example, through a neural net kind of testing process, that poses some serious challenges in terms of V&V when it's applied to a safety system or function.

And so, I'm sure you recognize the challenges that are out there, but, as part of this framework, you're going to, I suspect, have to have some common understanding of what you mean by "intelligent." There were a few keywords in your definitions, the different levels, that I think would benefit from a little more attention and definition.

MEMBER PETTI: Well, this is Dave. I'm just wondering if this whole adaptive issue implies that there's another dimension to the four categories. Because it could be adaptive in any of those four categories. And so, are the definitions sufficient? It's necessary to have them, but are they sufficient in terms of capturing all the key important variables?

MEMBER MARCH-LEUBA: Yes, before you answer that, let me emphasize that, when I have human operators in the control room, I train them real well. I'm training four procedures, and in the four procedures I hit them with a hammer. But I want them to learn from their mistakes.

I'd rather have an operator who has been running my plant for 20 years that has learned that, when that green light goes yellow and the other one -- it's not in the procedures, but they know what is going to happen next. So, adaptive is not bad. It's just difficult to regulate, and I don't envy your position.

On the other side, I do envy your position. And I wanted to tell you -- I'm digressing here -- but this is a batch of armor. There are so many questions and so many comments in an ACRS meeting. This doesn't happen often. And it's because you guys brought to us a very interesting topic, very lively, with real application. And so, you should consider yourselves happy, not happy, but lucky.

Happy, too. So, especially lucky that you are working on this project. And this is a career-defining project, and you should put your mind into it because it's important. Thank you.

CHAIR BIER: Anyway, I think we do have to move on to get through the rest of the presentation.

MR. DENNIS: All right. So now, we're moving on to -- this is, again, Matt Dennis -- and we're moving on to slide 9.

And so, this may answer a few of the comments/questions that were just about the structure that we're going to put in place with Goal No. 2.

So, Goal No. 2, the objective is to ensure cross-agency leadership in AI with a centralized approach. And I fully recognize that there are issues that come along with implementing anything like this with a lot of varied opinions. But what we've found so far is, as we mentioned, there's a lot of activities in AI that are going on at the agency. And so, this is our best attempt at this point to wrangle all of that together in a structured format, and as best we can.

MEMBER BROWN: Cross-agency would certainly fall into the category of never ever getting anything done.

And I hate to say that, but that's as soon as you bring in DOE and five or six other agencies, it will be --

MR. DENNIS: Just to clarify, this is only

for our internal -- for NRC.

MEMBER BROWN: Well, you said, "cross-agency."

MR. DENNIS: I'm sorry.

MEMBER BROWN: Does that mean within the agency?

MR. DENNIS: Intra-agency.

MEMBER BROWN: Intra-agency? Okay.

MR. DENNIS: I think Teri wanted to chime in here, too.

MS. LALAIN: This is Teri Lalain. So, for the Steering Committee, the core members in it are across the NRC offices. Because, as we're talking about today, there's a lot of potential areas where AI may be of interest. And we need to align, prioritize, and organize around what we think are going to be the first and most likely.

We also recognize from the conversation we've had here, and conversations prior, that there are other agencies that are working similar topics. So, having the ability to pull in an external member, if we want to talk about a specific type of risk, a specific type of analysis, learn their lessons learned, and gain from those insights, we see value in having external participants, where it is the right

information at the right time to help us continue moving forward on our journey as well.

Thank you.

MR. BLEY: This is Dennis Bley. One thing I'm very surprised wasn't on your first slide, and it's not on any of them -- and some of our discussions have brought this up -- what the heck do you, the NRC staff, really mean by AI? Are you talking about the things that have been traditionally developed and researched under artificial intelligence? Or are you including things like Jose brought up with reactor protection systems that is a deterministic piece of software, but that does some of the things one might want from things you list in the levels of AI on your slide?

MR. DENNIS: I'll go back and answer that question from the Strategic Plan itself. In preparing the Strategic Plan, we tried to craft a definition that was encompassing as best we could. And so, I won't read it out loud here, but the AI definition that's in the Strategic Plan is in the Executive Summary.

But one of the points that we wanted to get across at the end of the definition is that "The AI Strategic Plan focuses on a broad spectrum of sub-

specialties...which could encompass various algorithms and application examples which the NRC has not previously reviewed and evaluated."

MEMBER MARCH-LEUBA: I'm not going to say anything.

(Laughter.)

MR. DENNIS: Well, there's more definition about that, too.

CHAIR BIER: One comment that I would add with regard to AI is I think the agency may want to think about definitions that are functional rather than like semantic. Because what today is considered AI may, 10 or 15 years from now, be called computer programming. But if it's adaptive, for instance, then the same issues would apply 15 years from now, where it's no longer called AI because your kid's toy is doing it already, but it still needs to be regulated in kind of the same way.

MR. BLEY: I think what you're really talking about is automation, not -- some specific part of automation like artificial intelligence, but I'm not completely sure, even with your definition.

MR. GREEN: Hi. This is Brian Green, Human Factors Team Lead.

We recognize that there are some

challenges with the definitions. And some of our research partners are proposing different ways of defining this. I don't know that any of it is published yet. So, I'm reluctant to get into it too much.

But I think, Dennis, you're making a good point, and I think it's similar to one that Jose has made, that there is automation that does not use AI that can do things that are adaptive.

There was a long-term research project initiated in 2012 that focused on what we call adaptive automation, which is systems that can change the level of automation based on various parameters, but it didn't use any sort of machine learning to support it.

So, there are multiple dimensions that we're all trying to sort out at the same time, and that is a part of what's going on, is that definition process.

MEMBER MARCH-LEUBA: Sitting here -- this is Jose again -- I am starting to get a little confused or a bit concerned. Are we confusing AI with machine learning?

And if you go downstairs to the street and pick five people and make a focus group and start

talking about it, those five humans will tell you that an AI system is machine learning. And it doesn't need to be. It doesn't need to be.

I can make an argument that the reactor protective system is a beautiful AI system. It's fully automatic and makes good decisions because the human brain was involved designing it, instead of just machine learning.

So, there are applications where the machine learning works great. For example, you can put all the parameters of a pump or a thousand pumps and learn when bedding is starting to make noise. And then, you can use machine learning to diagnose the pump. There are other applications where really it's not worth it.

So, the difference between machine learning and AI is not the same thing.

MR. DENNIS: On the subject of definitions for that, to answer that a little bit, in one of the slides we'll show there's overlapping concentric circles. I'll get there in a minute, I promise. But, basically, machine learning, in our definition, is a subset, with artificial intelligence at the very top.

In the Strategic Plan, we have a little pictorial that shows that. We didn't include that in

the slide deck here. But AI is at the top of everything, and machine learning, natural language processing, and other things are under that umbrella.

MEMBER PETTI: I want to take you to a different one, because I'm worrying about the time. We may not get there.

The Strategic Plan struck me as a very high-level, sort of transactional document. These are some of the things we're going to do; these are our goals.

Do you plan to develop an implementation plan that really gets to some nuts and bolts about -- you know, it's a huge challenge. How are you going to regulate?

I mean, I just worry that here we are at 100,000 feet, but the real issues are at like 5,000 feet. And how are you going to get there?

Having a Strategic Plan is good, but I think those implementation plans are really where the rubber meets the road. And what's your timing on that?

MR. DENNIS: Well, the short answer is, yes, we will have an implementation plan. Luis mentioned that a minute ago, a little bit ago. And the timeline on that is we're working on it right now

and developing the implementation plan.

The first thing was to finalize, because, again, this is still a draft technically. We have to respond to the public comments and finalize the Strategic Plan.

Luis, do you want to add something?

MEMBER BROWN: When we first built the first nuclear power plants, we built the plants, and then, we figured out how to regulate. We didn't figure out how to regulate building the Nautilus. We didn't figure out how to build the Duquesne Power and Light Shippingport Plant. We built it, and then, figured out, oop, we've got to have all these things, as we found out as we went through the building.

I would think the way you want to really figure out what you want to do is find some application where the process would provide some positive benefits and go decide to do it. Get somebody to build it, and then, figure out, oh-oh, these are the things we need to do to make sure it comes out right.

You can't regulate without knowing what you're regulating. I think that's a bad way to go. That's the wrong word. That's not an efficient way, or even a useful way, to go about doing something.

You can't regulate what you don't know you're regulating.

And I'm sorry I've interrupted you.

MR. BETANCOURT: No, that's okay. This is Luis Betancourt. That's helpful, to answer your question. But, going back to your question, I think the idea of the actual curriculum was to get the discussion on AI going. We understand that it is very high-level, but we are now in the chicken-and-egg situation that, if we don't put something out there, how can we have that conversation?

The implementation plan, that's where we're going to be having the nuts and bolts, and that's where we need to have the transparency. But we need to put something out there, so industry can also tell us, okay, this is what we are planning to do.

To your point, Charlie, the tabletops, like the workshops, like that's where --

MEMBER BROWN: But you have to get within what would be useful in your applications.

MR. BETANCOURT: All right.

MEMBER BROWN: What would be useful? I don't know about the tabletops. I just built stuff; that's all I ever do, okay, and then, break it, make sure it works.

Find out something that industry can use that can improve the operation of the plant. Go do it, and then, figure out how you make sure it comes out right and works right. If you don't have something -- I'm hard over on that. I just don't like to see all the plans, strategic, blah-blah; we're going to do this and implementation on this. We'll be doing that for 10 years, and you'll have nothing to show for it except about a thousand pages of paper.

I'm being a little sarcastic, but, I mean --

MR. BETANCOURT: We know, Charlie, and again, like we understand that we need to work on the guidance, but we need to put something out. This is the first of many discussions. To your point, like we don't want to have this like a 10-year project, because, then, we are stifling innovation, if you really think about it. We're stifling like the industry trying to use AI. So, we understand that we need to do more, but this is the first step for us to get there.

So, I'll stop here because I know that we are running short on time.

MEMBER REMPE: Well, I would have blamed industry and said, "We've got to wait for them to

submit something to go further." I mean, you guys have kind of inferred that, but I'd say, "Hey, until they come in," you've got a good step and a path forward, but I would say, you know, we don't want to go ahead of them and do something they don't want. I'd blame it on them.

But, anyway, I know we're way over schedule.

MR. GREEN: This is Brian Green, the Human Factors Team Lead for NRR. I do want to say that there are designers who are coming to us, current designers. And again, this is proprietary, and I don't think we need to go into the details here. But they are starting to ask questions about certain applications of AI. Now, I've not seen anything that would go near a control room or anything that would be safety-related, but they are starting to push out into doing some functions at the plant. And there is hesitancy to do so because they're nervous that we're going to regulate it away.

So, industry is doing work in this area. They are progressing, and I think it's fair to say they're treading cautiously, waiting to see what our next move is.

So, we are not doing this in a vacuum. We

are getting some feedback about potential applications and where that's going. So, again, those aren't up and running yet. So, to Charlie's point, we're not quite there to see it and to have the great data on it, or at least not in my experience. Perhaps others have, but --

MEMBER HALNON: Yes, this is Greg. And that's exactly the point I made earlier about looking at each other; who's going to make the first move? And the business cases for these folks to put significant resources in developing a system -- testing, resources, all those dollars, all the extra stuff -- you're not going to get off dead center unless the regulatory regime is defined well enough, that they know that it's not going to, like you say, designed and regulated away.

So, there's got to be some specificity to some extent. And that table, level 1 through 4, is a good one to revolve around. I'm sure you'll come up with other diagrams, or whatnot, to revolve around.

Until there can be some certainty in the regulatory aspect, it's just like the nuclear plants we're looking at now, because of the work we're doing on the new Part 53, people aren't sure they're going to use it or not. So, until it's very specific, the

designers don't know what they're going to be designing to.

So, that's really important, and I think that's something that the industry and the NRC need to come together and figure out what to do. Because, otherwise, you're going to be staring at each other for years and it will be the 2009 renaissance of nuclear plants, where we thought we were going to get 60 new ones and we got two, because of the uncertainty in the regulatory regime.

MEMBER MARCH-LEUBA: The thing is, seeing past history, no matter what the plan says, if I am the first applicant that comes into this building with a plan, with an application, I will be scared to death because the NRC is going to regulate me. I'm going to get so many RAIs.

So, I was thinking right here -- and it's not my job to do it, and I don't know how to implement it -- but this is a perfect -- perfect -- example of where the DOE should fund the first prototype. And I don't know how to get that done, but I'm just putting it in there. I mean, they're funding all kinds of other stuff. This is something that Gorman should foot the money for, for the first implementation, not via the risk of my plant because I'm not one to do it.

I don't have the money.

MR. GREEN: I think there's some building blocks we can do with the non-safety and the business systems, and what you guys are talking about, learning and getting it in place, so you can have an intelligent conversation about it.

All that's important. Will it get us from that level to a level 3 in a safety system someday? Maybe, but, again, to get there, we have to have some sort of certainty and specifics of the regulatory regime.

CHAIR BIER: Anyway, I think we do need to move on.

MR. DENNIS: I will try to get through all the remaining slides in three minutes to the best of my ability.

But Brian did mention that a keyword that we want to communicate here is that the NRC is not working in a vacuum. That's one of the key messages we want to get across. We're maintaining awareness with these examples here. We have participated in communication forums and we do have working partnerships with three -- the DOE, EPRI; DOE and EPRI MOUs. We just started off a U.S., Canada, and UK trilateral engagement on high-level principles for

evaluating AI in nuclear applications. And then, we're also participating in IAEA consultancy meetings, which I think you'll hear a little bit more about later.

So, I guess the message I want to get across here is that we are strengthening and expanding our partnerships already, and we are serving and participating in lots of things to get awareness of what is coming down the pike as far as applications in AI, or AI that are being used or envisioned in the nuclear industry.

Next slide. We're on slide 11. I mentioned earlier that we have AI at the top. So, on the right you can see the concentric circles that, basically, give you a breakdown of some of the characteristics or key attributes of people who are data scientists, data analysts, program analysts. So, you'll see machine learning, artificial intelligence, natural language processing, deep learning, all under that data science level with data at the center of everything.

So, we've mentioned several times that the NRC, in order to maintain its readiness to evaluate AI applications, we need to focus on developing the critical skills, which include workforce planning,

staffing, and training the staff in these areas.

One thing I want to get across here is that we're not planning on targeting or training the entire workforce, but those who are interested in up-skilling, or through qualification plans, training themselves, or getting conversant in AI technology in order to be prepared for the future.

Right now, we have quite a small team of -- we have Trey, myself, the people on the phone. We have a handful of people, through the Office of Research and across the agency, that are conversant in this area, but as far as technical expertise for doing anything in the future, we need to be prepared.

So, moving to the last goal, Goal No. 4 -- we're on slide 12 -- and that is, one of the things in this, and I'll go to the one, two, three bullet, the test cases, proofs of concept, and pilot studies. I think this has been mentioned several times.

But one of the aspects of Goal No. 5 is to pursue those use cases, which includes implementing or doing a proof of concept to test out the regulatory framework that we envision. So, we need to learn, measure, and evaluate our readiness to implement that regulatory framework. And that does require industry engagement or DOE participation in areas where we

think that a future focused initiative may come in.

And so, part of our preparation is also having our Future Focused Research Program and the University Grant Program, which are crucial in developing and maintaining our AI capabilities.

So, we continue to invest in AI research that is already being conducted that's not necessarily part of the Strategic Plan, but will be supplementing that.

And then, also, we need to have our own AI ecosystem, so that staff are prepared with the data that we have internally, as well as data that may be coming as part of an AI application, and the tools. This is just an example of some of the tools that we need to have to learn, understand, test, and evaluate AI applications.

And so, the next slide, slide 13.

That brings us to my last one. I apologize for wrapping it up quickly, but the goals here with the status and next steps is I already mentioned that we are currently in the process of doing through and finalizing the AI Strategic Plan. Our goal is to get that out by spring 2023; go through the comment resolution, and publish the finalized Strategic Plan.

We'll be hosting a technical session at RIC 2023 on artificial intelligence, and then, we'll also be, again, resuming those data science and AI public workshops in the spring or summer of 2023.

So, with that, I think that wraps it up.

CHAIR BIER: Okay. One minor comment on slide 11, where you said data is at the center of everything. Not all artificial intelligence approaches are data-heavy. As Jose and Dennis I think have both mentioned, if you look far enough back, you have rule-based AI systems which are not popular right now, but which your data consists of sitting around and asking an expert, "What would you do in this situation?" And that's still AI, even if it's not data-intensive. So, just a clarification.

We are now, essentially, at 10 o'clock or a little after 10 o'clock. Are there other must-ask questions right now? Or do we want to move on to the EPRI presentation? Or do people need a short break? Where are we at?

MEMBER KIRCHNER: A break.

CHAIR BIER: Okay. Can we make it a 10-minute break and come back at -- well, let's make it 10:15, so a 12-minute break. Okay? Thank you. See you all soon.

(Whereupon, the above-entitled matter went off the record at 10:03 a.m. and resumed at 10:18 a.m.)

CHAIR BIER: Okay, we are going to go ahead and get started. The last couple members will be trickling in soon. This presentation is from EPRI by Thiago. I don't know if you want to just introduce yourself at the beginning, and then we will go from there. Thank you.

MR. SEUACIUC-OSORIO: Thank you, and good morning, everybody. I am Thiago Seuaciuc-Osorio, I am with EPRI. And I'm here to discuss some EPRI AI projects. And I'm going to do this mainly by example.

At the end here of the set, there's going to be a slide that has a number of different applications. It's just going to be titles, and I can talk very briefly about them. The goal there is really just to provide a scope so you have an idea of the scope of the different applications that EPRI is looking at.

But the main presentation is going to focus on one example, which is AI-assisted inspections. And the reason why this example was chosen, there's a -- there's a few reasons. One, it is the level maturity, it is an area where we had a

successful field trial and we can see that this can be -- being deployed, implemented in the field quite soon in the future here.

It is also an application that is over a regulated activity in the power plant, so there will be regulatory aspects to this here. So we start answering the question of, you know, how do we regulate AI in the power plants when they get applied.

You know, earlier in the day we heard the term of, you know, low-hanging fruit. We're going to see that this is one of the low-hanging fruits. There's a lot of characteristics in this case that make this a little bit easier to answer the question about how do you qualify this for use in the field.

Following this, you know, we're going to see -- I'm going to discuss an effort where what do we do when it's not a low-hanging fruit. You know, how do we start going for those a little bit higher hanging fruits when, you know, those things that make this easier perhaps to regulate and to apply. When those characteristics are not there, how do we go about that.

So I'm going to discuss a effort, a collaborative effort that we trying to tackle that question. So next slide, please.

So the example here is assisted analysis of inspections. In this particular case, ultrasonic inspections. Just as background here, UT inspections that are very important as the scope of an NDE program. Lots of them will use a lot of -- large volumes of data that need to be reviewed during an outage.

And the idea here is that the machine learning tools can assist in the analysis of the data. Now assist here, it's stressed there because the idea here is that the machine is not making a decision. Assist here means that the machine is just providing areas for the inspector to review, okay.

The final decision, this is the key point here, the final decision will still rest with the qualified inspector. Yes?

MEMBER MARCH-LEUBA: So how do you prevent the operator from relying too much on the machine and review only those things that the machine flags and not others?

MR. SEUACIUC-OSORIO: Yeah. Well, that is the idea. We'll get -- we'll get to this, if you allow.

MEMBER MARCH-LEUBA: And what's a UT?

MR. SEUACIUC-OSORIO: Ultrasonic

inspection. I'm sorry, UT means -- I am sorry. You know, after moving to the U.S., I've had this thing of also having to learn, deal with acronyms everywhere, and I think I got a little bit too used to it. So I apologize. All right, next slide, please.

So you know, the first thing that I want to cover this is why, why do we even want to do this. What's the value, what does it bring to the industry, right. That is a question that, you know, if we're going to worry and go through the trouble, you know, what are we going to gain.

So this is just to paint a picture here for you. So today the examiners were going to go and they're going to do inspection. In this specific case here we're talking about the penetration nozzles on the reactor -- on the reactor vessels that. There are, you know, 70-plus of them. So every outage that they need to do that, there is a lot of data they need to go through.

In that illustration there on the left, you have the inspector going through a lot of records. Most of those records actually have no flaws, no indication that need -- need their attention. This represents by the white records. Only a very small fraction of those is actually going to have something

that the inspector needs to look at. Those are the red ones there.

The problem is that he has no prior knowledge of, you know, which ones are the ones that have something, which ones are the ones that don't have something. And he has to go through all of them, and he has to, you know, distribute his time and energy equally across all of them.

Where, you know, if he knew which ones, he would be, you know, he would be much better off in applying most of his energy on the ones that he knows are to be the most problematic ones, or relevant, let's put it like that.

And they do this around the clock. There's a lot of data to go through, and they have to do that. They have to maintain the lab while focused in attention for extended periods of time, which is difficult to do.

The idea on the AI-assisted inspection is that the AI would come in and it would be able to pick out those red records for the inspector to look at. So now instead of having a lot of records to go through, he only has a select few. Those are the ones, if the machine is doing its job correctly, those are the ones that have the indications that need to be

carefully reviewed and disposed of.

So now, he can focus his energy on those. And the likelihood of, you know, fatigue or that momentary lapse of focus that may have -- it have -- it minimizes the chances that it will cause a problem.

MEMBER HALNON: This is Greg. Is there another block here that talks about if he's looking at the more critical indications, potential indications? There's still an interpretation piece of this that -- and that's why we have reviewers and we have level three reviewers because they're more experienced to look at this stuff.

Is there another block that's -- that's even more intelligent or less susceptible to human error, even from the -- looking at the indications and comparing them to code requirements and whether or not they're indications to begin with?

MR. SEUACIUC-OSORIO: Yeah, so this is what the level three is doing here on the right. So that person looking at those records is the level three. Even though those records have been flagged for a review, most of them are still not going to have, you know, relevant --

MEMBER HALNON: Okay, so it's a real wide filter basically.

MR. SEUACIUC-OSORIO: Yes, yes.

MEMBER HALNON: Okay.

MR. SEUACIUC-OSORIO: And we'll see here from, you will see when we get to the field trial, you're going to see everything because we did a careful analysis everything that got flagged. And we had the analyst go through and say why was it flagging. You're going to see that most of the things that get flagged are still not -- not relevant.

But they are things that need that expertise from the level three to say that, oh no, this is not a flaw, this is not relevant.

MEMBER HALNON: So we're not replacing the expertise of the level three.

MR. SEUACIUC-OSORIO: No.

MEMBER HALNON: We're giving it more focus.

MR. SEUACIUC-OSORIO: Yes. Now, part of the idea here is that, or part of the value we're going to see here, these have to be done twice, right. You have two independent reviews of this data set to be done. And the idea is that the AI-assisted inspection may replace one of those.

MEMBER BALLINGER: This is Ron Ballinger. A timely example. How can this -- there's a number of

things that are happening. A lot of this, for nozzles and stuff you're looking for thermal fatigue. But we've had a couple of really pants-down problems with stress corrosion cracking which we didn't see but was still there in the data, still there in the data.

And then not only that, but the methods by the censors that are being used. The technology has evolved with time.

So how does what you're thinking about doing provide alerts, if nothing else, that yeah, the data doesn't show that we've got a fatigue crack that's growing, but there's something else here that we weren't necessarily looking for but we need to flag this because it's not fatigue but it might be something else.

MR. SEUACIUC-OSORIO: Can we hold that for when we get to the results of the field trial? Because I think that was going to give you an idea of the things that are being flagged. Because the idea here is exactly that, that the tool is not only flagging stress, corrosion, cracking.

MEMBER BALLINGER: Because this did not happen in France by the way.

MR. SEUACIUC-OSORIO: Okay. The tool is trained with a very -- it's very sensitive to all

those things. It's not trained really to differentiate between, for instance, something that may be a manufacturing -- manufacture defect or a condition that is there versus against a flaw or a stress.

It's not trained to do that. We don't want, today, we don't want the tool to do that. That's what we want the qualified level three to do.

So once we get to the examples here on the -- on the future, you are going to see that all the different things that the tool is pulling. You know, some of it is just a little bit higher noise in the weld. Some of it is, you know, fabrication conditions that were there.

We're going to see that there was a even geometric condition that it was not trained for but it was present on some of the nozzles and he flagged that. As well we're going to be able to discuss that and hopefully help answer your question. Is that okay? All right.

CHAIR BIER: So just to clarify, what kind of records is the computer actually looking at? Is it images or text reports, or?

MR. SEUACIUC-OSORIO: Very good. It's ultrasonic data, ultrasonic data, yes.

MEMBER BROWN: You've got to be an idiot savant to look at that stuff.

MR. SEUACIUC-OSORIO: It's tough, it's tough, yes.

MEMBER MARCH-LEUBA: At the higher level, there are two types of errors, false positives and false negatives.

MR. SEUACIUC-OSORIO: Yeah.

MEMBER MARCH-LEUBA: So I see in the next slide that the human is in charge of verifying that there are no false positives.

MR. SEUACIUC-OSORIO: Yes.

MEMBER MARCH-LEUBA: He's injecting his knowledge and identify that the machine did a 95 false positive. But the computer, the AI, is 100% level four identifying false negatives. It's rejecting and every one of the white records in that slide.

MR. SEUACIUC-OSORIO: Yes. So that is a good point. I wouldn't say this is a level four application.

MEMBER MARCH-LEUBA: I would. They get thrown out and the operator doesn't look at -- well, your slide.

MR. SEUACIUC-OSORIO: But the AIs just make -- you're right, the AI is making a call on what

is not relevant. And this is why this is one case, and this is why I said that the tool was trained with a very low sensitive -- sensitivity.

And if we were to put the AI, just itself, to a test, you'd never pass because of false positive. Because you'd have way too many falses.

MEMBER MARCH-LEUBA: False positives.

MR. SEUACIUC-OSORIO: Yes, false positives, thank you. Which is what we want in this case. So the job and the burden of proof here is in showing that indeed the AI is not throwing anything away that has a valid indication.

Because at that point, you're right, at that point the inspector wouldn't even have a chance, an opportunity to call that. So that is the burden on this case. It's totally biased to detection.

MEMBER MARCH-LEUBA: From the risk-informed regulatory mindset, you need to figure out what is the probability that one of those white records in the slide really was positive. And what are the consequences if you miss it.

And then you have to go into, do we have any handle of how to evaluate risk of all of these false negatives? Or do you just say your sensitivity is so low that you don't let anyone go through? In

which case you're not doing anything.

MEMBER HALNON: Jose, this is Greg. False negative based on the UT data or the AI inability to --

MEMBER MARCH-LEUBA: Whatever this example is showing, it's showing that there are 20 records in there. The machine decided that the white records are negative.

MEMBER HALNON: I mean, because there's a difference. There's a different -- if you got bad UT data.

MEMBER MARCH-LEUBA: It doesn't need to be UT, I'm talking about the generic methodology.

MEMBER HALNON: Okay, so it's an AI issue.

MEMBER MARCH-LEUBA: In a print job, you're reviewing on a torque on a screw.

MEMBER HALNON: Because the UT data is different than the AI algorithm. But you're talking about the AI algorithm.

MEMBER MARCH-LEUBA: No, I'm talking about the process. Just, yeah, you have a 20 records of whatever and you run it through the AI, and it says you don't have to worry about those 15. Look at these five that look weird and see if they're really positive. But the 15 don't look at them.

So you need to be able to evaluate what is the probability that you will have a false negative and what are the consequences of that. And again, for some things like -- for some things missing one of those records has no real safety consequence. So you are allowed to have a 90% probability.

For other things, it has a very high consequence. So you then find fine lines

MR. SEUACIUC-OSORIO: Yeah, and what we're going to see here us that what we're doing is we are -- you know, because the question then becomes how do we know that we're not missing anything. So you know, we're doing the tests we can. So we tested it on previous OE, previous operation experience, in cases where we know that the inspector has missed things.

You know, and then you realize, oh, I have something and you start looking back in the previous inspections and you see oh, it was there, it could have been identified. So we run that through the tool, and the tool -- if we see that the tool did flag those indications, even, you know, all those years before.

Or also, you know, and you will see here in the idea of the qualification that gets proposed here, it's the very same thing that the inspectors

themselves are subject to where they do need to show 100% detection on the procedure. These are the things that we can do.

You know, going forward and thinking of implementing this, this is why the initial idea is that we need two reviews of this data, two independent reviews, right.

So this becomes one of those and the other one by the qualified inspector still is with the inspector going through the entire volume of data to help us get a bit more comfortable in seeing, you know, do we -- if we go that way and we have those cases where we're doing those two, you know, is there anything that they -- that was found in the record, in the white records here in the example, right, or not. Or everything that the inspector sees --

MEMBER MARCH-LEUBA: It comes out with the term Dennis brought up every year validation of your progress.

MR. SEUACIUC-OSORIO: Yes.

MEMBER MARCH-LEUBA: Which is really difficult to do with one of this.

MR. SEUACIUC-OSORIO: Yeah.

MEMBER MARCH-LEUBA: And the other thing we've been using a lot is low-hanging fruit. So if

you have a system in which the consequences of missing one of those white records being red is not that bad and you can live with it, then that's low-hanging fruit and can do it.

If you cannot afford to miss one, then you cannot use the system.

MR. SEUACIUC-OSORIO: Yeah, so --

MEMBER MARCH-LEUBA: And the problem we have when we do digital I&C for protections systems is that software typically, when you implement procedures by a computer, the computer does better than you do every single time. The problem is that once in a blue moon, it fails catastrophically. Whereas humans fail more often, but they don't do those bad things. And that's something that has to be considered.

MR. SEUACIUC-OSORIO: And we'll get to this. When I mentioned the low-hanging fruit I was really talking from a perspective of the qualification. Because this already has a qualification framework in place that we can insert this into that and still following all those -- all the same -- all the same rules.

So a lot of the qualification questions in this case may be easier to answer. Next slide, please.

MR. HECHT: This is Myron Hecht. Before, I'm sorry, before you move on to the last slide, I wanted to respond to Jose and also your answer, Thiago.

When we've developed our own machine learning applications, specifically in the area of natural language processing, we developed what we called confusion matrices where we basically, using tagged data in our training sets and in our test sets, we were able to come up with a table where we had basically the two labels on, you know, the vertical -- the row and the predicted labels on the columns, and we were able to determine what the percentage of true and false were.

Here it's a fairly simple confusion matrix. You just have, you know, a positive and negative and whether they were actual or real. So having done that, and what were the results of your sensitivity and recall.

MR. SEUACIUC-OSORIO: There we go. So that is a very good point. So in this particular case, a confusion matrix will not really make a lot of sense, just because you're going to see a lot of false calls in that case there.

We talked about how this is a very much

leaning towards the recall part of this, right, the detection. We want to make sure that we're not missing anything.

And we're not too concerned with false, the false positives because the inspector would dispose of those, especially because the rate of false positives here is still much, you know, is still much smaller data volume than the original data volume.

So yeah, while we've been looking not directly as confusion matrix, as I said, because you're just going to see lots of false positives there. We have really been looking at the detection, the recall, right. Does it flag all the regions that need to be flagged, meaning all those white records don't have anything of concern there. And this is what we've been looking at.

And the other thing that we've been looking at is the call rate or flag rate, which is, you know, how much of the data has it been -- is it flagging. Because to say that, you know, there's nothing on the white records but I'm still flagging 90% of my data, it hasn't helped me much, right.

So we're looking in those two terms. Has it been detecting everything or flagging everything that should be detected or flagged, and how much of

the data has it -- has it flagged. Because we don't expect to be a lot to contain indications, so we want to see a reduction in that.

CHAIR BIER: One other question. In the area of AI in medicine, it's kind of typical that you have two radiologists reading, say, a mammogram. And people have found that it's more accurate if you have a radiologist and a computer because the second human may have the same blindspot as the first human, whereas the computer sees something different.

But that really is where the focus is on accuracy and the human makes their own judgment first before they ever see the computer judgment. Otherwise there is this tendency of the human to say, oh well, the computer must be right and I just go along with that.

And it seems like if you're using this for efficiency, to have the human not look at all those white records, it's a different setup, a different assumption than using it to improve accuracy. And the two may have a tradeoff against each other.

MR. SEUACIUC-OSORIO: Yeah, that's a good point. It is slightly different from the example from the medical field, right. Because there you are actually, the computer is actually making a call.

And here the computer is not, you know, the intention again is not that whatever the AI flags here is not determined to be a flaw because no plant would ever want to do that anyway because it would have much more trouble than it's worth anyway.

So but the idea really is that what we gain from I think, and I think we're see that with the field trial, what we gain from the fact that we have one of those two independent reviewers, one is still looking at the full data set. One is only looking at the select portion, you know, the portion that the AI selected for review.

What we gain for that isn't reliability because the likelihood that the human looking at only the smaller portion is going to miss something is much lower than for the other one who had to look at the entire data set. So I think there is that reliability point that you gain there.

There are other things you gain. There is -- there is some efficiency because the analysis can be done much faster. There is items with workforce, you know, and the fact that now you don't need seven qualified inspectors on feud, and it is an area where it's becoming more and more challenging to get that staff to do that, so it helps with that.

There is a cost benefit that comes with that. So there are other benefits as well. And I think I have a slide on a few other benefits or things that the utilities may use for this that will also help.

MEMBER HALNON: So did you -- did you do a study with the reliability of it's like a couple level threes that partied all night and gave them this data set and gave the same data set to the AI? And maybe another level three. Did you do a -- do a comparison to show that this was better?

MR. SEUACIUC-OSORIO: The study that we did was based on operating experiences, as I mentioned. So we had one case where, you know, we had a flaw that went, you know, that finally got -- that got detected.

And then when we went back and started looking at previous outages or previous inspections on the same components, you start seeing that, oh, that should have been detected much earlier because those signals were already there and it should have been caught, you know, six years before or six inspections before.

MEMBER HALNON: Okay.

MR. SEUACIUC-OSORIO: So what we did was,

okay, let's put all this history of data through the tool and see what it flags. And we saw that even from the earlier inspection, the region that had the flaws, they were flagged for review.

Now, flagged for review doesn't mean detected, it means detected from the AI perspective. The inspector still has to review that.

But the idea is that they would only have to review in that case, you know, like ten percent of the data, which were something if I recall in that case, versus -- so the likelihood that they would have caught that early at that stage is we imagine that is higher because he can focus his time on that.

MEMBER HALNON: Okay. (Audio interference) take a data set and set here and then look at this, you'd see how it came out and then you run the AI process through.

MR. SEUACIUC-OSORIO: Yeah, it's really easy, the difficulty is in getting the data to do this. You know, we have qualification data that we could use to build, but that's not the data that's in the plan.

MEMBER HALNON: No, I understand you want to get your own data, there's plenty UT data up there that is data from a raw UT, even the steam generator

UT stuff that we used to do.

MR. SEUACIUC-OSORIO: Well, but then that wouldn't directly apply to this, right. I mean, we can test the whole idea of the process in a way. So I think, if I'm understanding your comment is to test the process with the whole --

(Simultaneous speaking.)

MEMBER HALNON: With raw data, give it to two level threes like you normally would and then take that same data finally and give to the AI and then compare the results.

MR. SEUACIUC-OSORIO: Okay, so but not so necessarily to this application, it's the process in general.

MEMBER HALNON: Well, I mean, it just -- I'm looking at it very simplistically.

MR. SEUACIUC-OSORIO: Yeah.

MEMBER HALNON: You know, it doesn't seem like it's that difficult to prove the reliability. If I, you know, if you take that approach to it. It may come out to be the same, and you say, well, why put the money in it.

MR. SEUACIUC-OSORIO: Well, in that case, well, point taken, yes.

MEMBER DIMITRIJEVIC: Hi, this is Vesna

Dimitrijevic. So in my time when I walk a risk-informed in-service inspectors, the EPRI publish a report I think Frank Ammirato was the author. And I participate that on the, you know, maximizing reliability of inspection and balancing false calls, you know, there is the missed calls.

So that was the some -- there is some previous study done on how to maximize reliability of inspection. And of course that also depends on the size of flaws you are interested in the inspection, so. I know there is EPRI report about that because I participate in that, but I don't exactly remember what was the title of.

MR. SEUACIUC-OSORIO: Yes. Yes, so there's been -- there's been work on that. But I don't think that, to the point made here, those do not necessarily involve the addition of AI in the process to assess that.

MEMBER DIMITRIJEVIC: But that will help you on the what level of activity you can set your, you know, the AI participation on that. You know, to -- because you want to minimize the, you know, the red ones and you want to make sure that you don't have too many missed calls, so.

MR. SEUACIUC-OSORIO: Yeah, yeah. It is

definitely something to look into.

On this slide here are just showing two ways that AI in this case here, for ultrasonic data, can be helping.

On the left you have a screen model, which was what was done on the field trial for the -- for the nozzles, where there you see the entire data set and you can see the red highlights is where the AI is flagging that something is there and the inspector should go in. And they can click on that and they will get a more detailed view of that region to look.

And you also can see there, you know, it's a much smaller portion of the entire volume of data. But it does tell the inspector exactly where the flaw is. He will go there, he will see the region that's flagged, but there's no indication within the region what triggered that, it's the analyst needs to review that.

On the right is a more detailed model where instead of just flagging entire regions it actually goes ahead and puts a box around the flaws. You know, it's a more complex model.

In some cases it may be required, a lot of times it's not and it's not worth the effort or you know, it's more complex which means it also requires

more work and more data to go into that. Next slide please.

And here is some of the uses, and when we're talking about the value of this for the utilities of course. Ultimately what we want to do is have this is as a credited part of an inspection as a secondary review, for which qualification is required, as we'll discuss.

But even when we talk qualification, there's a number of things that a tool like this can help the utility and bring value to them. One of them is oversight. Utility staff can have a chance to review the data that they're getting from their vendor.

You know, when you go from 100%, so hey, I only need to look at these 2% of the data. I can actually get that done. And it's feasible to think that the utilities can leverage that then to get oversight of the inspection being performed on their -- on their assets.

And there is usages around reviewing previous outage data. In the case of the vessel head application, they can prioritize the inspection order because if there is a nozzle that's a little bit more suspect, they want to do that earlier in the spec, in

the outage, rather than later. They can assess needed resources, you know, how many people do I need to have available for that.

And you can also, it's a great tool to familiarize the examiners with the expected conditions of the plant. We'll see -- we saw that in the field trial, it was very useful to be able to go through the last outage results in a very short amount of time, so after the outage has begun.

And in four hours are able to cover the -- all the previous data. And they already kind of know, you know, what to expect, what is the condition, and things like that. Next slide, please.

Just showing here what the implementation looks like. The UT data is saved directly into a little machine learning box that you see there. That box detects, hey, I have new data. It will go to work and it will produce a report for the inspector to review.

All data remains local, which is significant for the utilities. There's no software needed, there's no extra training needed, nothing like that. Next slide, please.

Okay, so the field trial here. This is where earlier in the year we had the opportunity to go

to a power plant and carry out an inspection using this tool. So this was carried out in parallel and completely independent from the qualified vendor inspection.

So the vendor was on site doing their inspection as normal. And EPRI and our contractor were there as well on a different room going through the data as well but using the AI-assisted tool.

And we wanted to see, you know, is the implementation that we have proposed, does it work in the field, how does it perform. And we want to ask to compare the results to the vendor inspection.

So 78 penetrations were evaluated using this. In four hours, so we -- after the outage had begun, as I said, but before this specific inspection begun, the EPRI team had the chance to review previous outage data. So that meant all 78 penetrations were reviewed in only like in four hours. Because again, we're not looking at all the data.

In that time, we were able to see that, hey, there is a given geometric condition that's possibly being flagged. And we know that there is an oh, never mind, because it's -- we know it's a geometric condition that the model wasn't trained for.

So we were able to on site retrain the

model to learn that is a geometric condition which reduces 60 -- 60% further reduction in the data, so at the end of the day we only have 2% of the data for review.

Once we did this retraining, we had benchmark data available to us there so we could make sure this retraining is not affecting detection performance and things like that.

Average of four minutes per penetration. This was done much faster than of course what the inspector, the qualified inspection was doing.

So what we can see here, to the question that was asked earlier about what is AI triggering, we did exactly a detailed analysis of every single flag that the AI had what was the reason that was leading to that -- to that flag. And that's what the plot on the bottom right there shows.

And you can see there that a lot of this are things like J groove weld interface response for mechanical induced indication, fabrication, two-vent noise or lack of data. So you can see that is triggering if I have a bad data condition that for some reason I missed data, it gets triggered. And things like that.

So you know, you can see that over 50% is

related to either a mechanical induced indication, which may be bad probe contact or something like that, to weld interface, which are things that we want the inspector to go through. We want them to make sure that that's the case.

The red bar there, which is about 13% of the flagged regions, is the part that, okay, for these flags we could see no reason whatsoever. These were real just noise. And that's what we would consider real false positives from the AI perspective, the things that were flagged that really did not need to be flagged.

Well, we had 4.4 miles of data in total, that's what the qualified inspector needs to review. With the AI tool, that becomes less than 500 feet of data for review. That's the difference in the amount of data there.

All SCC, stress, corrosion, cracking, locations were flagged. In this case, there was indication that was -- I think there was also previous indications that had already been repaired. And but the model was even able to flag those as well.

At the end of the day, the analysis that we conducted -- we conduct -- the EPRI conducted with the tool led to the same results as the one from the

qualified vendor, meaning there was nothing that the qualified vendor indicated or saw that the EPRI reviewers didn't. Which is, this is the indication that in all those white records, in this case at least there was nothing really to mention.

MEMBER BALLINGER: But in this case, you were always looking for SCC.

MR. SEUACIUC-OSORIO: No. I mean, that is the name --

MEMBER BALLINGER: You would have -- (simultaneous speaking) correct?

MR. SEUACIUC-OSORIO: I would hope so. Now let me -- let me --

MEMBER BALLINGER: Both brings the tango, as they say.

MR. SEUACIUC-OSORIO: Yes. No, I mean, what I'm saying is, okay, the model was training mainly with SCCs. There were some thermal fatigues as well, okay. Did we have new thermal fatigue flaws to test this on? No, okay. In the lab specimens that we tested on, yes. It was flagged.

So what I want to correct is and make clear is that we were not specifically looking for SCC. There were additional, like thermal fatigue was also in the training data set. And even from this

plot that you see everything that got called, you can see that not only SCCs being -- is being called.

You have noise, you have lack of data, you have things like, you know, if there is interface responses on the weld. All those things are being called as well, and they're not SCC.

MEMBER HALNON: So where is the artificial intelligence piece, where is the learning piece? I mean, I get that it's filtering out records that don't have indications. That doesn't -- where's the learning? I mean, how's it getting better, how is it -- I guess I'm trying to understand.

MR. SEUACIUC-OSORIO: Okay.

MEMBER HALNON: You can call it AI as in -- but it seems pretty static.

MR. SEUACIUC-OSORIO: It is, and it needs to be for the reason that was brought up earlier about, you know, if you're going to allow this to evolve over time. And we'll get to that when we get to qualification here. But the learning piece happened before this.

So in order to come to this stage where we can get and have the model ready to go in a field trial, there was preparing the model, which included learning. The model had to learn on a number of

examples.

Those examples included field data from nozzles that are no longer in service, so we can use those to train the model. It involved some thermal fatigue that not from the field but developed in the lab. And it also involved some virtual flaws that are made based on those field flaws because we don't have enough field flaws to train -- to train a model.

So the learning here happened before. So before we got to the field trial, we trained the model with all this -- a lot of this data that I just mentioned, which we tested it on the OE that I mentioned. We tested on EPRI qualification samples and all that to make sure that it was ready for this.

Once we come to the qualification, in a little bit you will see that for a qualification framework to be valid, you have to bring that model to the qualification, and at that point the model is frozen.

Anything that you do to the model, even this retraining that we did here to the model on site, on a qualification framework would not be possible. Because if you do that, you are going to automatically invalidate the qualification you have to requalify.

Now, the requalification may be a simpler

process depending on what changes were made and things like that. But to do effective qualification, a model needs to be frozen.

MEMBER HALNON: So it is static, once you get it to --

MR. SEUACIUC-OSORIO: Once we get to this point.

MEMBER HALNON: Yes.

MR. SEUACIUC-OSORIO: Yeah. All right. Next slide, please. So here the qualification prospects, and we're going to talk about this. So a key point here is that this leverages an existing ASME qualification framework.

Okay, so this is -- there is already a qualification framework that's regulated, and we're leveraging that for this. We're not creating new things, we're just using the same thing that there is.

We looked -- this is a procedure type. So there's two different types of qualifications there, a procedure and a personal qualification. For a procedure qualification, it requires that all flaws within the scope need to be detected.

It's a 100% detection requirement. Because you need to show that the procedure is capable of detecting everything that is within the scope of

the inspection. And after that, those are followed by personal qualifications.

So we're looking at the same thing here. We want the model to detect all flawed regions because we want to give the inspector the opportunity to call those flaws. So the test that we did here, the model was exposed to more than five times the amount of PDI here, that's PDI is performance demonstration initiative.

More than five times the amount that they require for a traditional procedure qualification. So we really went above and beyond there what today what procedures are tested against. And the model flagged all the flawed regions that were within the scope of the procedure, which is what we wanted to see.

Yes?

MEMBER BALLINGER: So I know a little bit about what -- about the procedure down in Charlotte and stuff and the E Center. The flaws that people are training on are pretty big. And I'm curious as to whether you could use the AI side to detect things that the -- that are smaller.

I mean, that's the issue, finding flaws that you weren't trained on. You passed your level three -- you made all the pass, but the -- my

understanding is that those -- that that data, that library that's available, they're pretty deep flaws. I mean, they're more than 20%.

MR. SEUACIUC-OSORIO: Ten percent.

MEMBER BALLINGER: Ten percent.

MR. SEUACIUC-OSORIO: Yeah.

MEMBER BALLINGER: Yeah. And so what we're seeing now, what we did see, were less than 10%. I mean, eventually we found them. But so I'm wondering whether you could -- whether AI could use -- could be used to expand down the data that gets called as suspicious to in effect increase your sensitivity where the level three guy wouldn't have seen them and was qualified.

MR. SEUACIUC-OSORIO: Yeah. So the sizes in the qualification, they are dictated by the code. And that's why you're not going to see anything below 10%.

So but that is a very interesting thing. So here the idea is that we're putting, you know, we're proposing this AI-assisted approach and we're putting it to the same level of qualification that inspectors that are qualified to do it today, you know, the same level of requirements as they are. Now --

MEMBER BALLINGER: What I'm saying is it probably has implications for SLR, for long, much longer term.

MR. SEUACIUC-OSORIO: Yeah.

MEMBER BALLINGER: Where, yeah.

MR. SEUACIUC-OSORIO: Yeah. Well, I'm just making the point that, you know, the idea here is that we are putting this to the same test that we put the people that go inspect -- inspect it today, right. So this is the point of we're leveraging that existing framework and putting it to the same standards that people are today.

So to your question about can this be used to get a little bit, you know, to the smaller -- well -- so --

MEMBER BALLINGER: A little bit of learning is what Greg was thinking.

MR. SEUACIUC-OSORIO: Yeah. So the model can be trained on small flaws through -- now, the smaller flaws are mainly going to be done through the virtual technology because we don't have -- unless we have them, right. So yes, it is possible, and we can -- we can test that.

But still in the -- the qualification stuff is not going to test the capability because

that's outside the rules of the qualification program. So yes, we can train the model in smaller flaws to some extent, which makes it likely, you know, that you would detect. And we can test that to the extent that we have those flaws available.

But the qualification itself was not going to test that because that's outside of the scope of the roots.

MEMBER BALLINGER: You're kind of artificially increasing the POD for smaller flaws in effect?

MR. SEUACIUC-OSORIO: Well, do you mean --

MEMBER BALLINGER: The probability of detection --

MR. SEUACIUC-OSORIO: Yes, but --

MEMBER BALLINGER: But with virtual flaws.

MR. SEUACIUC-OSORIO: Yeah, yeah. No, well, so use virtual flaws in training, we never use them on testing. That's an important distinction. Every testing we do is only on real flaws. It's difficult to talk about POD for flaws that small because we do not have them.

It's, you know, if we want to test and if we go to the library and the -- with all the specimens and the -- we're just not going to have them because

they're outside of the scope. So it's really, that's an open question today.

And I work with POD and I get this question all the time, what happens below 10%. We don't know, we don't have the data to support that.

MEMBER BALLINGER: But with phased array, you may be getting it.

MEMBER MARCH-LEUBA: Yeah, just a comment. When you work on minimization problems, the hardest thing to do is define what the weight functions are. How do you weight one thing versus the other.

And I see on your way of thinking the way you talk that your goal was to minimize the amount of time the operator has to spend reviewing false positives. You want to compress -- compress the things the operator has to do.

If I put a black hat as a bad guy reviewing this, I'm working from the safety point of view, I would have preferred if you develop a model that identifies and learns how a good pipe looks like.

And you don't show it any flaws, you just show it good pipe, and it learns what a good pipe is. Identifies any deviation from good pipe. In which case, you can identify things that you didn't feel it.

And there is no incentive in doing that

because it's going to increase the amount of time the operator has to be review false positives. But it would more safe. And you have to weight where you're going just for safety.

MR. SEUACIUC-OSORIO: Yeah, so, you know, I would say that the goal here, yes, that we want to help focus the, you know, help the inspector focus in the regions that matter, yes, while maintaining 100% detection. We -- you know, there is that.

MEMBER MARCH-LEUBA: It's impossible to guarantee 100%.

MR. SEUACIUC-OSORIO: Yes, it is impossible with inspector --

MEMBER MARCH-LEUBA: As you set up a question, I was asking myself why when I was in the office of the Subsecretary of DOE, are you willing to bet your firstborn son that this is going to work. And I was asked that, and I said yes, because I was -- I was sure. But that's only because I'm very confident of the water.

I don't think -- I mean, you have to ask yourself that question when you work in the nuclear industry.

MR. SEUACIUC-OSORIO: Yeah, no but, the same issue, of you can't guarantee 100% detection, it

goes through AI or non-AI. With just the inspectors in the field today, you cannot guarantee 100% detection. And as a matter of fact, you know, there is -- oh, we just showed that there have been missed indications, right. So.

MEMBER MARCH-LEUBA: It depends on the -- I mean, going back to risk-informed, it depends on the consequences of you making a bad decision.

MR. SEUACIUC-OSORIO: Yeah.

MEMBER MARCH-LEUBA: And as I said, if the consequences are large, maybe concentrating a network that learns how the good pipe looks like, and then identify things that don't look like a good pipe to me. I don't know if it's bad or not, but something doesn't look good.

MR. SEUACIUC-OSORIO: Yeah.

MEMBER MARCH-LEUBA: It may be more reliable.

MR. SEUACIUC-OSORIO: Yeah. We would have to see the value in that as well because --

MEMBER MARCH-LEUBA: It's going to cost you more money.

MR. SEUACIUC-OSORIO: Well, yes. But you know, what a good pipe looks like, it's probably going to change from plant A to plant B. And it would be

very difficult to get a library of all of -- all of those.

MEMBER MARCH-LEUBA: And from plant A to plant B, you're going to have a good sense of what the flaws look like?

MR. SEUACIUC-OSORIO: Well --

MEMBER MARCH-LEUBA: That's what Ron is saying. I mean, you can identify a two-inch one but not a half a millimeter.

MR. SEUACIUC-OSORIO: Yeah, but I would -- I could probably argue that, you know, if you are looking for the flaws you're -- I think you're more likely to find them than if you're just --

MEMBER MARCH-LEUBA: I don't think.

MR. SEUACIUC-OSORIO: No, I mean yes, you're right. If you're trying to just look at, determine what a good pipe is and find everything that deviates from that, yes, you definitely should find the flaws --

MEMBER MARCH-LEUBA: There's something that smells funny here.

MR. SEUACIUC-OSORIO: But you would have way too many false positives --

MEMBER MARCH-LEUBA: You may not have an incentive to implement such system.

MR. SEUACIUC-OSORIO: Yeah, and at the point really raises the question is it is even worth doing it, right. Because this is one of the things that I said if we did this and the flag rate was still in the 90% or whatever, I'm still looking at pretty much most of my data. So why bother, right?

MEMBER MARCH-LEUBA: Yeah. And don't get me wrong, I mean, I'm not throwing cold water at you.

MR. SEUACIUC-OSORIO: Yeah.

MEMBER MARCH-LEUBA: It's because we like it.

CHAIR BIER: Yeah, if you can go ahead, I know there may be some questions on your last slide also.

MR. SEUACIUC-OSORIO: Yeah. Next slide. Now, here we go. We're talking about qualification, and this is a proposed qualification scenario. This is not necessarily what it's going to be like, but it's what it could look like, okay. And starting to answer the question how do we qualify this.

Again, there's the leverage that puts that process is already in place, right, and we're just inserting the I in there. As I mentioned, the algorithms would have to be provided before the qualification process and be frozen at that stage. So

any modifications after that would invalidate this qualification.

So the prerecorded data goes through the AI model, the AI model will kind of identify the areas for review, like we've been talking about. And here the qualification body will make the determination, right. This is a test, this is a qualification test.

At this point the qualification body will look at all those, the things that got flagged, and they'll saw were all the flawed regions flagged, yes or no. And if they were not flagged, if they were not detected, I'm sorry, it's a automatic fail.

There is no reason to continue because it's impossible to reach 100% effection because the inspector's not even going to see those areas. So that's an automatic fail for the procedure.

If the -- if all the flawed regions were flagged we go to the next step, which is where the human comes in. He would review the flagged indications, that what we have on the right.

He will make his calls, he will do -- he will work only on the -- on that subset of data, and he will make his calls just like he does normally. It is going to say whether it's a flaw or not.

So here, you know, mostly are the false

calls, false positives would be eliminated. And then the qualification body would make a pass or fail decision based on the same criteria that we use today for both false calls and detection.

So the inspector still has the chance to make false calls because he may not eliminate all the -- all the false calls, false positives in the data. And he also has the chance to make misses because even though the flawed region was detected, he may say that oh, you know what, this is -- this is not a flaw.

So the same criteria would be applied at that point. That would be what -- how we can leverage the qualification as it is today to qualify a procedure that includes the AI-assisted analysis.

Questions here?

MEMBER MARCH-LEUBA: Final comment.

MR. SEUACIUC-OSORIO: Yeah.

MEMBER MARCH-LEUBA: I love your work, I think it's worthwhile, and you guys did a good job. And it's, as I said, worthwhile. I am 99.5% convinced that the computer data system will detect more flaws than 100% human data system. Because the humans get bored and they miss things. Computers never do that.

But same thing that we've done with software-driven digital I&C, we're talking about one

case in a million where the computer misses something but misses it badly. And that has to be considered when we do the regulation, is that a one in a million when everything goes completely wrong.

And that is so difficult to define or validate. Just a word of wisdom, somebody doesn't tell us and look wise.

MR. SEUACIUC-OSORIO: You know the, with the initial at least approach that this would be a secondary review, if there is a catastrophic miss from the computer, those things that are like this is so obvious, how did you miss this, this is something that the human that's actually doing the other review should catch, right.

Any other, any questions here?

MEMBER KIRCHNER: Yeah, this is Walt Kirchner. Do you keep a separate set of data like the data set on the left and separate that out and then use that to test the algorithm?

MR. SEUACIUC-OSORIO: In the --

MEMBER KIRCHNER: How do you V&V this?

MR. SEUACIUC-OSORIO: So are you talking about --

MEMBER KIRCHNER: Yeah, how are you validating that this algorithm actually works well?

MR. SEUACIUC-OSORIO: Yeah, so in -- you're asking in the context of this slide here talking about qualification?

MEMBER KIRCHNER: Yes, do you separate out enough data from that middle, let's see, one, two, the third AI flags areas for a review?

MR. SEUACIUC-OSORIO: Yeah.

MEMBER KIRCHNER: Do you have a large enough data set that you then can test it?

MR. SEUACIUC-OSORIO: Yeah, so this would be the same data that's provided to, in a qualification test, to inspectors today. So yes, there's actually, you know, this is typically very flaw-heavy. You know, there's more flaws here than the inspector would see in the rest of their lifetime in a power plant.

So to answer the question, I think that the best thing to say is this is the same data that's used in qualification today. That's the data pool here that you see from left there and then in the middle here. And it is considerably flaw dense.

MEMBER KIRCHNER: Well, you're comparing it to the human success rate in looking at it. What I'm asking is do you hold back data for validation purposes, similar data, and then test the algorithm to

see, to verify and validate that the algorithm is independently, so to speak, working correctly?

MR. SEUACIUC-OSORIO: Well, that's testing the algorithm itself, and yes, we've been doing that. Even before we got to the field trial and before saying that we were ready to do that, we tested it with as much data as we had available and could use for this purpose.

MEMBER BALLINGER: Yeah, but what I -- what I think Walt is saying is that did you have a training set and a validation set?

MR. SEUACIUC-OSORIO: Yes, yes, yeah. And --

MEMBER KIRCHNER: That was done at the second box -- it's built into the box, the second box, the AI model is applied. So you did V&V on the algorithm when you built this box.

MR. SEUACIUC-OSORIO: So, yes, the model was tested following traditional procedures of validation and testing, and you know, everything that we've been doing and testing it against. So there was validation data test, there was a test data set. We tested it on additional sets coming from the field as they become available and we can do that.

MEMBER KIRCHNER: And then is that part of

your formal QA program?

MR. SEUACIUC-OSORIO: The part that would be done under QA here is actually the procedure. And the model is part of that. Right? So the qualification here would come in and the model that comes in and is frozen is qualified at that point. Does that answer the question or no?

MEMBER KIRCHNER: I'm just wondering whether you have independent data sets that you can then V&V the model against before you then go out and use it on a new set of data. Or is this very restricted to just UT testing for -- so the UT testing probably has its own quality assurance program and controls.

MR. SEUACIUC-OSORIO: Yes.

MEMBER KIRCHNER: So it's only for a specific technique, a specific machine --

MR. SEUACIUC-OSORIO: Yes.

MEMBER KIRCHNER: -- for the UT, et cetera, and then you V&V the model before you build the box there on the second object on the --

(Simultaneous speaking.)

MR. SEUACIUC-OSORIO: Yeah. Yes.

MEMBER KIRCHNER: Okay. Thank you.

MR. SEUACIUC-OSORIO: Yeah. The model is

going to be specific to -- in this case, to the UT data. Is it specific to a given procedure? It's a common procedure in the industry, but it is specific to a given procedure. So it's very specific. If we wanted to expand this to different procedures, a different model would have to be used. In this case, this is the one procedure used throughout the industry. So --

MEMBER KIRCHNER: Okay. Thank you.

MR. SEUACIUC-OSORIO: In a case where multiple procedures exist, then each case needs its own model. Yes.

MR. BURKHART: Chairman, this is Wyatt Burkhart from the ACRS staff. Just a time check that we have about -- less than 45 minutes on the presentations. We have another presenter. And a reminder that we're at the strategic plan phase. So I think this is all good discussion, but I just want to make sure you're aware of the time constraints, and we have another meeting this afternoon in a different Subcommittee.

MR. HECHT: This is Myron Hecht. I just wanted to make a short comment. Hopefully it won't take very long. That is that we're now -- in light of the last discussion and particularly Walt's comment,

we're now moving from qualification of the algorithms to qualification of the data. And of course, that's a fairly well-known problem in machine learning, that the data sets have to be qualified or protected or whatever needs to be done to ensure their integrity as well as the algorithm. That should be part of the strategic plan, in my opinion. What's being done to look at the data, the training data?

MR. SEUACIUC-OSORIO: All right. So next slide, please, and we're going to try to wrap this up here in a couple of minutes. I said I would have a slide just that provides a little bit of scope of the different activities that EPRI is doing. I'm not going to have time to cover them here. Most of them, I'm not the SME to even discuss them.

But as you review this, it can give you an idea of the different areas where EPRI's working. And their insights there -- most of it is related to natural language processing and trying to retrieve information from a large database more easily. And automation -- this is pretty much what we discussed here, the example there.

Prognostics is trying to look at the future and there's things like the FAC there is flow accelerated corrosion. We're trying to leverage

machine-learning techniques to get a better estimate of the wear rates so we can get better information on which components we have to inspect and when. So I'm not going to go into details here. This is just to provide an idea of the scope of the different efforts across EPRI.

Next slide, please.

I'm going to cover this real quick. Larry, keep going. There you go.

As a note, mistakenly, this set here was in the agenda, but it was not distributed. So you haven't seen this. I do apologize for that.

MEMBER MARCH-LEUBA: We just received it.

MR. SEUACIUC-OSORIO: You did? Okay. During meeting?

MEMBER MARCH-LEUBA: Yes.

MR. SEUACIUC-OSORIO: Yes. So we apologize for that. What I am missing is original set.

MEMBER BROWN: So am I.

MR. SEUACIUC-OSORIO: Next slide, please, Larry.

So I'll cover this real quick because of the time. And this is just for awareness of this effort. The idea here is -- we discussed how -- in the example that I just gave, answering the question

of qualification and how do we know it's good may be a low-hanging fruit or a little bit easier because we can use an existing framework that's already in place, and we can leverage that.

There is a question, though, how do we handle cases where that is not the case? So you heard from the NRC on their plan, and there are things to look at. What is left, the level of autonomy that is desired? What is the safety implication of the desired application? And things like that. So we are starting a collaborative approach between industry, the AI community, and regulators to start looking at that.

We've already discussed how some traditional V&V for software may not be applicable or may not be sufficient for all the things that the machine learning and AI brings. So this is an effort to start answering the question, what do we do? How do we address those?

Next slide, please, Larry.

We had a kick-off meeting earlier. This is where the IAEA is involved as well. There were a number of recommendations from this consultancy meeting, and I'm going to focus on the first one here on the right, which is develop a guidance document on

the deployment of error solutions for the nuclear industry.

Next slide, please.

And what we're going to start doing here in the efforts next year is we want to do two things. We want to gather use cases and desired use industry applications, so we're going to work with the industry and AI community and different folks to see, what are the use cases? Where do we foresee using AI in nuclear power plants, and to do what? And that will tell us what level of autonomy are we looking at, what are the safety implications of these different applications, and things like that.

And you also look at the unknown challenges of AI as they relate to deployment. And here, it really is -- the idea is let's gather all, and I'm going to use a term from integrity assessment. But anyway, we're going to use -- we want to look at the figure modes. How can AI fail? What are the things that it fails? We have things that we see and hear, like bridge on this, right? It works really well, and then all of a sudden, you give it something slightly different, and it goes off the rails.

We already discussed the -- we may have performance drift. We discussed earlier the thing of

if you are learning all the time and the model's changing continuously, that leads to its own set of challenges and problems. And we want to -- the idea here is bring all those to the table. We have a view for these are the type of applications. We have a view for these are the type of problems that we can expect.

And then we can begin assessment and see, okay, what are the problems that apply to which applications and where? Because they do not apply equally across the board. Explainable AI or having an explanation from what the AI is calling may not be necessary all the time, or it may not be necessary just at the same level of the time. So where is that necessary and to what levels?

So the idea is we have to put all that on the table. We have a large group of people going over this to discuss from different organizations and draw these generic guidelines and also identify what are the research gaps.

MEMBER MARCH-LEUBA: Will you consider this guidance document what I call software obsolescence? You spend all this money, time, and effort developing a model and developing it, and suddenly it only runs on one particular CPU that is no

longer available. So you need to consider that it needs to work for the next four years.

MR. SEUACIUC-OSORIO: Yes. That is a good point.

MEMBER MARCH-LEUBA: Don't waste time, but keep it in mind.

MR. SEUACIUC-OSORIO: Yeah. So we have a plan where we have a authorial board, which is a smaller group of people who are actually leading this work, and with a review board where that starts to get the results, review it, provide feedback, and things like that.

Next slide, please, is just going to have an overall timeline for the efforts that we are proposing here. The main thing here, as I said, is the effort's going to start with gathering those examples of applications and challenges or figure modes, putting them together to start drawing generic guidelines and things like that, culminating at the end of the year where we want to have this document, at least the final draft, complete and ready for publication.

In light of the time, I will stop here and see if there's any other questions or comments. I apologize. I know it was rushed here, this last bit

here.

CHAIR BIER: Okay. If there are any urgent questions, we can take one or two now. Otherwise, we do have a presentation from Exelon waiting in the wing. So -- okay. Well, thank you very much -- sorry. Constellation. Sorry. Wrong company. Anyway, thank you very much. It is very interesting work.

So okay. So we have new slides and new presenters. So --

(Off mic comment.)

CHAIR BIER: Yeah. Are the Constellation folks here in the room? Yes, it looks like.

MR. SZOCH: Yes.

CHAIR BIER: Okay. Good.

MR. SZOCH: Richard Szoch, and Gene Kelly is on the line.

CHAIR BIER: Okay. Great. Thank you, Richard. And I think we can go ahead and get started.

MR. SZOCH: All right. Well, good morning, everybody.

CHAIR BIER: Thank you.

MR. SZOCH: We are Constellation now.

CHAIR BIER: Yes. Sorry.

MR. SZOCH: I make that slip every day.

And Exelon -- I've been in the nuclear industry -- I'm Rick. So I've been in the industry for 41 years. And I think we've talked about with the NRC this is an exciting time, and -- was Exelon, now Constellation, fleet of 21 reactors -- are really looking at this whole topic, not just looking at it, but practicing it and trying to put it into place to bring value to the business both from an economic standpoint, risk, and safety standpoint.

So what I want to talk about today is what our big-picture vision, beyond analytics -- let's talk about digitization of the way we run our plants, both business process and operations. When I say operations -- not in the control room, but all of the support staff behind that operator, every belly button on our stations, is an operator. Everybody is responsible for safety.

So what can they do and how can we leverage these technologies to improve efficiencies, the way we do work, to focus on the right thing like we talked about in the previous discussion, focus on the safety-significant issues and the efficiency issues of the station and bring value to the business?

So I'm going to paint the big picture. We can talk about current applications of analytics and

then what we have in mind for the short- and long-term future. Sound good? Looking for some body language. All right. Thank you.

So, as I mentioned, Gene Kelly is on the phone. He's a manager in the risk management area, and I'm the manager of nuclear innovation. We have a centralized group. This is how serious we take this. We have a centralized corporate group now where we've taken some graybeards from the stations that have been there, done it -- engineering, ops, maintenance -- and apply that basic knowledge to how we can use technologies to take the way we operate the plants to the next level.

Next slide, please.

So, in words, this is our vision for where we want to be in the future. But I say vision -- we're doing most of this now, if not all of it. We're automating our work processes to get to a streamlined workforce, mostly business processes. When I use the term smart process -- as we speak, we're deploying what we call electronic procedures.

An operator of 21 reactors -- we have procedures that are standardized. They're almost exactly the same where it makes sense for every reactor that we own. We have the same number of

people. We have the same organizations. It's all built around that standardization to drive consistent decision-making, and that improves the way we operate and the way we maintain our safety margins.

So, now, we're taking that to the next step and improving that through digitizing our procedures so that the worker in the field will have a digital version of a procedure, and when they do a step instead of a circle and a slash -- okay. Initial I've done that step, and then take that document to the document control. It's all done electronically, and the historical data is saved forever. And they can go back and look at how it was done before, results of previous test or procedure, evolutions took place -- it's all digitized now.

Our theme is making decisions as we talk today through the use of AI -- or I'm going to use the word AI and analytics interchangeably, not to -- I don't want to get into the definition. You do talk to somebody different, get a different answer. But I'm going to define it in my terms. I think it lines up with our discussion this morning.

The big umbrella is analytics. We're using that, and there's shades of gray with AI under that. And I'm going to use those terms generically.

But we're making decisions based on those technologies.

MEMBER MARCH-LEUBA: Let me interrupt you. I know we're late, but let me -- I'm very worried about the concept of cybersecurity. Whenever you have a system that is that valuable and it is distributed among your 21 plants, and you run from iPads on a Wi-Fi network inside the plant, how vulnerable are you to a cyber attack?

MR. SZOCH: I'll address that in the next slide.

MEMBER MARCH-LEUBA: Oh. Okay.

MR. SZOCH: I promise. Hold me accountable to that.

MEMBER MARCH-LEUBA: I will.

MR. SZOCH: But the short answer is we have a very robust process for that, and we run through all of our data storage and our data communications through that process. That's the short answer.

MEMBER MARCH-LEUBA: Even the best guys out there get hit, every day.

MR. SZOCH: Okay. All right. So field workers' instant access to that information that we're talking about across those networks -- predicted plant

performance -- I'm going to talk more about that a little bit, but that's basically using these technologies to predict equipment failure well in advance of actual degradation of the equipment. I'm going to talk a little bit how we're using analytics to the effect today.

We have a centralized information backbone. I'm going to show that to you in a bit. That's very important as well. And then behind all that is a culture of innovation where every person at the station can come up with ideas and put those ideas into action. So it's not a top-down driven -- it's a culture of thinking outside the box to improve the way we operate, to improve safety, and to improve efficiencies of the units.

Next slide, please.

CHAIR BIER: If I can interrupt again, with this emphasis on digital, I want to say that digital has a lot of advantages. Obviously, nobody's going away from digital. But it can also have some disadvantages. I think some of the folks who are reactor operators talked about -- you miss your feel for the smell and sound and, oh, wait, that thing is hot, because you're looking at your screen.

And as a former teacher, I will also say

that there's a difference between asking a student to do something on paper and pencil versus asking them to do something on a screen. There's a lot more guessing on the screen and a lot less staring and figuring it out that you get on pencil and paper. So just wanted to raise that issue.

MR. SZOCH: I appreciate that. As an ex -- well, I still am an INC engineer from back in the day and post-TMI modifications. And I remember going to the controls, and this is our chance. We had all these lessons learned from Three Mile Island. Let's digitize this thing. This is in the '80s, right? And you know what? Not the right thing to do, to your point.

And quite frankly, we're staying away from the control room in this discussion. We're not touching the control room, we're not touching the boards, and we're not touching the algorithms behind the way we control the plant. I'll just cut to the chase right now. There's no reason to do that. There's not a good business reason for us, for Constellation. And from a safety perspective, there will be marginal gains, and the risk isn't for that right now.

But let's learn from the business

processes, put those in place, use the analytics, get the value out of that. And from there, maybe we can expand further. But the short term and even immediate -- the next generation of analytics is not going to impact operations of the unit. It's not worth the risk, and the value's no.

MEMBER MARCH-LEUBA: My earlier comment on cybersecurity had nothing to do with safety of the plant, biological releases from the plant, as with gaining control of your network of 21 plants and shutting you down.

MR. SZOCH: Security.

MEMBER MARCH-LEUBA: You can call it security.

MR. SZOCH: I mean, a general term, but --

(Simultaneous speaking.)

MEMBER MARCH-LEUBA: Maintaining operation. I can blackmail you for a lot of money if I gain control of your system and you cannot operate your plant anymore. So you have to take that extremely seriously and never, ever dismiss it.

MR. SZOCH: Absolutely. So I'm going to use that to launch us into this discussion here. With that in mind, we are embarking -- and I'm going to use another term that -- you talk to every belly button in

this industry; you get a different answer what a digital twin is. So I'm going to define it because everybody has a constitutional right to define it. I'm going to do the same thing.

So you'll see those left three arrows. See that digital twin cartoon, if you will? You see the nuclear plant to the left. It is a physical plant. To the right, with the zeros and ones, that's my depiction of a digital plant or digital twin. What that really means is we're taking the initiative to automate key processes. What I mean, processes -- that's basically maintenance and business processes.

In the middle there, data-driven analytics, which is really the focus of our discussion today. And thirdly, the data-enabled worker -- in other words, having the worker in the field with the information readily available so they can do testing, plan procedures as in the example I mentioned earlier.

And to the left of that, we have this hub, that round circle that really is all of our -- it's a centralized data approach. It's a data-centric platform. This is embedded -- we're partnering with a key software provider to deploy this platform not only across our nuclear but our non-nuclear side of the business as well, where we have all of our data in

one place. It's easily accessible by the users.

And on the nuclear side, we're embedding as we speak appropriate security to ensure that our cybersecurity procedures and guidelines are put in place as we go through the design of this -- what we call data hub. It's not only cybersecurity but export controlled as well. So we're taking extreme precautions to do that.

And some of the technologies don't lend ourselves to that. So we've had to turn away many technologies because of the cyber risk and the export control risk to housing that data and utilizing it both in the field and at our desks.

And then, finally, with regards to that station digital twin, you might ask, so what? What are you doing with it? So, in the bottom, with that information, you'll see down at the 6:00 at the bottom of the slide, corporate station monitoring. That's a depiction not of a control room, but we have a control center, if you will -- I don't want to even use control -- a monitoring center, a diagnostic monitoring center.

It's in our Kennett Square. It's outside our Philadelphia location. It's a room about the size of this, surrounded by screens, where we have a core

group of engineers that can look at any major piece of equipment in any one of our 21 reactors, and it has advance notice of equipment health well beyond what a control room operator has. The control room operator has limits, procedures, automatic alarms, a plant computer. That all stays. Plan to be operating the plant just as we are today.

But our monitoring center allows us -- hey, this pump bearing is -- I'm seeing a slight degradation. And based on the historical data that I have and the trained algorithms that we put behind us, if we keep going like we're going, this thing is going to fail in six months based on data. And by the way, this data is collected throughout our 21 reactors, throughout the industry, and even in some non-nuclear applications, we use data to train these models.

So individual picks up the phone, doesn't call the control room -- calls engineer. Hey, this is what I'm seeing. Can you validate? They go out in the field to do their validation and many times say, you know what? We're going to do preemptive maintenance. We're going to take advantage of a maintenance window and avert a failure and do maintenance on the equipment, take it out of service in a safe way. And I've got time to do it.

That's value. That's saving money on equipment. It saves transients, shutdown. It saves trips. And operations is informed. They don't even have to be involved with the decision other than the fact that we let them know what we're doing and why. That's probably the ultimate example of the value that analytics brings to the business both from a safety and economic standpoint.

The other two examples -- I talked about the mobile worker to the left there with a field worker having this data available. That's in place now. And of course, all the staff engineers and even maintenance technicians -- they have access to this historical data to supplement the processes they used to do work.

Next slide, please.

So here's our current examples of what we're doing today, and I'll talk about that at a high level. I just mentioned this advanced pattern recognition. I didn't use that term, but for those of you that are familiar with that, advanced pattern recognition -- I'm going to jump ahead and -- can you jump ahead, please, and then we'll come back?

Advanced pattern recognition has been in place for a good ten years at Exelon, now

Constellation. But we've -- I don't want to use the term again, but we've advanced that technology even further. And what that is is every piece of equipment has a thumbprint or a pattern. So it's like getting in your car. Your car has a feel, but if you went into your neighbor's car that's the same model, it's going to feel different. Even if it's the same model, same year, it doesn't feel like yours. When you rent a car, you go back to yours, it feels different.

Every pump -- reactor coolant pump, pump A, B, C, and D -- they all have a different pattern because they've all had a different break-in period. They might have had different replacements on their motors, on the seals. So they have a different pattern. So that pattern is recognized, and there's multiple points. There might be flow. There might be seal pressure. There might be bearing temperature. There might be oil diversity.

You look at all of those parameters, and that defines the pattern. And if you see multiple sources of that pattern change, through historical data and through machine learning we determine, wow, this thing is failing, and this is how it will fail. This bearing may burn out, or it may wear out. And this is the type of wear that you're probably seeing.

The analytic is down to that level where it can actually help you quickly determine the maintenance that's required. So that advanced pattern recognition is in use and, as I mentioned in the example of the M&D center, notifies the stations.

Can we go back one again? Thank you.

The next one that's really -- this one's really my favorite, ILT, the initial license training analytic. As some of you know, to train our operators, that's an 18-month rigorous process to train an operator so that he comes out of that training by passing the NRC exam and obtaining an operator license to be in that control room. So every station takes that very seriously.

Recently, I'd say one and two years ago, our training program didn't meet our own standards with regards to pass rate. We were having employees or potential operators fail that NRC license exam. So that investment into that student is lost if they fail the exam, because there's no second time around.

So we worked with our training organization, and we came up with one solution. There's multiple solutions so we can improve performance in that area. But we developed an analytic that looks at student performance starting on

day one, where when a student takes a test and may pass all these tests throughout -- but this analytic will detect either theoretical and/or system weaknesses based on the answers to specific questions.

So we categorized all of the questions on the exam in different fundamentals or different technical areas, and we can detect slight weaknesses in an individual. It might be maybe this individual -- his or her math aptitude isn't where it needs to be, or thermodynamics is their weakness, or a certain system -- electrical breakers. Maybe there's a weakness there, but they can still pass the exam.

So we're looking for those fundamental weaknesses earlier on so that we can detect potential failure later. This is an example where we were able to take our historical data -- so we took ten years of training data and to train this analytic. Then we went back and tested it on data because we knew who passed or failed, and it worked.

And we had criteria. We wanted some kind of a success rate. And don't quote me on the number, but it's on the order of we want 95 percent alignment to what actually happened. And we were able to achieve that through many iterations of this analytic.

So, now -- and by the way, as we speak,

we're deploying this to the fleet this month, this month and next. Now we have a tool that -- what was done manually before by training instructors where they would pay attention to the tests and talk to students and remediate students, now they have an analytic that serves that need.

We are very confident that through implementation of that, we'll either avert that student failure or prevent it from happening and/or recognize the weakness sooner so that we can give somebody else a chance to join the class and we can recover. It costs on the order of half a million to a million dollars to run a student through the program. It's not cheap. That's the economic piece. But we certainly don't want to challenge a student or borderline students in the program. We want to capture that early.

Maintenance rule functional failure is another example of an analytic that's been in place. The maintenance rule -- I'll explain it for those that might not be familiar with it. It's a program that's regulated which helps us -- the program will help us, the operator, detect weaknesses in the way we maintain equipment. So, for example, if we have failures either high end or unexpected, that results in what we

call condition reports or issue reports where we evaluate, why did the equipment fail? And if it did fail, what was the weakness in our maintenance program that allowed that equipment important to safety to fail, or high-safety-significant equipment to fail?

So, normally, that review of the program is done manually through looking at these condition reports and issue reports. We have an analytic that does -- I'll simplify it, but it's a word recognition and a category recognition analytic that takes a look at what the basis of that failure was, whether or not it's safety significant, and whether or not we need to go back and re-look at the program. It doesn't replace a human, but it quickly pulls that data together for the reviewer, the engineer, to determine, where are my risks in my program and where do I need to focus?

Same thing with the new work screening analytic. I'll talk some more about that in a little bit. We just applied -- and by the way, those analytics -- they are the same baseline analytics. In other words, we built the framework for the maintenance rule functional failures. You know what? This will work for looking at other types of condition reports. We can categorize those automatically.

And then we also found, hey, we're doing these license extension reviews going from 40 to 60 years on several of our plants, and that's a very rigorous review of our Corrective Action Program. We're applying analytics to that so we can look at historical data and figure out where we need to focus our reviews and provide justifications for continued operation long term. And again, it supplements the human in the decision-making and collating a very complex and very, I'd say, numerous amounts of data.

I talked about preventative maintenance a little bit, but preventative maintenance optimization is analytics we put in place similar to what I mentioned, but there are others that we put in place that optimizes the maintenance frequency. In other words, there's maintenance on the right equipment at the right time to avert failure.

And we've even looked at equipment where if I do this as a group in this one system, I can take that system out of service less often if I group it correctly. So that reduces risk. It reduces operator challenges and risk for the plant because I'm taking the system out once versus multiple times to take care of a pump in one window or a breaker in another window or another piece of equipment in a third window.

Let's do it all at once, do the testing all at once, and then return that system to the operator. Everybody wins.

It's a very complex thing to do to do that alignment, to maintain safety risk, and to maintain your maintenance rule functional failure criteria. But when you use the analytics, it makes it simple, and it puts it in the right window. I'll tell you a customer here is operations, but ultimately public because actually, you're doing maintenance in a safer, low-risk-profile way by not challenging the systems.

MEMBER MARCH-LEUBA: It's late, and you're having too much fun. What fraction of all these programs you're -- very worldwide programs are very successful, I'm sure. What fraction of them are machine learning, and what fraction of them are human-driven pattern recognition?

MR. SZOCH: The advanced pattern recognition is machine learning. They continue to build on themselves. The --

MEMBER MARCH-LEUBA: So it's self-adjusting?

MR. SZOCH: Yes.

MEMBER MARCH-LEUBA: It starts with something the human started with --

MR. SZOCH: Right. It's self-adjusting, and that's been tested robustly throughout many industries. Let me think a minute on the training analytic. That's not automatic machine learning, but it's continuously monitored and adjusted as we go through that. That's the plan, and the same thing with maintenance rule functional failure.

And the next four are relatively -- in the terms we used earlier today, they're static, but we're monitoring, and we are planning on and we have adjusted the analytic based on monitoring and feedback because we've been running these in parallel, and when we run them in parallel with the human systems, it's not a month or a week. It's for a year or two is what we plan. We plan on weaning ourselves away after we have that degree of confidence in --

(Simultaneous speaking.)

MEMBER MARCH-LEUBA: Yes. I just wanted to get a feeling of what fraction was it.

MR. SZOCH: It looks like about 25 percent.

MEMBER HALNON: So, Rick, we talked earlier that there's no -- none of these are obviously safety-related applications.

MR. SZOCH: Correct.

MEMBER HALNON: However, they do influence, if not drive, safety-related decisions on taking systems out of service, replacing equipment. Maintenance rule certainly acts through safety-related applications as well. Has the NRC engaged on that piece of it from the standpoint of the decision-making as a result of this AI?

MR. SZOCH: Yes, especially on the maintenance rule since that's the most mature and it's been in place. But we've done that and deployed that in a very transparent way with the NRC, and they've looked at it specifically. This is back when we were Exelon.

They've looked at it and reviewed it. We've reviewed it with them, and it's reviewed when we do the annual and biannual reviews of the program. It's constantly revisited.

MEMBER HALNON: Okay. Since you mentioned it, APR is more closer to the machine learning or the AI that we're kind of dealing with. What about that? Have they got into that? Has that been a subject of this table talk and research that they're starting to get into?

MR. SZOCH: Not to my knowledge, it has not, and I'd say for two reasons. It's been around a

while, number one, but number two, it's just purely informing the engineer. Everything is still validated through monitoring and looking at maintenance programs. The decision-making is simply enhanced and made better. It's not replaced.

MEMBER HALNON: Okay. But that's clearly an area that can be learned. I mean, when someone's trying to get into AI and what the application is and what it can do for us --

MR. SZOCH: Yes. Yeah.

MEMBER HALNON: -- that's certainly an area.

MR. SZOCH: We agree.

MEMBER HALNON: Okay.

MR. SZOCH: Yeah.

Next slide, please. We talked about that. Next slide, please.

I'm going back to my favorite one again, but this is kind of just a graphical representation or user interface that we plan on giving either this form or something different. As we speak, we're talking with our training organization, and we're refining this. But you can glance at this and say, oh, I have three spikes. Those three spikes in data, those are students that are at risk because of some of those

fundamental weaknesses.

So this is type report that would be given to our training organization on a weekly basis for some periodicity so they can monitor training effectiveness and whether or not an individual is at risk. In this case, you'd have three that were at risk, and the trainer would pull the data, which will be made available then, and perform, again, remediation much earlier in the process than we did in the past.

Next slide, please.

We talked about maintenance rule functional failure --

(Simultaneous speaking.)

MEMBER KIRCHNER: Pardon me. Pardon me. This is Walt Kirchner. Was that last slide -- it strikes me that's personnel information. Are those fictitious names or --

MR. SZOCH: Yes. Yes. Yes, of course.

MEMBER KIRCHNER: Oh. Good, because these spikes have been made public. Thank you.

MR. SZOCH: Yes, sir. Thanks. I probably should have said that up front. I thought somebody might worry. Yeah. Absolutely. Yes.

Next slide, please. No, you're good.

Functional full analyze. We talked a little bit about that, so I'm going to skip. I think we ran that to ground. Next slide, please.

This is an example -- and I wanted to pause on this because I heard some discussion on it this morning. To the right, big picture, you'll see some orange bars there and some blue. We have a Committee today that reviews condition reports, or some plants call them issue reports, on a daily basis. And we screen them for safety significance and impact on maintenance. There's multiple screening criteria. In fact, you can count them out there. There's -- I don't know -- 10 or 12 that are listed there.

The analytic that we use that does that now takes a first shot at it. By the way, it takes a first shot at it, and as we discussed earlier, I think at Thiago's example, it informs the reviewer to where to focus more. So you got to focus -- in this case, focus on the orange ones because the analytic may not be as reliable. I didn't get as much of a match.

It can sense the matching that it gets with the data. It might be just a 60 percent match or a 50 percent match on the word recognition, if you will. I'm oversimplifying, but that's basically what it does. And the blue, you got a pretty solid match.

So start with the orange, review those rigorously, and then continue your review from there. So, again, it allows focus for the reviewer, not to ignore the blue, but to make sure you really focus on the orange or the high-impact ones that the analytic didn't have as high of a match.

And when you're doing these types of analytics, they can get very complex. This one has like seven or eight different word recognition models in it. And we've been training it for about six months now, and we have found that -- like we talked this morning, the result -- and we've compared it against actual physical human reading. And it's better. We got to the point where it does better than the humans do.

It has word recognition -- the way our issue report process is, and sometimes it's a yes or no. It's pretty black and white. Some of this stuff's black and white, and the machine's not going to mess that up as much. The human will. So, from that standpoint, we saw an improvement there with regards to quality of review.

MEMBER HALNON: So, Rick, I was going to ask -- Theresa Sutter at Scientech introduced me to a term many years ago called Faint Signals.

MR. SZOCH: Oh yeah.

MEMBER HALNON: And is this part of that where you're looking for those faint signals of early --

MR. SZOCH: Yes.

MEMBER HALNON: Just like APR, looking for early indications of trends?

MR. SZOCH: APR is like a key example of faint signals. I like somebody else using that term. We use it all the time. But that's a faint signal that instead of hoping and praying that it goes away, you take action on it, and it's sticking --

MEMBER HALNON: Good.

(Simultaneous speaking.)

MR. SZOCH: -- in between the eyes.

MEMBER HALNON: That's good.

MR. SZOCH: What's really cool about it is -- and I know I'm excited about this because it's really advancing -- is that the customers, the power plants, they're responding to that. So we get the phone call, hey, I got this faint signal. Yeah. Well, I think you caught me last month. And I'm going to go after it. I'm going to go after it tonight and get that feedback. So that feedback loop and the confidence thereof is continuing to increase.

CHAIR BIER: Yeah. I just wanted to interject briefly. I'm very glad to see that focus because on the regulator side, a lot of what we see is, oh, our preventive maintenance signal is saying we can skip this maintenance. And it's nice to see the ones that are in practice that say, no, we're going to accelerate that because it looks fishy somehow.

MR. SZOCH: Exactly. Make decisions on data. And on the example that I brought up -- to build on that, if I may -- on the preventive maintenance optimization, in some cases, we found based on historical data that you could do maintenance more frequently. We stretched it out too far. We found some of those. And then, if you're going to do it less frequently, then I could put it in the optimum window. So it really truly optimizes it from a safety risk, not just an economic, basis.

Next slide, please.

All right. So, in the future, this is the type of thing we're thinking about. I talked about the work order of a PM optimization. PM is preventative maintenance. WOs are work orders. So, sometimes, a preventative maintenance activity is done under a work order, and I'll explain that in a second.

But if you can detect these failures or

have these faint signals early, why not take it to the point where, okay, faint signal. It's probably a bearing. It's probably a Babbitt based on -- Babbitt on your bearing that's going because of the indications that I'm seeing -- automatically creates a work order, it sends it to the planner for review, and then it goes to maintenance to automate that whole thing.

And the lead planner, towards the end of the process, goes, oh, yeah, all this makes sense. Today that's all manually done. And historical data -- oh, I think I've done this before. Let me go back and check my files. This would all be automatic, including the assessment of what the problem really is. The analytics, at least in the APR arena, have become that mature.

And there are some non-nuclear facilities that are already doing this. The work orders are automatically generated based on the indications seen in the field. And I think that's really cool, especially if you can do it with a high degree of reliability. But that's future, and we're building up to this. So that's our vision.

The next one is -- I think this one's big. The Station Oversight Committee, or the New Works

Screening Committee example that I provided to you, as well as the maintenance rule functional failure -- those are our toes dipped in the water for the CAP, the Corrective Action Program. Those are both looking at Corrective Action Program documents and basically looking at the data and taking action from the trends of data.

That data is pretty black and white. It's yes, safety significant; no, it's a piece of equipment. But we want to expand that more broadly to the Corrective Action Program. More broadly means process issues, human performance issues, and equipment issues. Process issues and human performance issues are softer data. It's much more difficult to apply these analytics for that type of soft data because it's interpreted by individuals before it's called data, if you will.

In other words, if I observe you doing work in the field and you didn't use place keeping when following a procedure, and if that happens over and over again, you can miss a procedure step. Well, that's my interpretation. So there's that human element. Some people may call that place keeping. Another person might call that attention to detail, and it's categorized differently. And it's hard to

collect that data in a consistent way.

With machines -- I'm talking physical equipment -- it's either hot or it's cold. It's over temperature, or it's under temperature. It's either high vibration or not. But when you get into human performance, it's not as black and white. But there are tools out there that we can fine-tune that and get those analytics to work to some degree.

So we're not giving up on that science, and neither is industry. We're looking at that. We've spent a lot of money on it so far and trying to make some of those analytics work. We're not happy with that yet. We're not going there yet. That's something in the future.

Talked about equipment performance, how to schedule predictions. We have done this where we're able to take historical data and predict future schedule performance in an outage. And again, that really is a safety issue. Here's why. We have found through experience, and I think you would all agree, that if you deviate from an outage schedule, you are challenging the risk of that shutdown condition of the unit, and you're challenging your safety systems. You're challenging all of your lineups if you do anything that's out of order to the original schedule.

So we've developed an analytic that predicts future schedule challenges that may avert or cause us to change the schedule well in advance based on previous outage schedules that we deployed throughout our fleet of 21 reactors. And here's an advantage of being standardized. We do all our outages the same. We use the same procedures, basically the same systems, PWRs and BWRs.

But we learned from that, and we can actually predict scheduled performance. More tuning to be done. We haven't deployed it yet. It's a tough one, but we're getting very close on making that work for us, such that the management team and what we call our Outage Control Center that manages that outage, can reliably use that data to make decisions. So that's why I had that out as future.

And then, finally, I mentioned before and I'll end with we're not going into the control room. Just don't see the value in it. It's not worth the risk, especially with the maturity of artificial intelligence analytics today. We're also -- Constellation, we're advising on an SMR overseas, and we're pretty much -- we've been asked to provide this digital vision consultation. And that's our current stance today, even for future designs.

MEMBER HALNON: Rick, you mentioned you're not going in the control room. Is there any work on shutdown risk and/or risk monitors that the control room uses relative to clearances, tagging systems, work order systems, and grabbing all that data and saying, hey, the confluence of plants are lining up; you better watch out?

MR. SZOCH: Let may do a repeat back. Are you talking about looking at taking certain systems out of service --

MEMBER HALNON: Well, like --

(Simultaneous speaking.)

MR. SZOCH: -- with regards to operational --

(Simultaneous speaking.)

MEMBER HALNON: The online risk monitors -- if we look at the tagging on and the work order problems and stuff like that, but tying in APR and other things to tie into PRA-type calculations and --

MR. SZOCH: We haven't gone there yet.

MEMBER HALNON: Okay.

MR. SZOCH: We're not there yet. Lots of talk about that. And one example that I brought up with regards to optimizing maintenance in the optimum window is a baby step, very small baby step, in that

direction. But a lot of talk on that.

MEMBER HALNON: Last question for me is -- excellent presentation. Appreciate all the information. Huge help. What does the industry organization look like? Where is the clearinghouse? Is there a central group like CAPOG (phonetic) for corrective action, or how are you guys organizing around what you're doing and others are doing to make sure that it's not a confusing spaghetti dish going to the NRC saying, figure out what you're going to regulate?

MR. SZOCH: It's really -- I'm looking over at Jim. It's really through the NEI. We've got the Committee, and Jim facilitates that for us, Jim Slider with NEI. And off of that, we've spurred off several other industry groups to have these discussions. But these types of discussions occur in Plant Modernization Committee that we've been involved with, with EPRI and NEI, which -- and the same people attend, so there's lots of crosstalk there.

MEMBER HALNON: Let me ask NEI real quick.

Jim, do you have a working group, I mean official NEI task force? How are you doing that organization inside NEI?

MR. SLIDER: We formed an Innovation Task

Force, Greg, and it convened earlier this year. And that is our forum for gathering the industry innovation managers like Rick, as well as regulatory folks who can advise us on strategy for addressing regulatory barriers. And we also have EPRI and Idaho National Laboratory Light water Reactor Sustainability representative.

So it really is, as Rick said, a clearinghouse for all of our industry innovation --

MEMBER HALNON: Okay. And what working group does the task force report to?

MR. SLIDER: The NEI's Regulatory Issues Working Group.

MEMBER HALNON: All right. Okay. Thanks, Jim.

MR. SZOCH: Next slide, please.

I think I covered all these points. Any questions? Anything else I could help you with?

MEMBER MARCH-LEUBA: No real questions, but I want to emphasize again cybersecurity, cybersecurity, cybersecurity. This morning, I was joking with one of my students: there are three levels in the knowledge, and you know it. On level one, you don't know anything, and you admit it. Level two, you think you know everything, but you don't. Level three

is you know just a little bit and you're scared to death.

I'm on level three in cybersecurity, and I think you're level two. Please take it seriously because you are a super fat target. People don't realize how serious this problem is. If we can fire up the Tor Browser now and go into the dark web, you will find on there credentials to log into 500 different companies for sale. And the average price is $2,000. Yours will be much higher.

You can find for sale phishing kits for 100 bucks, and that's why you receive so many emails phishing. And that's how everybody goes and they bypass the multifactor authentication. They bypass all the credentials. You are sold out if you are connected to the web. And you are connected to the web, and you have hundreds if not thousands of logins. Oh my God. Please talk to your cyber guys.

MR. SZOCH: I will. Believe me.

MEMBER HALNON: That one is tough.

MR. SZOCH: Absolutely.

CHAIR BIER: Well, and just, I think, yesterday in the news, there was a story of a U.S. company that turned out to actually be a Russian company with a pretend mailing address in Maryland,

and their app is used by CDC and the Army.

MEMBER HALNON: How many companies do you want me to open tomorrow?

CHAIR BIER: Yeah. Yeah. So, anyway, other questions for Rick? And then, after that, I guess will go to public comments.

Okay.

MR. SZOCH: I appreciate the feedback and comments. Thank you very much. And we're committed to continue to be transparent with where we are and where we're going. I'm looking forward to talking to --

MEMBER HALNON: For the record, I want to support what Greg was saying. This was a fantastic presentation, it's a great program. I mean, I started my career doing that, pattern recognition, okay? So I love it. I love that you're doing it. It's just you have to protect it.

MR. SZOCH: Thank you.

MEMBER HALNON: You have to protect it.

MR. SZOCH: Thank you.

MEMBER BROWN: I wanted to spring a little bit off of Jose's comment when he emphasized the cyber issue. The perception of cyber issues varies in terms of how you implement that in a plan with these levels,

coming inside the plant and within the plant, the various strata of level one, two, three, and four, et cetera.

People fail to recognize that in fact the first level of defense, safety and even critical non-safety systems, is the equipment itself. Every piece of equipment, safety or critical non-safety, that communicate bi-directionally, not into a network of in kind, is compromised its ability to withstand any type of intrusion.

If a system does not need to communicate bi-directionally -- all it has to do is send data out -- it ought to be one way, data diode hardware, whatever the terminology you want to use. And right now, everybody conflates incorporating that into the systems as a cyber issue, not a design issue when you build the stuff.

You can't come in five years later and say, oh, holy cow. I now have to change all that. That's dumb. Now, where you require bi-directional in order to measure a thing, then you have to make sure it has its adequate protections at the next level. But you shouldn't say, I'm going to depend on that fence. It's like physical security. You have to get inside the fence. Then you've got to get inside the

plant. Then you got to go to the main control room, and then you got to get the key to go open the door to where the instrumentation is, your systems are.

That is the first level, and it ought to be impenetrable if you don't need to do anything else. And once you get outside the systems, now it's in the big cyber world of how you do it. But depending on a network or another network to protect you, and the thought that you can use wireless for these critical -- is just -- I'll use a nasty -- insane. Okay?

MR. SZOCH: Thank you.

MEMBER BROWN: I'll stop right there. I had to get my pitch in while I have some industry people around.

MR. SZOCH: That's good. No, I appreciate that good feedback. Appreciate it. Thank you.

CHAIR BIER: Okay. Thank you very much.

A reminder for everybody, the speaker -- what?

(Off mic comment.)

CHAIR BIER: I'm about to do that. We're not adjourning yet.

Just a reminder for everybody who's been listening along, stay on for public or whatever -- the ACRS is not planning to write a letter summarizing or

responding to these presentations. So the comments you heard today are some excellent suggestions, but just from individuals that people can consider at their leisure.

And now I will open it up if there are any public comments.

So, hearing none, we will move on. I wanted to thank all of the speakers and Christina Antonescu for her work in organizing this presentation. Obviously, the topic is of great interest to most of us here on the Committee, and I'm sure we will have lots of opportunity to interact in the future. But just like the NRC is investing in trying to get prepared for what's coming that way, ACRS is trying to do the same. So it's very helpful to us to hear what you folks are working on.

And with that, if there are no further comments, I think we are ready to adjourn. What time do we regroup this afternoon?

MR. BURKHART: 1:00.

CHAIR BIER: 1:00. Okay. So we get a very quick lunch, but ten extra minutes for lunch. Okay. Thank you all very much. Bye.

(Whereupon, the above-entitled matter went off the record at 12:13 p.m.)

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

Overview of the Draft NRC Artificial Intelligence Strategic Plan Fiscal Years 2023-2027

Advisory Committee on Reactor Safeguards
Joint Digital Instrumentation & Controls and Plant Operations Subcommittee Meeting

November 16, 2022

Presenters

- NRC Technical Staff Presenters
  - **Matt Dennis** – Reactor Systems Engineer (Data Scientist), RES/DSA/AAB
  - **Dr. Trey Hathaway** – Reactor Systems Engineer (Code Development), RES/DSA/AAB
- NRC Management AI Champions
  - Dr. Teri Lalain – Deputy Division Director, RES/DSA
  - Luis Betancourt – Branch Chief, RES/DSA/AAB

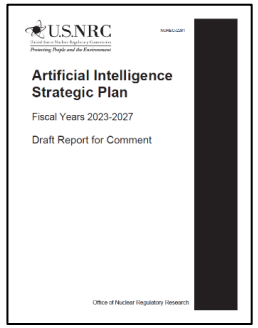
Agenda

- Artificial Intelligence (AI) Landscape and the NRC
- Draft AI Strategic Plan Development Background
- Draft AI Strategic Plan Overview
- Status of Draft AI Strategic Plan and Next Steps

Artificial Intelligence Landscape and the NRC

NUCLEAR INDUSTRY (EXTERNAL)

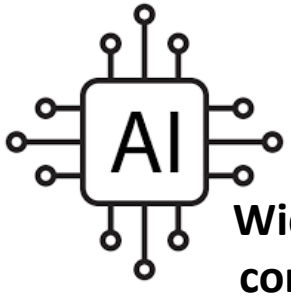
Industry wants to use AI



Draft AI Strategic Plan to prepare staff to review AI

OTHER CONSIDERATIONS AND OPPORTUNITIES (EXTERNAL)

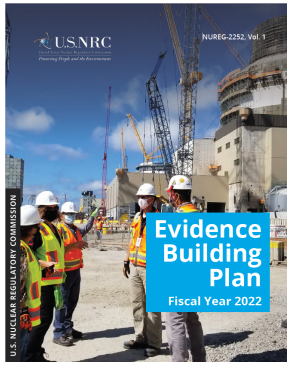
OMB EO 13960 and reporting requirements for implementing agencies



ACTIVITIES

Wide range of AI meetings, conferences, and activities

INTERNAL TO THE NRC



NRC Evidence Building Priority Questions

Internal interest in researching AI-based tools ranging from AI-embedded in commercial applications to custom programming

Draft AI Strategic Plan Development Background

- Formed an interdisciplinary team of AI subject matter experts (2021)
  - Leveraged insights from three NRC Data Science and Artificial Intelligence Regulatory Applications Workshops*
  - Engaged across the agency (OEDO, NMSS, NRR, NSIR, OCIO, OCFO, OIP, and Regions)

- Proactively researching AI usage across the nuclear industry, Federal government, and international counterparts
  - Maintaining external awareness via participation at forums, workshops, meetings and the NRC/EPRI-DOE MoUs

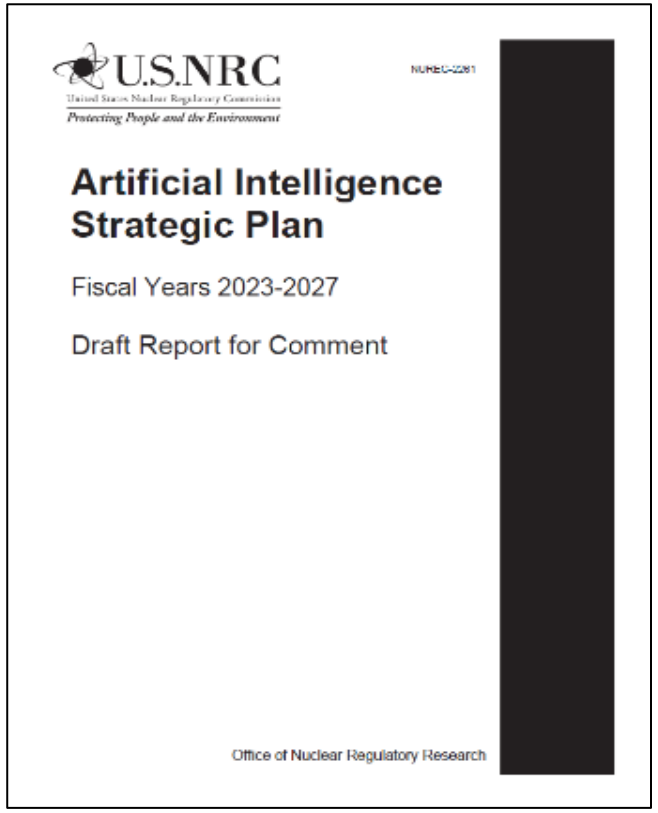
- Early stakeholder engagement and data gathering to inform the Draft AI Strategic Plan development
  - Finalized Draft AI Strategic Plan and continued internal alignment and briefings (Spring 2022)
  - Published Draft AI Strategic Plan for comment and held comment-gathering public meeting (Summer 2022)

*Available at https://www.nrc.gov/public-involve/conference-symposia/data-science-ai-reg-workshops.html

NRC is not alone in considering AI and we're taking a proactive approach

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

Draft AI Strategic Plan Overview


U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

NUREG-2261

Artificial Intelligence Strategic Plan

Fiscal Years 2023-2027

Draft Report for Comment

Office of Nuclear Regulatory Research

** Available at [ML22175A206]*

Vision: Continue to keep pace with technological innovations to ensure the safe and secure use of AI in NRC-regulated activities

High-Level Outcomes

- AI framework to review and evaluate the use of AI in NRC-regulated activities
- AI skilled workforce to review and evaluate the use of AI in NRC-regulated activities

The Draft AI Strategic Plan consists of five strategic goals

- Goal 1: Ensure NRC Readiness for Regulatory Decisionmaking
- Goal 2: Establish an Organizational Framework to Review AI Applications
- Goal 3: Strengthen and Expand AI Partnerships
- Goal 4: Cultivate an AI-Proficient Workforce
- Goal 5: Pursue Use Cases to Build an AI Foundation Across the NRC
- Appendix B on acknowledging internal uses of AI

Notional AI and Autonomy Levels in Commercial Nuclear Activities

Human Involvement

Level	Notional AI and Autonomy Levels	Potential Uses of AI and Autonomy in Commercial Nuclear Activities
Level 1	**Insight** Human decisionmaking assisted by a machine	AI integration in systems is used for optimization, operational guidance, or business process automation that would not affect plant safety/security and control
Level 2	**Collaboration** Human decisionmaking augmented by a machine	AI integration in systems where algorithms make recommendations that could affect plant safety/security and control are vetted and carried out by a human decisionmaker
Level 3	**Operation** Machine decisionmaking supervised by a human	AI and autonomy integration in systems where algorithms make decisions and conduct operations with human oversight that could affect plant safety/security and control
Level 4	**Fully Autonomous** Machine decisionmaking with no human intervention	Fully autonomous AI in systems where the algorithm is responsible for operation, control, and intelligent adaptation without reliance on human intervention or oversight that could affect plant safety/security and control

Machine Independence

Common Understanding of the Level Key for Regulatory Readiness

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

KEEPING THE END IN MIND – DETERMINING THE DEPTH OF REVIEW

Goal 1. Ensure NRC Readiness for Regulatory Decisionmaking

AI Research

- Determine approach to assess AI (e.g., XAI, trustworthiness, etc.)
- Development of AI standards and identify where gaps exists

Framework and Tools

- Clarify the process and procedures for AI regulatory reviews and oversight
- Consider options for long-range changes for AI regulatory reviews and oversight that may require rulemaking

Communications

- Public meetings to inform key activities
- Agency-wide internal communications and coordination to harmonize AI activities

Outcome: Develop an AI framework to review the use of AI in NRC-regulated activities

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

ENSURE CROSS-AGENCY LEADERSHIP IN AI WITH CENTRALIZED APPROACH

Goal 2. Establish an Organizational Framework

NRC AI Steering Committee

Cross-agency strategic alignment and direction

Centralized coordination of resources, priorities, and use case analyses

NRC AI Community of Practice

Lead best practices for reviewing requests that use AI technologies

Provide agencywide awareness on active and potential use cases

Facilitate the sharing of knowledge and lessons learned

Outcome: An organization that facilitates effective coordination and collaboration across the NRC to ensure readiness for reviewing the use of AI in NRC-regulated activities

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

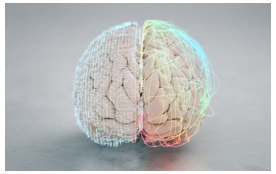
GAIN VALUEABLE INFORMATION TO BENCHMARK AI ACTIVITIES

Goal 3. Strengthen and Expand AI Partnerships

Maintaining Awareness

EXAMPLES INCLUDE

NIST AI Risk Management Framework standards development



Data Science and AI Regulatory Applications Workshops

Communication Forums

EXAMPLES INCLUDE

Public meetings to inform key activities

National labs, vendor and academic AI-focused workshops and seminars

GSA-sponsored AI Community of Practice

Working Partnerships

EXAMPLES INCLUDE

NRC/DOE-EPRI MOUs on Data Science, AI, and Operating Experience

US-Canada-UK trilateral engagement on AI

IAEA Consultancy and Technical Meetings on AI

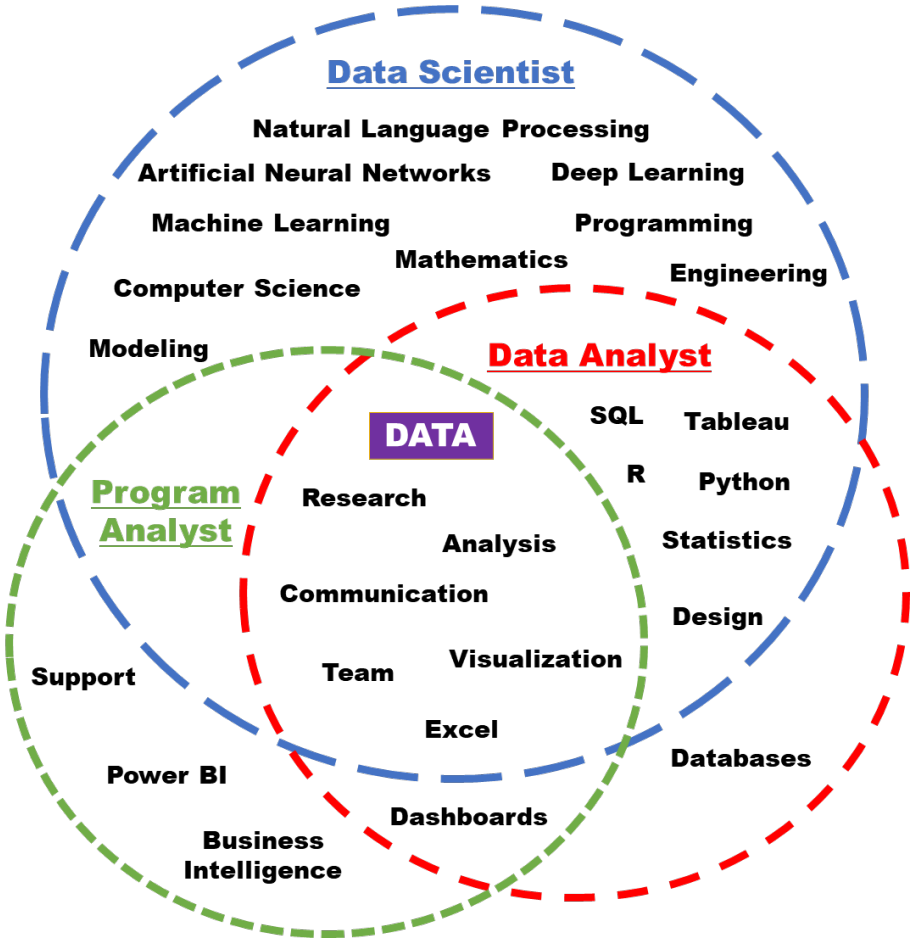
Outcomes: (1) Maintain awareness of industry plans, (2) establish communication forums to discuss future plans and regulatory needs, and (3) effectively partner with other agencies on AI topics of mutual benefit

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

ACQUIRE, DEVELOP, RETAIN, AND TRAIN AN NRC AI KNOWLEGABLE STAFF

Goal 4. Cultivate an AI Proficient Workforce

- Focused on developing the critical skills for the AI workforce of tomorrow
- Workforce Planning
  - Perform gap analysis, assess position descriptions and review hiring needs
- Staffing
  - Working toward the right skills at the right time
- Training
  - Develop qualifications and up-skilling opportunities



Outcome: Ensure appropriate qualifications, training, expertise, and access to tools exist for the workforce to review and evaluate AI usage in NRC-regulated activities effectively and efficiently

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

CREATE AI ECOSYSTEM TO PREPARE FOR INCOMING EXTERNAL AI USAGE

Goal 5. Pursue Use Cases to Build AI Foundation Across the NRC

AI Ecosystem

NRC Enterprise Data Warehouse

Example AI tools needed to learn, understand, test and evaluate external AI applications

python™

scikit learn

TensorFlow

- Focus is on preparing for incoming external AI usage
- Create an AI Ecosystem
  - Integrated development environments, common data science tools, and regulatory sand boxes are needed to conduct AI use cases
- Test Cases, Proofs-of-Concept or Pilot Studies
  - Learn, measure, and evaluate readiness to implement regulatory framework
  - Requires industry engagement and participation
- Continue Investment in AI Research
  - Future Focused Research and University Grant Program are crucial to developing and maintaining AI capabilities

Outcome: Develop an ecosystem that supports AI analysis, integration of emerging AI tools, and hands-on talent development for reviewing AI applications from the nuclear industry

U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

Status and Next Steps

- Finalize AI Strategic Plan by Spring 2023
- Host AI Technical Session at the Regulatory Information Conference in March 2023
- Host Data Science and AI Public Workshops in Spring/Summer 2023

BACKUP SLIDES

Acronyms

- AI - Artificial Intelligence
- AICoP - Artificial Intelligence Community of Practice
- AISC - Artificial Intelligence Steering Committee
- DOE – U.S. Department of Energy
- EPRI – Electric Power Research Institute
- FR - Federal Register
- FY - Fiscal Year
- GSA – General Services Administration
- ML – Machine Learning
- MoU – Memorandum of Understanding
- NRC – U.S. Nuclear Regulatory Commission
- OCIO – Office of the Chief Information Officer
- OCFO – Office of the Chief Financial Officer
- OEDO – Office of the Executive Director for Operations
- OIP – Office of International Programs
- NMSS – Office of Nuclear Material Safety and Safeguards
- NRR – Office of Nuclear Reactor Regulation
- NSIR – Office of Nuclear Security and Incident Response
- XAI – Explainable Artificial Intelligence

Public Meeting and Comment Period Summary

Draft AI Strategic Plan Public Meeting Summary

- ENABLE INNOVATION. The nuclear industry does not want regulatory oversight to stifle innovative AI uses
- ACCELERATE AI PLANNING. Significant Industry interest in the implementation plan
- INSTITUTE REGULATORY FRAMEWORK. Defining the regulatory oversight required at the various levels of autonomy is crucial

Preliminary NRC Observations on the Draft AI Strategic Plan Public Comments

- AGENCY APPROACH. The NRC needs a unified, agency-wide approach to harmonize AI research, use cases, governance, planning, and communication
- AI DEFINITIONS. The NRC needs to establish consistent AI terminology and definitions
- AI STATUS. Operating reactors do not appear to be planning safety-related AI applications in the next five years, and advanced reactors run the gamut of unqualified intentions
- AGENCY COMMUNICATION. The industry desires frequent and transparent engagement requests communicated metrics showing progress on Strategic Goal #1 and would like further details on how NRC may use or benefit from other industry or federal agency AI guidance
- AI LIFECYCLE. The NRC will need to consider AI life cycle maintenance, data availability, human-AI interaction in operation and user training, suitable test and evaluation methodologies, domain adaptation eschewing universal model usage, and model degradation

The NRC is proactively considering all the feedback from stakeholder engagement for implementation of the goals of the AI Strategic Plan

EPRI AI Projects

Key Example:
AI-Assisted Inspections

Thiago Seuaciuc-Osorio
Principal Technical Leader

Advisory Committee on Reactor Safeguards
Joint Digital I&C and Plant Operations Subcommittee Meeting
Wednesday, November 16, 2022

© 2022 Electric Power Research Institute, Inc. All rights reserved.

Assisted Analysis of UT Inspections

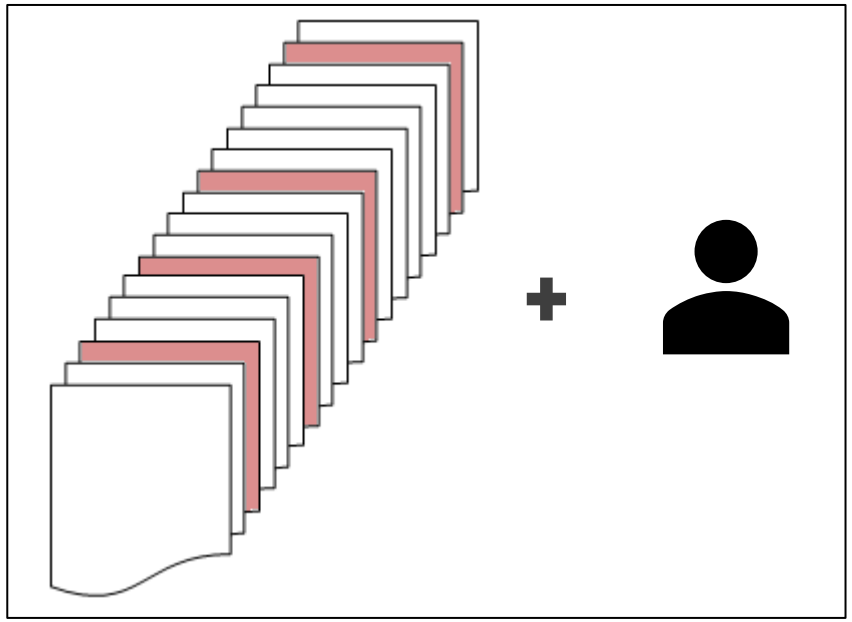
- UT inspections are an important part of the scope of an NDE program
- Some inspections are challenging or have large volumes of data
- Machine learning tools can potentially **assist** in the analysis of the data
  - Increase reliability
  - Decrease analysis time
- *Assist* means AI flags regions for review: final decisions still rest with the qualified inspector.

***Goal:* Develop auto-analysis tools to assist in UT inspections**

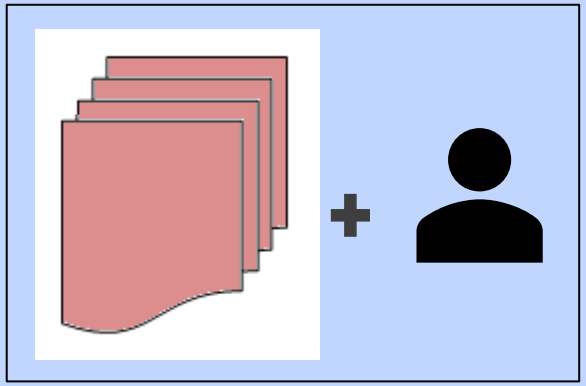
© 2022 Electric Power Research Institute, Inc. All rights reserved.

How Would AI Assist in UT Inspections?

- Current Inspection
  - Examiners distribute their energy across a high volume of (mostly benign) data



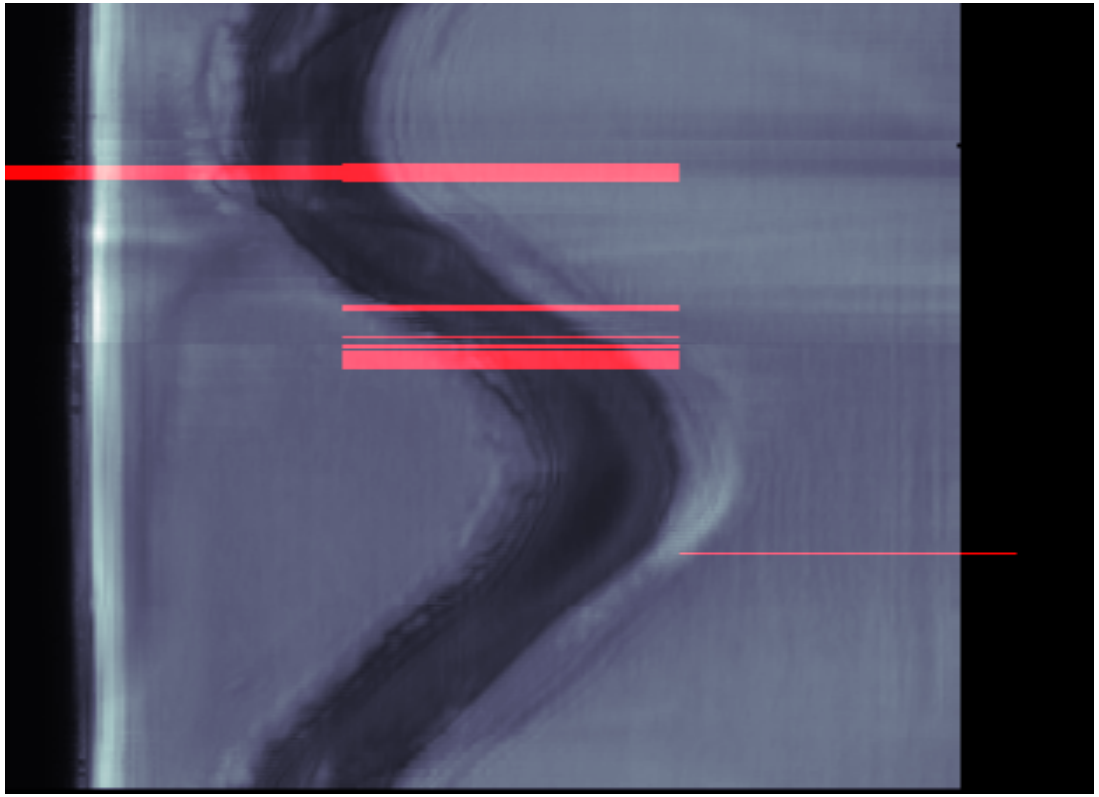
- AI Assisted Inspection
  - Examiners focus their energy on the regions that require more careful review



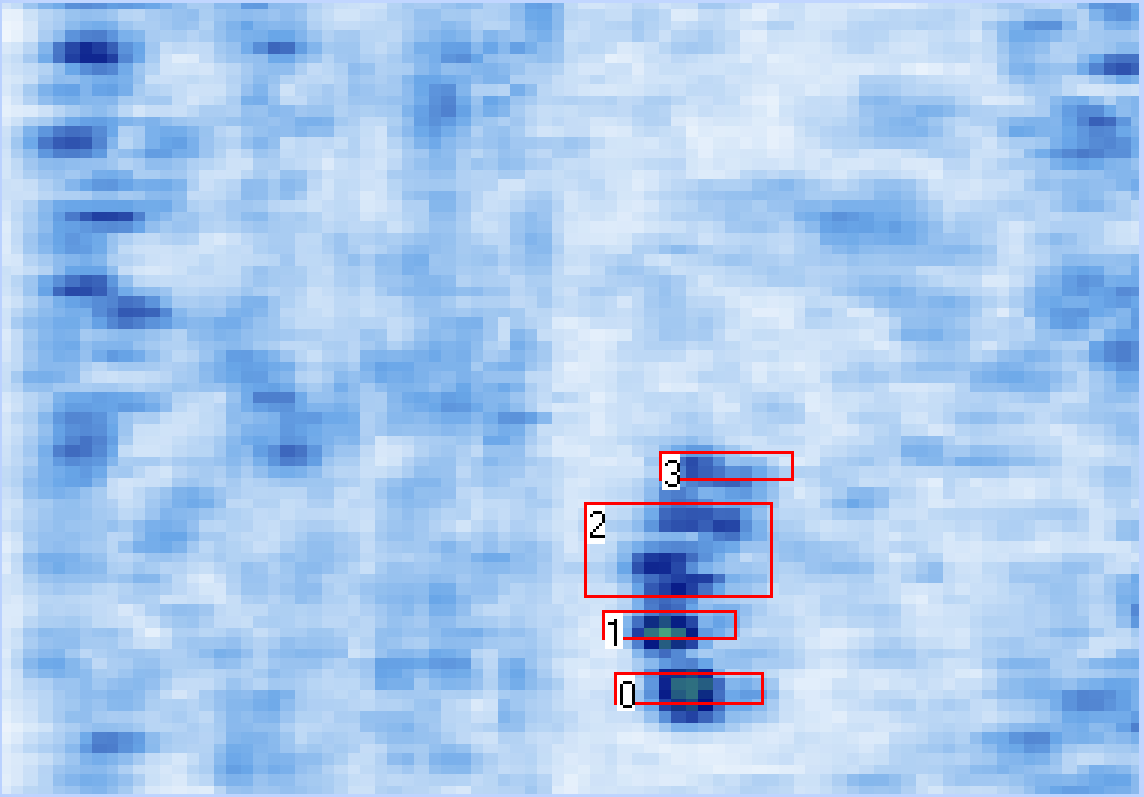
© 2022 Electric Power Research Institute, Inc. All rights reserved.

Two Ways to Utilize AI

- “Screening model” highlights areas of interest



- ”Indication model” pinpoints indications for evaluation



© 2022 Electric Power Research Institute, Inc. All rights reserved.

Envisioned Uses

Requires Qualification

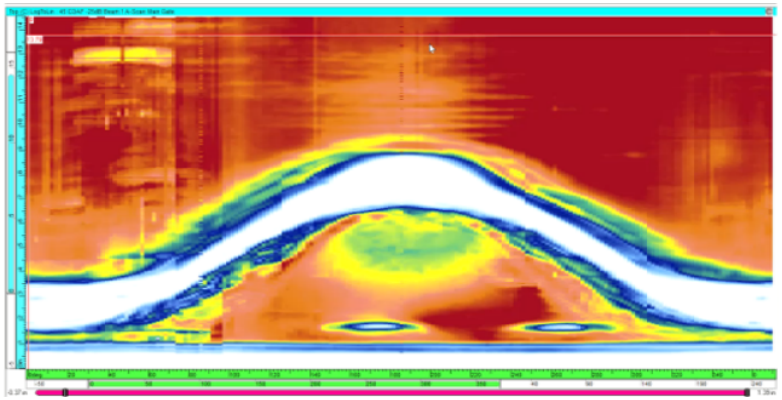
- Credited as part of inspection as a required (secondary) review

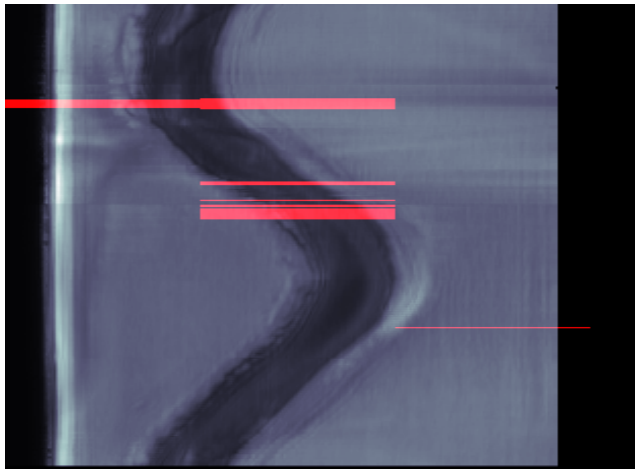
Without Qualification

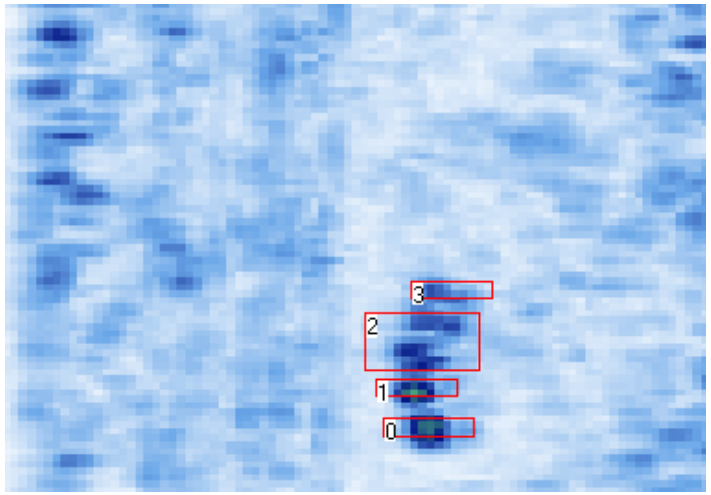
- To aid utility oversight
  - Utility staff data review (not credited as part of the mandatory inspection)
- To review previous outage data:
  - Prioritize inspection order (RVUH application)
  - Assess needed resources
  - Familiarize examiners with expected output/last examination results

© 2022 Electric Power Research Institute, Inc. All rights reserved.

Envisioned Implementation







1 UT data is saved in ML Box through local connection

2 Box automatically detects and processes UT data…

3 … and generates output for review

All data remains local; no SW needed

© 2022 Electric Power Research Institute, Inc. All rights reserved.

Spring 2022 Field Trial

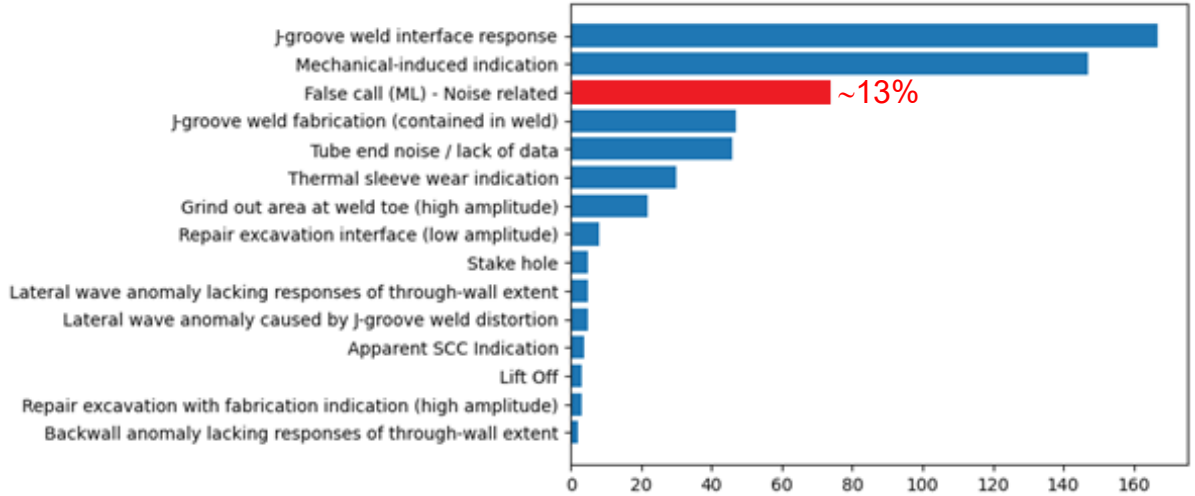
- 78 penetrations evaluated using AI
- AI flagged **~2%** for review
  - After onsite re-training led to 60% further reduction in amount of data for review
- Average of **~4min** per penetration
- All flagged regions were categorized
  - High weld interface disturbances & mechanically induced indications account for most calls (~56%)
  - Only **~13%** of flags due to noise
- The model successfully **flagged all SCC locations**

AI-assisted analysis yielded same results as the one from the qualified vendor

Amount of Data Requiring Review

Pre-AI		Post-AI	
4.4	miles	463	feet
7.0	km	141	meters

All values are approximate



© 2022 Electric Power Research Institute, Inc. All rights reserved.

Qualification Prospects

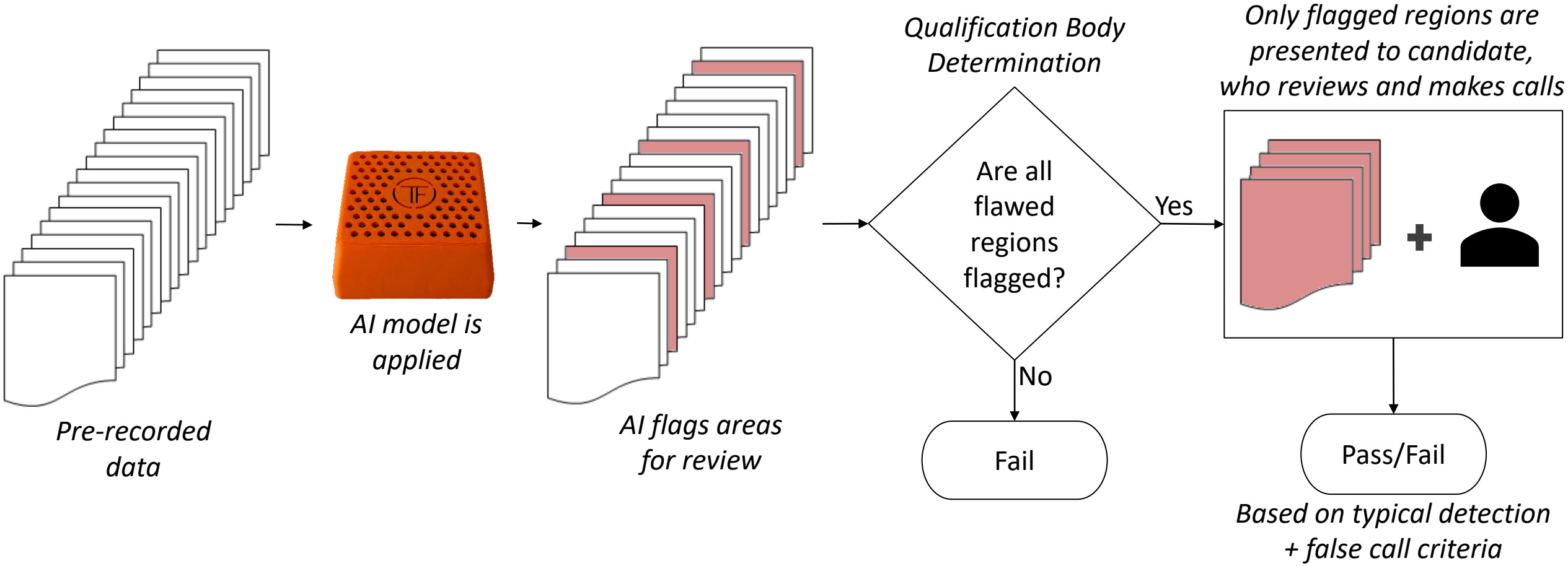
- *Leverages existing ASME qualification framework*
- Looking at this as a "procedure" type:
  - Model should flag **all flawed regions** for review by inspector in "personnel" qual
  - Inspector has opportunity to detect all flaws
- Model was exposed to more than **5x** the amount of PDI data required for a traditional procedure qualification
  - Model accurately flagged all flawed regions

All observations point to AI being qualifiable when paired with human analyst

© 2022 Electric Power Research Institute, Inc. All rights reserved.

What initial qualification *may* look like

- Procedure would be updated to include a defined process for the AI evaluation
- AI algorithms for qualification would be developed and provided to the qualification body: model is **frozen** at this stage



© 2022 Electric Power Research Institute, Inc. All rights reserved.

AI Applications across the Nuclear Industry

Insights

- Event Management Response Tool
- Relief Request Index
- Nuclear NLP Dictionary

Automation

- Automating analysis of NDE data
  - UT, Visual, eddy current
- Analog gauge reader
- Automating vibration analysis

Prognostics

- Plant historian data analysis
- FAC wear rate prediction

Optimization

- Decision logic for source term reduction
- Decommissioning planning

© 2022 Electric Power Research Institute, Inc. All rights reserved.

Together…Shaping the Future of Energy™

© 2022 Electric Power Research Institute, Inc. All rights reserved.

Constellation

Data-driven Decisions using Artificial Intelligence/Machine Learning

Advisory Committee on Reactor Safeguards

Joint Digital I&C and Plant Operations Subcommittee Meeting

November 16, 2022

Rick Szoch, Senior Manager, Nuclear Innovation

Gene Kelly, Senior Manager, Risk Management

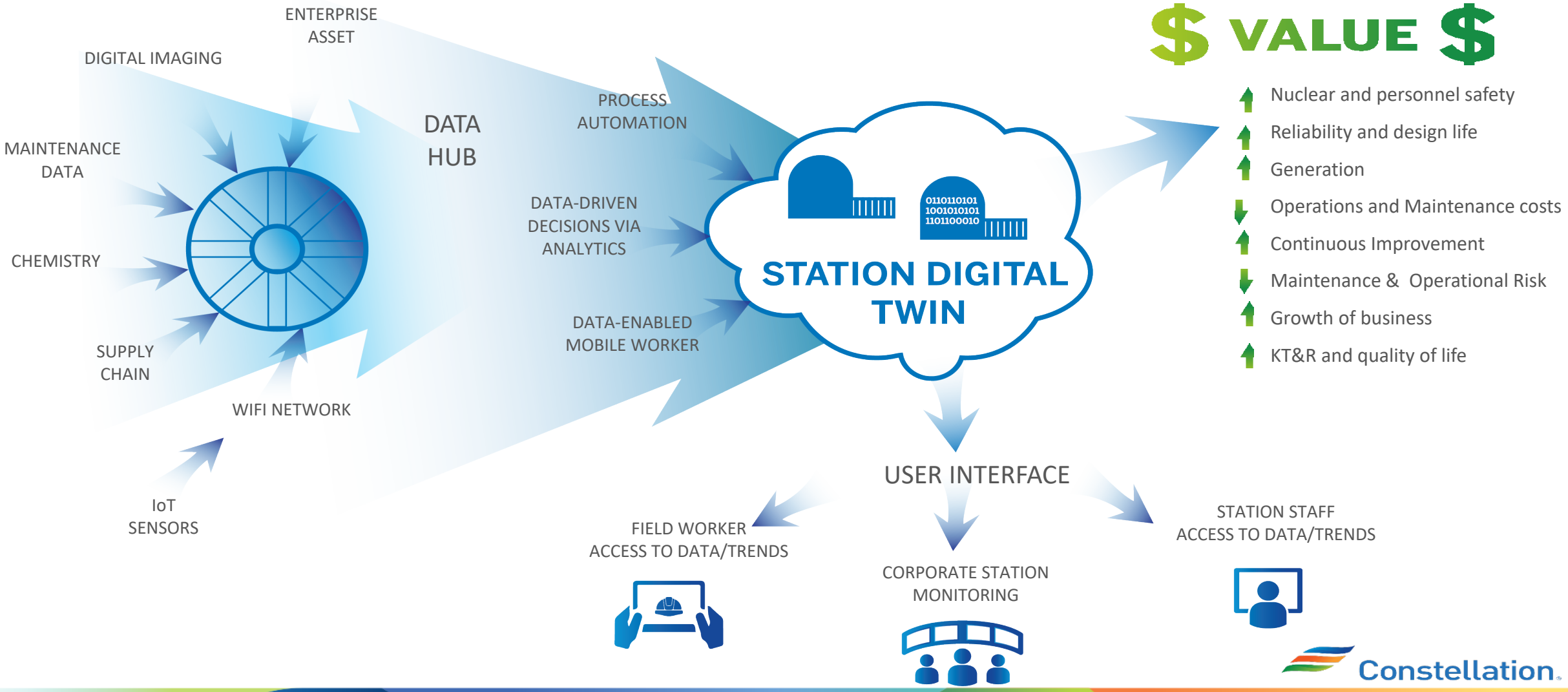
Constellation Energy Generation

Vision for a Digital Generation Business

Automated Work Processes	Streamlined workforce
Smart Processes	Electronic procedures/work instructions
Decision making through analytics/AI	Faster, more accurate decisions by right level
Field workers instant access to information	No time wasted, improved HU
In-field training through virtual reality and video	More wrench time
Predictive plant performance	Zero transients/trips
Centralized information backbone	Ability to leverage additional technologies
Innovative culture embedded	Workers continuously improving

Constellation.

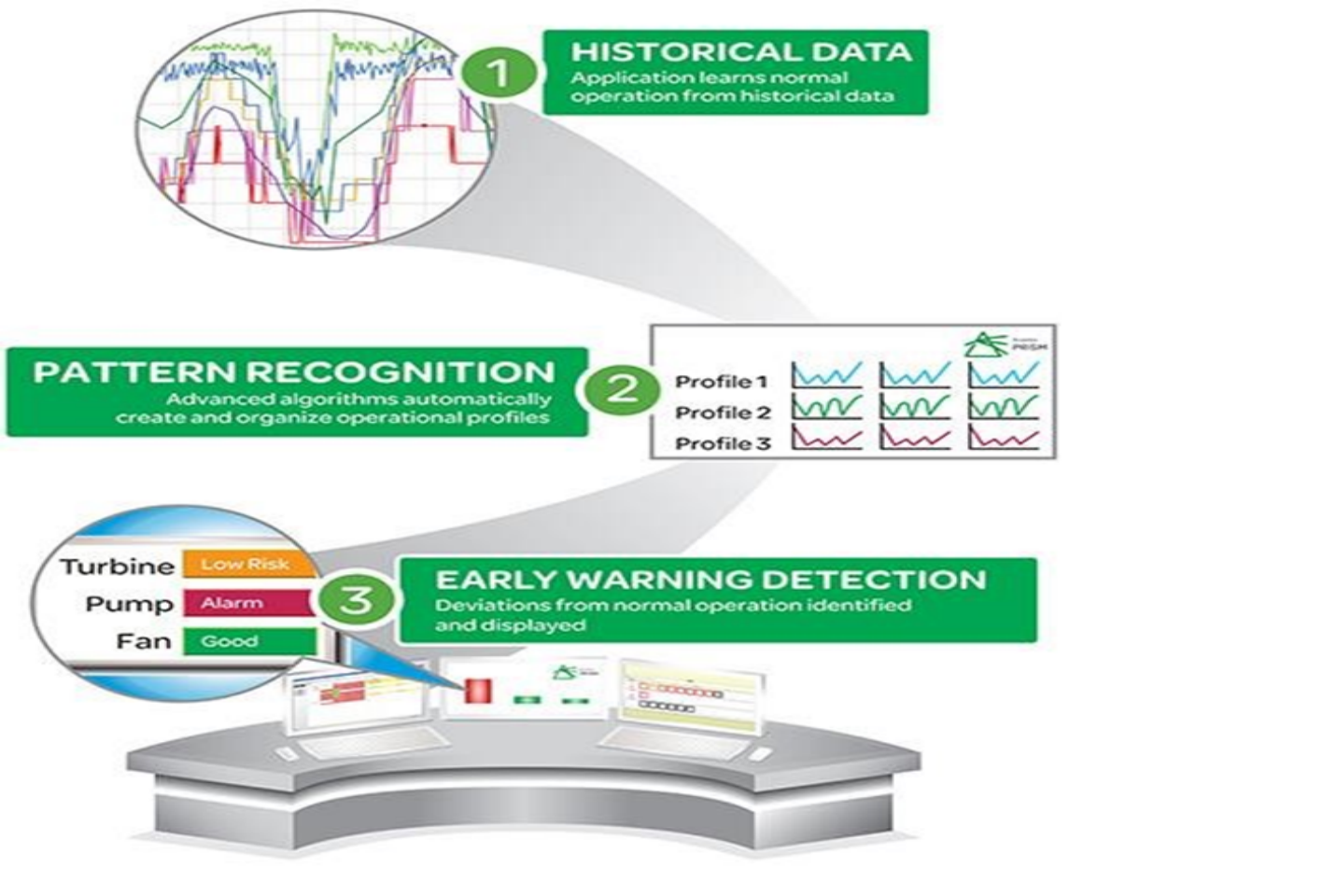
CONSTELLATION NUCLEAR DIGITAL TRANSFORMATION



Current Analytics Examples

- Advance Pattern Recognition (APR) – *averts equipment failures*
- Initial License Training analytic – *improves student pass rates*
- Maintenance Rule Functional Failure analytic – *streamlines categorization*
- New Work Screening analytic – *streamlines condition report categorization*
- License Extension reviews – *automates data collection*
- Preventive Maintenance Optimization – *optimizes maintenance frequencies*
- Work Order optimization – *automates work order process*

Equipment Reliability – Advanced Pattern Recognition to Avert Failure



Constellation

Initial License Training Student Performance

Constellation - DEV

HOME | ACTIONS | REPORTS | HELP

Exam Risks +

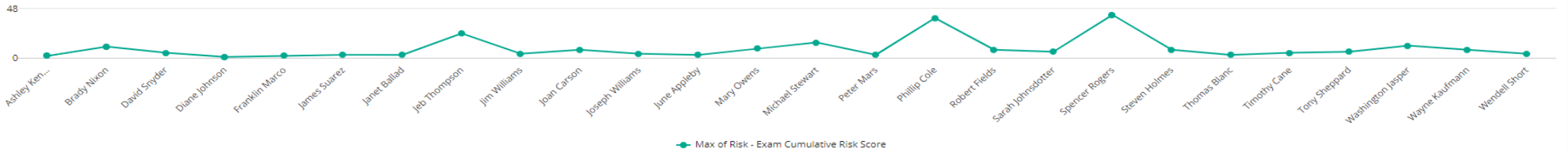
EDIT DASHBOARD

NO SAVED FILTERS

NO FILTERS APPLIED

SEARCH

Student Risks



Gaps Grid

Search Gaps Syncs | SEARCH

Gap - ID	Gap - Gaptopic	Gap - Examid	Gap - NRC Reference	Gap - Score	Gap - Station Reference	Gap - Student ID	Gap - Type	Person - First Name	Person - Last Name
1	High Pressure Coolant Injection	355363	201001.A1.07	5	DRE203LN001	5	Current	Louis	Armstrong
2	Turbine Trip	355363	201001.A1.08	4	DRE262LN007	5	Current	Louis	Armstrong
3	Recirculating Pump Trips	355363	201001.A1.09	4	DRE262LN003	5	Predicted	Louis	Armstrong

Constellation

Maintenance Rule Functional Failure Analyzer

- In place successfully for two years
- The software is *not* making failure determinations … it is pre-screening/flagging potential functional failure CR's for human review
- Biased towards high safety-significant component failures
- Equipment monitoring still occurs (i.e., human element backstopped)
- Confidence in results is gained through continuous feedback

Issue Report Screening Analytic

Displays the “why”

- Summarize data for user review
- Provide users detailed insights into how the predictions were made
- Indicate potentially similar IRs to the one being reviewed
- Orange means over the threshold for review; blue wouldn't trigger the need for a review

DataAdvisr

Home Analysis Results Admin Logout

IR 48 - CAQ

Select Model

Flagged for Review: Yes

DataAdvisr has identified this IR as a potential Significant Condition Adverse to Quality.

Last Processed On: 6-Sep-2022

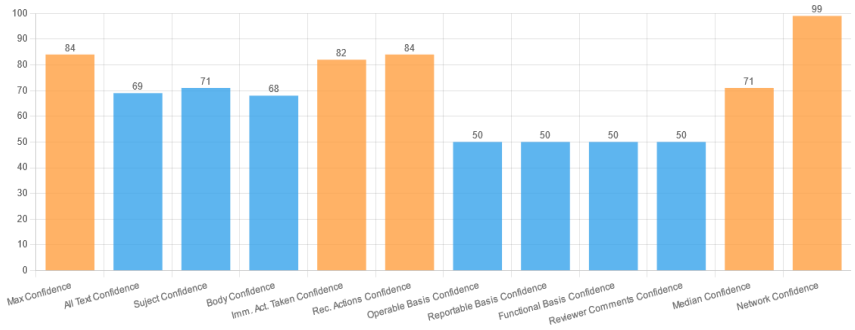
Location: Plant BWR1 | Unit: 2

Metrics Exceeded

1. At least one text field was highly indicative of the identified class.
2. The majority of text fields were more indicative of the identified class.
3. The combination of binary, numeric, and textual data had a high confidence.

Binary Field Triggers

Confidence Values



Highlighted Text Fields

Subject

Trip of 2B / 2C Circ Water Pumps

IR Body

At 05:15, the 2B Circ Water Pump tripped along with the 2C Circ Water Pump . Reactor power was immediately reduced by Reactor Operator due to condenser vacuum issues. Investigation of breaker shows no flags.

Body

Immediate Actions Taken

Recommended Actions

Operable Basis

Reportable Basis

Functional Basis

Phrases

Phrases Indicative that IR is CAQ

1. trip (8.55e-3)
2. water (3.94e-3)
3. pumps (4.05e-4)
4. shows no (3.79e-4)
5. water pump (3.71e-4)
6. circulation water (1.62e-4)
7. reactor power (1.48e-4)
8. pump tripped (1.09e-4)
9. circulation water pump (1.04e-4)
10. condenser vacuum (8.70e-5)

Phrases Indicative that IR is NOT CAQ

1. water (5.52e-3)
2. trip (1.29e-3)
3. shows no (5.39e-4)
4. water pump (3.50e-4)
5. pumps (2.99e-4)
6. circulation water (2.15e-4)
7. circulation water pump (1.21e-4)
8. circ (6.01e-5)

Constellation

Future Analytics

- **Predictive analytics – *Work Order generation***
- **Corrective Action Program – *Predictive trending to avert operational challenges***
- **Advanced equipment performance analytics – *avert equipment failure and operational risk***
- **Outage schedule predictions – *improve outage schedule adherence***
- **No intent to apply artificial intelligence to control the plant**

Summary

- Constellation continues to explore and deploy artificial intelligence (AI) that can improve or streamline business processes and data-driven decision-making
- Constellation considers AI to be useful as a decision support tool, not as a replacement for human judgment
- Constellation AI applications are designed to keep subject matter experts fully engaged in the decision-making process
- Current projects are giving Constellation experience with AI to inform future decisions on the potential value of AI in other areas