IDA Ideas

A podcast by the Institute for Defense Analyses

Episode 17

# Exploring AI and Machine Learning Capabilities

**Guests**: Arun S. Maiya

**Host:** Rhett A. Moeller

**June 2024**

IDA Product 3002620<br>
Distribution Statement A. Approved for public release; distribution is unlimited.

Institute for Defense Analyses<br>
730 East Glebe Road<br>
Alexandria, VA 22305

### About This Publication

The views, opinions and findings should not be construed as representing the official positions of the Department of Defense or the U.S. Government.

### For More Information

Arun S. Maiya, Information Technology and Systems Division
amaiya@ida.org, (703) 578-2790

### Copyright Notice

# Exploring AI and Machine Learning Capabilities

IDA Ideas guest Arun Maiya joins host Rhett Moeller to discuss artificial intelligence (AI) and machine learning: what they are, their capabilities and limitations, ongoing developments in technology and applicability in the defense industry. Arun is a researcher within the Information Technology and Systems Division (ITSD) of the Systems and Analyses Center, an IDA-operated federally funded research and development center (FFRDC). There, he has been heavily involved in projects related to AI and machine learning and developing techniques and software tools, including IDATA, ktrain and OnPrem.LLM.

**[Begin transcript]**

**Rhett Moeller**: Hello listeners. I'm Rhett Moeller and I'm the host of IDA Ideas, a podcast hosted by the Institute for Defense Analyses. You can find out more about us at www.ida.org. Welcome to another episode of IDA Ideas.

In this episode, we're going to focus on artificial intelligence and machine learning, a technology that has exploded into public consciousness in recent years thanks to last year's avalanche of easy-to-use generative AI tools. Although it may seem like last year was the beginning of AI, it's been developing quietly and used in numerous hard-to-spot ways for decades. Only we're now starting to see more public-facing implementations. There's a lot to cover, and a quick glance at the podcasting sphere shows that there's a surplus of shows that cover the ins and outs of the technology, but many of these presume an existing level of understanding that may not always be present. So today we're going to take a step back and consider some of the basics.

To give us this foundational look I am thrilled to introduce Arun Maiya, a researcher in IDA's Information Technology and Systems Division, or ITSD. He's been heavily involved with numerous data, natural language processing and computer vision projects and brings a wealth of experience to today's discussion. Arun, I'm excited to get this opportunity to talk with you about a subject that you're so familiar with. Can you tell us a bit about yourself?

**Arun Maiya**: Sure. Thanks, Rhett. I'm happy to be here. I'm a computer scientist in the Information Technology and Systems Division at IDA, and my work at IDA has generally involved advising government sponsors on matters related to AI, machine learning, data science and data analytics.

**Rhett**: Great. Well, welcome, and we have some good things to talk about. And really, one of the first things I wanted to focus on is it never hurts to understand exactly where you're coming from. So, we at IDA know that you're associated with a lot of open-source projects. Can you tell us some of the work that you've done that you're willing to discuss?

**Arun**: Sure. So, a lot of my work at IDA has taken the form of applied research and development where my team and I have developed techniques and software tools to address various problems. One example is IDATA, which is a document intelligence capability we designed to search and explore and analyze large collections of files, which has been applied to a wide range of sponsor problems internally at IDA on everything ranging from digital investigations of cyber compromises to policy analyses. IDATA is not open source, but we have developed other tools that we have made publicly available. One example is ktrain. So, ktrain is an open-source machine learning library designed to make machine learning and AI more accessible and easier to apply. You know, since being publicly released, ktrain has been applied to a wide range of use cases in industry, government and academia and has helped facilitate the application of machine learning to areas in which it previously hadn't really been applied much — areas like political science, psychology, communications and quantum software testing. Another example is OnPrem.LLM, which is a software toolkit that we developed to make it easier to run [OpenAI's] ChatGPT-like models on your own machines using non-public data.

**Rhett**: Great, but that sounds like a lot of work, a lot of experience. You mentioned ktrain, Arun, that's k-t-r-a-i-n?

**Arun**: Yeah, correct. And the package is available on GitHub.

**Rhett**: So, I think the basic question where we should really start our talk today is, what is artificial intelligence? What is machine learning? Is there a difference between them? What can you tell us about these core concepts?

**Arun**: So, you know, artificial intelligence is sort of a strange field for a multitude of different reasons. It's a completely different animal than pretty much any other critical technology of interest in the Department of Defense right now. And one of the ways you can see this is in the very way the field is defined. In most fields, the definition of the field is not typically a matter of debate. So, if you ask a quantum computing expert to define quantum computing, you know, it'll be a fairly uncontroversial answer. Not true of AI. There are heated and intense debates among experts [about] the ins and outs of what is considered and what is not considered AI. And, you know, what kinds of computational tasks require intelligence to solve. And part of the reason this happens is because artificial intelligence was never really strictly defined. The term was coined by a researcher named John McCarthy in the 1950s, and it's my understanding he coined the term specifically to get funding from the Department of Defense for his research. So, AI began its life as more of a buzzword, and to me, it's still a buzzword today. Now, if you look up the definition on Google or Wikipedia, you'll find a definition something like, "AI involves computer systems that can solve tasks commonly associated with human-level intelligence." But if you examine the definition you can begin to see some issues. So, you know, for one, intelligence itself is kind of a fuzzy concept.

**Rhett**: Right.

**Arun**: Moreover, you know, people tend to view human-level intelligence, and by extension, artificial intelligence, as something just beyond the frontiers of technology. So, you end up with these paradoxical scenarios where once someone solves a problem with artificial intelligence, people tend to not view the solution as artificial intelligence anymore. In fact, there was an article in The Washington Post a few weeks ago where they listed 10 technologies. And for each technology, they asked experts whether or not they considered it artificial intelligence. You know, interestingly, there wasn't a single technology where there was unanimous agreement on whether it was artificial intelligence, and the degree to which technologies are considered artificial intelligence tend to be correlated with how recent they are. So, there's very much kind of a moving goalpost in how we perceive and define AI. And by the way, these aren't just philosophical issues. They come up in very practical work, like horizon scanning of the latest artificial intelligence research, characterizing the AI workforce or characterizing AI education and training programs in the U.S., all of which are projects that IDA has worked on.

**Rhett**: What about machine learning?

**Arun**: So, machine learning is technically a subfield of artificial intelligence that focuses on systems that learn from data to perform tasks. Most AI today is powered by machine learning, so, for all practical purposes, they're kind of the same thing and they're used interchangeably. But machine learning is a comparatively more concrete concept than AI. So, one way to think about machine learning systems is they kind of look at data in the past, compute statistics about them and then use that to make predictions on new data in the future. And you know, by that broad description, you know, there are statistical techniques that have been around for maybe over a century that would be considered machine learning. There's an argument to be made that that's true. In fact, Craig Martel, the former [Chief Digital and Artificial Intelligence Officer] of the DOD [Department of Defense], once described machine learning as statistics at scale. You know, I would tend to agree with that description with one caveat. We understand a lot more about how statistics works than machine learning. You know, modern machine learning systems have billions of parameters. You know, we understand all the mathematical operations to calculate these parameters, and we even understand how to optimize and tune the settings of these parameters so they perform better. What we don't quite understand is how these billions of parameters collaboratively generate answers and predictions. And for these reasons, they've been called opaque, black boxes and inscrutable artifacts. And this lack of understanding of how they do what they do can pose challenges in applying, evaluating and regulating these kinds of systems.

**Rhett**: Right. So, Arun, you just mentioned some distinct challenges associated with artificial intelligence. Can you talk a little bit about some of the challenges surrounding development and using AI and ML [machine learning]?

**Arun**: Sure. So, there are a lot of challenges facing practitioners who are trying to apply machine learning. Before I talk about that, I wanted to say a few words about a different kind of challenge, one faced by decision-makers that sit at higher levels in the organization and make decisions on allocating resources to solve problems. And one thing I've been observing across different organizations is that it can sometimes be challenging for decision-makers to distinguish problems that are amenable to being solved by AI [from] problems that aren't. And one of the reasons for that is that it's not actually always that intuitive. There are problems in an organization that seem like you should be able to solve them with AI because they seem simpler than other problems that you already know are being solved by AI that you may read about in the press, and the answer all comes down to the data. If you want to solve a problem with machine learning, you need data that captures the expertise that you're trying to automate. So, for example, you know, one of the reasons we have really great face recognition models is because we've had lots of data to train face recognition models. There are lots of photos of celebrities and politicians on the internet, for example, that are captioned with who they are. There are other problems that seem like they should be simpler to solve than face recognition, and, for a long time we haven't had good models for those problems because, for one reason or another, we just haven't had good datasets for those problems. So, data is the lifeblood of machine learning. Because machine learning systems are data-driven and not engineered like traditional software, they also pose challenges when you're trying to apply these technologies to practical problems. One of the things you may have heard is that machine learning systems are brittle, and what's meant by that is that they tend to be somewhat sensitive to the data you train them with.

**Rhett**: Right.

**Arun**: So, for example, if there's bias or imbalance or incompleteness in your dataset or it's not representative of the thing you're trying to learn fully, this can cause problems in your model. For example, face recognition models often perform less well on people of color and women. And why? Because a lot of face datasets are biased towards white males. Machine learning systems also tend to be vulnerable to so-called adversarial inputs. So, for example, it's possible to manipulate inputs in such a way that the machine learning model will be fooled into making the wrong prediction—

**Rhett**: Right.

**Arun**: — in ways that humans would never be fooled. And part of the reason these kinds of things happen is because machine learning systems are essentially just huge correlation engines that are really great at recognizing patterns, and this ability to recognize patterns is a double-edged sword. On the one hand, it allows these systems to make great predictions and perform much better than older forms of pattern recognition and statistical techniques. On the other hand, sometimes your machine learning model will detect a subtle pattern buried in your dataset that you wouldn't actually want it to use when making predictions. And yet you won't even know that it learned these patterns because, um, these models tend to be more opaque. In fact, sometimes you don't realize the problem until the model has already been deployed, and it encounters an input that triggers one of these undesirable patterns —

Rhett: Exactly.

Arun: — into making a wrong prediction. You know this. We've seen this happen a lot with big tech companies where they'll release the model and then after deployment, it'll do something offensive or racist. So let me provide a more illustrative example of this. So, there are these medical researchers that were training a model, a computer vision model, to analyze medical images, to detect severe cases of some medical condition. And in the lab, in their experiments, they're getting great results, but when they deployed the model, it started making a lot of wrong predictions. And when they examine the model, they noticed that it was focusing attention on a serial number at the bottom of the medical images. And at first this was really confusing because the serial number was just a random string of letters and numbers until they realized the font of the serial number varied based on the hospital the medical image was from, and because all the severe cases were being referred to one particular large hospital during training, the model figured out they could use the font to predict whether or not a case was severe.

**Rhett**: Interesting.

**Arun**: Now this relationship between the font and the severity kind of fell apart in the actual operational environment, which is why it made wrong predictions. So, the lesson learned here is, number one: [if] your machine learning model can find a way to cheat during training, it probably will. And number two: it's really important to assess machine learning systems in a realistic setting within the larger context of the goal you're trying to accomplish. And one of the implications of this is that if you're pairing a machine learning model with a human to help the human make decisions, it's important to test them together because testing the machine learning model in isolation may not tell the whole story. When you pair a machine learning model or an AI with humans, you know, they can interact in weird and unexpected ways. IDA has actually done some work on establishing frameworks to test and evaluate systems where humans are paired with machines. But there's still a lot we don't know about this interaction, and more work is needed on the human-computer interaction aspects of artificial intelligence.

**Rhett**: Yeah, and given the seriousness of some of these applications, like medical diagnoses, it's critical to understand why it's important that these systems can explain themselves, that we can understand where they're getting the information from. That's excellent insight.

**Arun**: And by the way, the work doesn't end when you deploy the model. You have to continuously monitor the model. Why? Because if the data you're making predictions on doesn't look like the data the machine learning model was trained on, then the model will make wrong predictions. In general, if you talk to members of the test and evaluation community here at IDA, they would probably tell you that it's really important to establish a set of processes that span the entire life cycle of machine learning development, from curating and collecting data to training, evaluating and deploying models that will help you build a body of evidence that will allow you, your users and stakeholders to say, "Yes, I've justified confidence that this model is doing the right thing and not doing the wrong thing and being used properly by the human," to loosely quote a paper on trustworthy artificial intelligence authored by IDA researcher, David Tate, one of my colleagues.

**Rhett**: At the beginning of this discussion we mentioned generative AI, and obviously that's a big phrase being thrown around these days with ChatGPT and other major models that are making the rounds. Can you describe what generative AI is, and how does it differ from other forms of AI?

**Arun**: Sure. So, most real-world applications of AI (I think this is true even after the release of ChatGPT) involve a form of AI that might be referred to as decisional AI. So decisional AI models output decisions in the form of a category or a number, usually a category. So, for example, a fraud-detection model will look at a financial transaction and output fraud or not fraud. So, two categories. A face recognition model will identify a person in a photo where the number of categories might be the total number of people that the model is capable of recognizing. By contrast, generative AI, instead of outputting a predicted category, outputs new data that has similar characteristics to the data it was trained on. Although this is very different, under the hood, the way generative AI models work is not that dissimilar to decision AI models. So, for example, whereas a face recognition model will predict the identity of a person in the photo based on the visual features of the photo, or the face in the photo, generative AI models like ChatGPT will predict the identity of the next word in the response based on the words that came before it. So, it's conceptually somewhat identical.

One of the key differences, though, is that generative AI models like ChatGPT are autoregressive, which is a statistical term, but that basically means that their decisions are based on previous decisions it makes. So, for example, when you ask a question to ChatGPT, it uses the question to generate the first word in the response and then uses the first word in the response to generate the second word, and so on and so on. And it's because of this that the way you solve problems with generative AI is completely different than the way you solve problems with decisional AI. With decisional AI you basically have to set up your problem as a categorization task. So, if you're building a fraud-detection model, you basically bin all financial transactions as either fraud or not fraud, and your model will only ever output one of those two decisions. In generative AI you basically set up your problem as what is essentially a fill-in-the-blank task. One of the interesting consequences of this is that, unlike decisional AI models that can only ever solve the specific categorization problem that you trained it to solve, a single generative AI model can solve a wide range of problems. The only requirement is that you have to somehow represent the input as a fill-in-the-blank problem. And as it turns out, a lot of problems can be framed as fill-in-the-blank problems: answering questions, translating a sentence, you know, summarizing a document, generating code based on the description of what you want the code to do, and so on and so on.

Incidentally, generative AI models can also solve decisional problems, so if you want to categorize a product review as positive or negative, you provide the product review, you provide maybe an example of a positive and negative review, and then you say this product review is blank and it'll fill in the decision. I guess, more interestingly, generative AI models can solve problems that can't really easily be solved by decisional AI models, like recommending a course of action or a strategy based on a described problem scenario using all the historical data the model was trained on, or generating synthetic data for use in a simulation or a wargaming exercise. So, these models are, you know, flexible and really powerful and have lots of potential. But they also come with more challenges. So, for example, they're really computationally expensive to train and run. Their outputs are more complex and therefore more unwieldy, which makes them harder to correct and control as compared to decisional AI models.

**Rhett**: Right.

**Arun**: And sometimes getting them to work is more trial and error than engineering and more art than science. Despite these challenges, these models are flexible and have lots of potential, and it's for these reasons that the DOD has created a department-wide initiative called Task Force Lima to examine the potentials of generative AI and develop recommendations for how these models can be used in a responsible way. And as an FFRDC, IDA has been providing support to this initiative in various ways.

**Rhett**: So, we have seen more and more presence of artificial intelligence in reporting and new products being released, some products that claim they have AI and maybe don't. But it's obviously something that's very much in the consciousness of the American public and beyond. What do you think is driving that elevated, heightened awareness of AI?

**Arun**: So, you know, another really strange thing about AI is that if you look at the history of AI in this past, maybe 15 to 20 years, a lot of the catalysts and triggers for making AI successful or effective or popular or disruptive or game-changing have often been things on the periphery that are not actually related to the latest AI research. So, for example, most state-of-the-art machine learning models right now are powered by artificial neural networks, a particular form of machine learning. And a lot of what makes modern neural networks tick have been around for a while, in some cases since the 1980s and since the 1990s. So why were those techniques not effective then? Why are they only effective more recently? One reason is data. So, in the [early] 2000s, we saw a surge of digitized data on the internet. So, people were uploading photos and images to blogs, which can then be used to train computer vision models. We saw the creation of websites like Wikipedia that contained an abundance of high-quality text, which would later go on to be help train systems like ChatGPT. Now, when you train a large neural network and a large dataset, things can be really slow. Another thing people figured out in the 2000s is that you could use graphical processing units, or GPUs, which are these devices typically used to render graphics and video games and use them to train neural networks much more quickly than would otherwise be possible. The reason for that is because the math to train neural networks is the same as the math to render graphics in video games. When people started using GPUs to train neural networks on large datasets, they discovered that these techniques that previously hadn't worked very well started to work really well to the point where they're shattering previous benchmarks, and we're seeing that continue to happen even through the present day.

**Rhett**: And that's why companies like NVIDIA are enjoying the success that they are today.

**Arun**: Yeah, exactly. So, it's really, you know, been data and hardware that have been critical drivers of success in AI, and that also is something you can even see recently. There's also a third factor that I think has been critical to the success of machine learning that isn't often talked about, so, I wanted to briefly mention it, and that's openness. So, machine learning as a field has had a long tradition, maybe, say, past 15 [to] 20 years of open science. When a machine learning researcher submits a paper to a conference or a journal, they often make the code for their technique publicly available. If they use a dataset, they sometimes make that dataset available to facilitate reproducibility. If they train a model on a large dataset, they take the trained weights or parameters for that model and make them available so that other people can take that model and apply them to different problems. So, for example, if someone trained a model on millions of photos, through the use of something called transfer learning, I can take that model or the weights of that model and then fine-tune it on a much smaller set of satellite images than I would otherwise need and get a pretty, really good model to analyze satellite images. So, I think this sharing of code and data and model weights have, you know, supercharged the quick exchange of ideas in the field and has also fueled a lot of the advancement and innovative applications that we're seeing.

This is beginning to change. So, for example, OpenAI has not been transparent on the technical details of ChatGPT, and you know, when they publish research papers, they come under some criticism because they're less scientific and more like press releases. So, at the same time, they've made these systems available for a fee to the public for potential use and for malicious applications. And on the other hand, you know, Microsoft recently released, published, a paper on a model that can take a still photo and audio and generate video. And they purposely have not made that model publicly available because of the potential for misuse.

**Rhett**: Right.

**Arun**: I guess my point in all of this is that I think the machine learning and AI community, and society at large, is kind of grappling right now on how to balance open science and AI safety. And, you know, I think there's a lot we still need to figure out in terms of how to govern these AI models and systems that we have.

**Rhett**: Well, given that, and, yeah, I've been impressed with the advances where you can take a picture and pair it with a little bit of recorded voice text and generate a fully video-rendered face speaking realistically. Amazing stuff. Given the things that you have seen and where you have seen artificial intelligence in the past years, where do you see it heading? What trends do you foresee for this field?

**Arun**: Well, you know, even the best AI systems we have today have a lot of problems. So, for example, they require a ridiculous amount of data even to learn simple concepts. Meanwhile, you know a 10-year-old can learn to clear the table with a single demonstration, a high school student can learn to drive with only 20 hours of instruction, both of which are impossible feats for even the best AI we have today. The systems we have today that are most representative of what people think of when they say artificial intelligence are systems like ChatGPT. And if you look at what ChatGPT is doing, it's literally just predicting the next word based on the words that came before it. So, it's, you know, a kind of a glorified version of the autocorrect feature on our phones. And although they store a lot of factual knowledge, they don't really have a true understanding or a true world model of the complex physical and social relationships among entities, objects and people in the world. And this is actually fairly easy to demonstrate, even in the original ChatGPT.

They also entangle language and factual knowledge, whereas cognitive scientists will probably tell you that they're more separated, and they're also missing a lot of other things that we have, like formal reasoning, ability to plan, episodic memory and so on. So as a result of these limitations there are a lot of prominent, or several prominent, researchers that are putting out high-level proposals on entirely new architectures that we should be focusing on. But these high-level proposals are just that — high-level. To my knowledge, no one in the world knows how to practically implement these proposals in a way to solve practical problems. So, most of the community right now are focusing on lower-hanging fruit like multimodal models that can understand language, audio and video or embodied intelligence, which is to take AI and put it in agents or robots in such a way that these agents can explore and learn about the world in a manner that's more similar to the way we do, which some feel is a necessary ingredient to achieve the kind of intelligence that we see in nature and biology.

There's also a lot of interest right now in synthetic data, and it is somewhat a controversial area because there are papers that have shown that this can go horribly wrong. So, for example, if you take a generative AI model and use it to generate data and then go back and train the model on that data in kind of a loop —

**Rhett**: I see.

**Arun**: — the model will eventually go nuts. And, you know, biases in the original dataset get amplified with each generation until the model outputs data that doesn't really reflect reality anymore. However, there are more surgical, sophisticated ways to use synthetic data. So, for example, if you use generative AI to generate an answer, you can actually take that model and have it inspect the answer to detect low-quality or incorrect answers, and improve the answer. And that seems counterintuitive, because if the model can detect incorrect answers, why didn't it output the correct answer to begin with?

**Rhett**: Right, right. Good question.

**Arun**: And the reason for that is because generating data and going back and observing and making a decision about the data are two entirely different problems to the model, the latter of which is easier for the model. Companies like Anthropic have been leveraging this with great success. Synthetic data has also been used in medical imaging, so, you know medical researchers, for example, have generated synthetic medical images and then showed that when you train a computer vision model on both synthetic medical images and real images, it can outperform models trained only on the real medical images. Another way you can use synthetic data is to generate data from a large, capable model like GPT-4 from OpenAI, and then take a smaller model and then fine-tune that model on this generated synthetic data. And there are a lot of Y Combinator-funded start-ups that are doing this very successfully and building small, but capable models that actually rival the performance of these larger models on this specific task. In general, there's a trend right now in doing more with less. So, the original ChatGPT was 175 billion parameters. Fast-forward a year later, we now have much smaller models (that are also relatively more open) that you can run on a single machine. In fact, just last week I believe, Microsoft released an even smaller model that they claim is suitable to run on a mobile phone. So, I mentioned this earlier, but at IDA, we've developed a software package called OnPrem.LLM, which is designed to help run ChatGPT-like models on your own machines using possibly sensitive data. And one of the motivations for that work was to flexibly apply generative

AI in small, isolated air-gapped networks where you might be dealing with sensitive data and you might have machines that are more resource-constrained. You know, that work has only been possible because of the availability of these smaller, relatively more open models.

**Rhett**: Great. Sounds like a lot to look forward to. A lot of changes coming in a very short amount of time. You're talking major slimming in the course of a year, and who knows what the next year is gonna bring as more and more groups build on each other's work to see this entire technological field expand in new and unexpected ways.

Arun, thank you very much for sharing your time with us. I think you've given us a lot to think about and a lot of information at a very approachable level. So, thank you for sharing that with us and providing us a framework to get the necessary understanding of the basics of artificial intelligence and machine learning and the distinctions between the two. I think it's been most illuminating.

**Arun**: Thanks, Rhett. This was a lot of fun.

**Rhett**: As always, if you want more information on IDA and its ongoing work, please check us out at ida.org. We also have a presence on X [formerly known as Twitter] @IDA_org, and we have a channel on YouTube. You can find direct links to all of our online presences in this episode's show notes. The show is hosted by the Institute for Defense Analyses, a nonprofit organization based in the Washington, D.C., area. Once more, you can find out more about us and the work we do at IDA.org. Thank you for tuning in, and we hope you'll join us again next time as we discuss another big idea here at IDA Ideas.

## Show Notes

Find us on X, on YouTube and at ida.org. Learn more about the topics discussed in this episode via the links below.

Albert, Michelle G., Katharine J. Burton, Forrest R. Frank, Arun S. Maiya, Daniel Y. Nakada, James OM. O'Connor, Laura A. Odell, Robert M. Rolfe, Miranda G. Seitz-McLeese, Thi Uyen Tran, Anna Vasilyeva, Dale Visser, Andrew Wan. "IDA Research Notes – IDA Text Analytics." October 2018. https://www.ida.org/-/media/feature/publications/i/id/ida-research-notes---ida-text-analytics/idataresearchnotes.ashx

Lillard, Vincent A., Leonard D. Wilkins, Caitlan A. Fealing and John T. Haman. "DATAWorks 2022: Topological Modeling of Human-Machine Teams." IDA Document D-33031. April 2022. ida.org/research-and-publications/publications/all/d/da/dataworks-2022-topological-modeling-of-human-machine-teams.

Maiya, Arun S. "Eye on IDA: Arun Maiya on Machine Learning and Artificial Intelligence." November 2023. https://www.youtube.com/watch?v=fVmndWp_s74.

Maiya, Arun S. “Democratizing Deep Learning with the ktrain Library.” April 2020. https://www.ida.org/-/media/feature/publications/d/de/democratizing-deep-learning-with-the-ktrain-library/d-10559.ashx.

Maiya, Arun S. “ktrain: A Low-Code Library for Augmented Machine Learning.” IDA Document D-28847. October 2021. https://www.ida.org/-/media/feature/publications/k/kt/ktrain-a-low-code-library-for-augmented-machine-learning/d-28847.ashx.

Maiya, Arun S. “On-Premises Generative AI.” IDA Product 3000898. November 2023. https://www.ida.org/-/media/feature/publications/o/on/on-premises-generative-ai/on-premises-generative-ai.ashx.

Tate, David M. “Trust, Trustworthiness, and Assurance of AI and Autonomy.” IDA Document D-22631. April 2021. https://www.ida.org/-/media/feature/publications/t/tr/trust-trustworthiness-and-assurance-of-ai-and-autonomy/d-22631.ashx.

Wojton, Heather M., Brian D. Vickers, Kristina A. Carter, David A. Sparrow, Leonard D. Wilkins, Caitlan A. Fealing. “DATAWorks 2021: Characterizing Human-Machine Teaming metrics for Test and Evaluation.” IDA Document 21564. April 2021. https://www.ida.org/-/media/feature/publications/d/da/dataworks-2021-characterizing-human-machine-teaming-metrics-for-test-and-evaluation/d-21564.ashx.