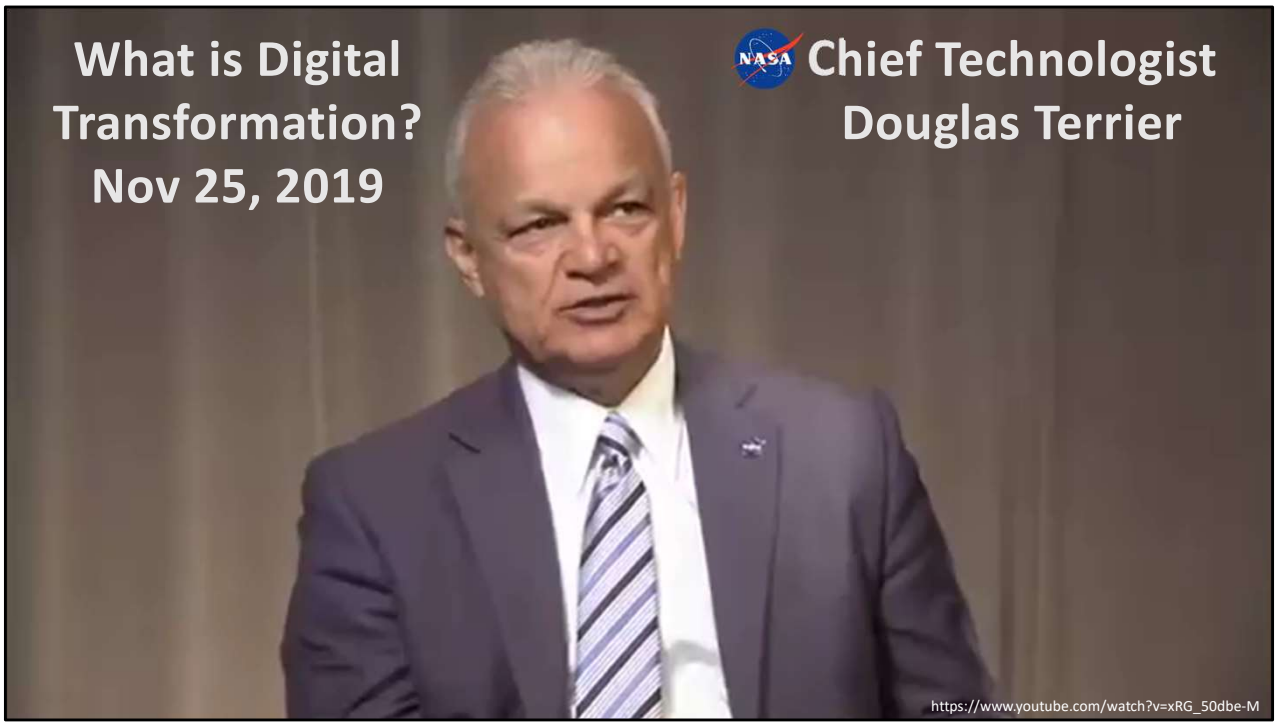


A YouTube video pulled from https://www.youtube.com/watch?v=xRG_50dbe-M

The video shows CTO Doug Terrier giving his definition of what Digital Transformation means for the agency. The video is set to play between 00:29 and 00:59.

That was Douglas Terrier, our Chief Technologist at NASA. And my name is Newton Campbell. Through SAIC, I serve under the OCIO Data Science Team at NASA Langley Research Center, where we combine our Digital Transformation thrust at NASA with artificial intelligence and machine learning. By doing this, we help scientists, engineers, as well as people in the enterprise, take on new technologies that are rooted in AI. Our work ranges from using AI for intelligent urban air aviation to detecting cancer-causing space radiation to developing intelligent dashboards that help the enterprise with decision-making. And as our group lead is helping to frame the AI Ethics strategy for the entire agency, we have the opportunity to be out in front with establishing processes and methods for advancing understanding of concrete ethical use on ongoing projects.

Digital Transformation at NASA

"One giant leap for technology"

arm DevSummit

Key Goals

- Accelerated technical and engineering innovation
- Increased efficiency and effectiveness of business processes
- Efficient, reliable, and safe mission systems and missions
- Real-time, data-driven decision making
- Agile workforce, facilities, and IT infrastructure
- Integrated collaboration and partnerships
- Advancement of exploration, discovery, and science
- Extended aerospace leadership

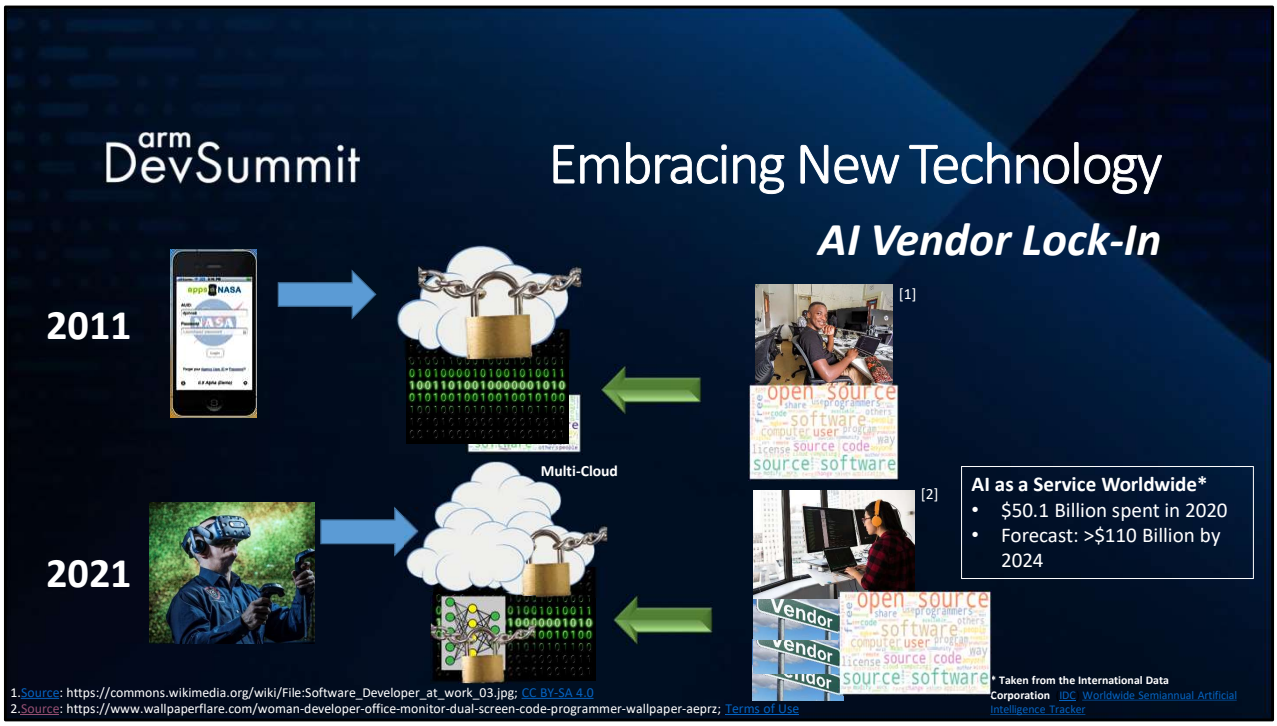
Key Technologies

- Cloud Computing
- Automation and Robotics
- Big Data/Data Mining/Analytics
- Artificial Intelligence (AI) and Machine Learning
- Agile software development/DevOps/DevSecOps
- Internet of Things (IoT)
- Integrated multidisciplinary modeling and simulation
- AR/VR/XR
- High-Performance Computing
- Mobile Access
- Collaboration
- Social Media

Currently, the US Government (USG) is going through a large-scale data Transformation effort. The GSA playbook is guiding all agencies to make their processes data-driven with the help of emerging technologies. At the same time, NASA leadership wants to share best practices in digital solutions, avoid duplication and gaps, ensure interoperability, and encourage cooperation among all stakeholders. Therefore, in the spring of 2019, NASA leadership approved an agency digital transformation strategy, with the Office of Chief Technologist leading coordination in partnership with the Office of Chief Information Officer. Key elements of this strategy, this plan, include digitally transforming NASA's data use, collaboration, model-based work, administrative processes, application of artificial intelligence, and workforce and culture.

Digital transformation will increasingly change the way NASA operates and enable the agency's missions to be completed more efficiently and effectively. And concerning this talk, allow future employees to use their talents in more innovative ways. Under NASA's Digital Transformation Officer, Jill Marlowe, the agency is making investments in nearly every sector to unlock any potential technology innovation: this comes through money put towards tech challenges, sponsorship of small technology programs to vet capabilities, and real strategic assessments regarding the use of AI

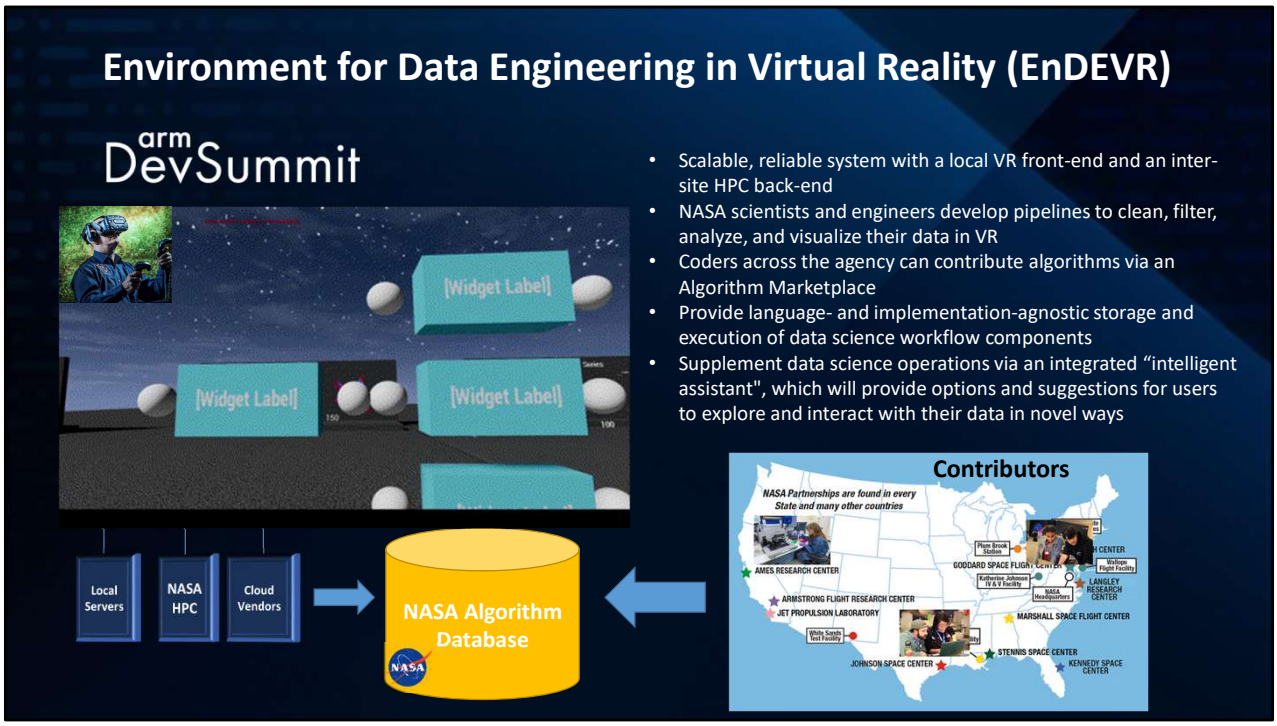
across the agency.



One issue that we face in digital transformation, particularly as it relates to integrating new advances in AI, is that the investment of time and effort in obtaining product licenses, support, training, and other product-specific services can unwittingly tie some of our projects to product vendors for the duration of a project. This is a concept known as “vendor lock-in. And its something that I saw a lot of Government agencies go through 10-15 years ago, when everyone started talking more about Big Data and the cloud, and I'm now seeing go through this with Diverse data when it comes to AI. Okay? When it came to the cloud, vendors were all about "We can take all of your data, as is, slam it into our system, and give you all the compute functionality that you could ever want. No maintenance or troubleshooting of platforms needed." Fast forward, ten years later, its now, "We can take all of the data that you have stored on that cloud instance that you're still paying a support fee for, intelligently fuse all of the other data that you've been trying to store for the last ten years (that’s still not structured), and use AI/ML to make sense of it."

In either case, the common theme is for the vendor to hang on as long as possible. Makes a little more sense in the cloud case, because, well, you are hosting data or a service on their platform. However, in the case of AI, you are tying the LOGIC of your project, not just the workflow, not just the platform, but the LOGIC of your project to

an organization with inherently limited capability. With SO many of these fairly new AI capabilities, you risk coupling your own software implementations tightly to a vendor's API or service. And suppose you fail to make incredibly nuanced purchasing and system design decisions. In that case, you can easily incur high operating costs for AI toolkits of vendors large and small, while simultaneously limiting your capacity for data science to the capabilities of that vendor. And those effects can be more impactful in organizations that don't have a trained data science staff on-hand to sufficiently augment those vendor capabilities. And that is just an impossibility for many organizations.



So with respect to these themes, we are building EnDEVR, the Environment for Data Engineering in Virtual Reality (we love our acronyms in the Government). This is funded under the ACT II Digital Transformation effort, local to NASA Langley Research Center in Virginia. The goal of EnDEVR is to develop an applied mathematics and data science ecosystem that allows users to command and investigate customizable data analyses from the virtual reality (VR) environment. A key feature of the EnDEVR ecosystem is that it is not dependent on any one proprietary data science platform, allowing users to extend its data analysis capabilities without restriction. EnDEVR permits a user to load and quickly preprocess data, develop a data processing pipeline from a VR environment, and kickoff computation of that pipeline on a back-end computing server. It then allows the results of this pipeline to be explored using existing VR capabilities. The user constructs a data pipeline, using hand motions, by selecting and organizing a set of data processing widgets (e.g. cleaning, time-alignment, interpolation, PCA, neural net training) provided in the EnDEVR Algorithm Marketplace. The Agency user community will provide algorithms to the Marketplace to grow it over time, adapt to new open-source or proprietary data analysis libraries, and inspire collaboration across Centers.

The full system is still under development and test at NASA. To support standard

NASA research, the full implementation of the environment will also contain:

- The ability to quickly pull code/scripts from a local machine, StackOverflow, or GitHub and incorporate it as an atomized function in the ecosystem
- The ability to load other data science software (RStudio, MATLAB, etc.) as 2D screens in this 3D environment, and the ability to quickly transfer data/results between environments
- The ability to quickly ingest data and run initial analyses based on user preferences
- An “Intelligent Assistant” to recommend algorithms based on mathematical transforms of the data

So much of the AI on-board EnDEVR is being designed an implemented by our development team this Fall. However, as this system is implemented, we as NASA community members are trying to be aware of the ethical implications of building such a system. One of the first of many studies that we looked to, prior to NASA's own AI Ethics Framework coming out, was ODNI's (Office of the Director of National Intelligence's) framework for ethics. And in asking ourselves some of these questions, we realized we had to focus an ethical analysis on certain components and features of the EnDEVR system.

The two primary AI-focused components of the EnDEVR system are the Intelligent Assistant and the storage and execution of arbitrary algorithms.
The EnDEVR Intelligent Assistant is a form of AI native to the EnDEVR platform. We are combining logical programming with Category Theory (a field of mathematics) to assess properties of user-uploaded data to determine the best mathematical representations for analysis and visualization within VR. We are also planning to use this capability to tell the user in specific scenarios that "Hey, you may want to try this algorithm or convert your data to this kind of mathematical representation to achieve your desired outcome." If unchecked, this can easily lead to certain biased recommendations and introduce blindspots to analysis.

And the second is the more general concept of democratizing artificial intelligence within the agency. Remember, one of the key features of the EnDEVR system is the ability for users to submit their own data science algorithms, using an open-ended array of languages, platforms, and third-party libraries. And this WILL, by design, lead to execution of unknown or loosely understood machine learning code in one's software or data analysis pipeline. Both of these are concerns that we must study further before full deployment across the agency.

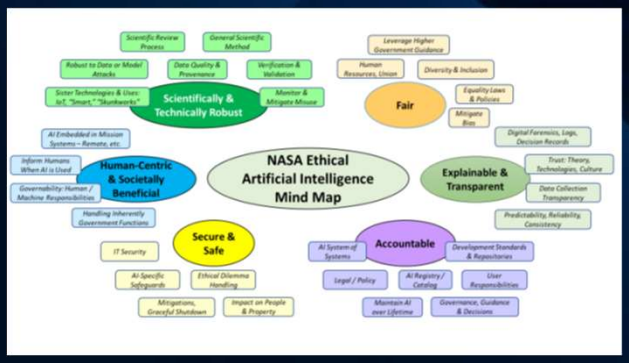
The third component of the system is future-facing, but it is an area of research that we want to pursue as something we can contribute to other programs. We have some pre-analysis work done and have some basic plans in place to pursue a framework for guided testing and intelligent validation within the EnDEVR system. The need for a component such as this arises from the high level of difficulty inherent in automating test scenarios within a human-centric environment such as VR, and as such, would be responsible for walking users through manual test scenarios. Machine learning would decide what scenarios should take priority here. And again, its something where we can transition methods and results to other programs. But we also want to understand any ethical ramifications behind using machine learning for this kind of testing.

arm DevSummit

AI Ethics at NASA

- NASA Framework for the Ethical Use of Artificial Intelligence released in April 2021
- Applies principles from expert studies and other Government agencies to NASA mission and operations

Defense Innovation Board	Gartner	ACT-IAC
Responsible	Secure and Safe	Bias
Equitable	Fair	Fair
Traceable	Explainable and Transparent	Transparent
Governable	Accountable	Responsible
Reliable	Human-centric, Societally Beneficial	Interpretable

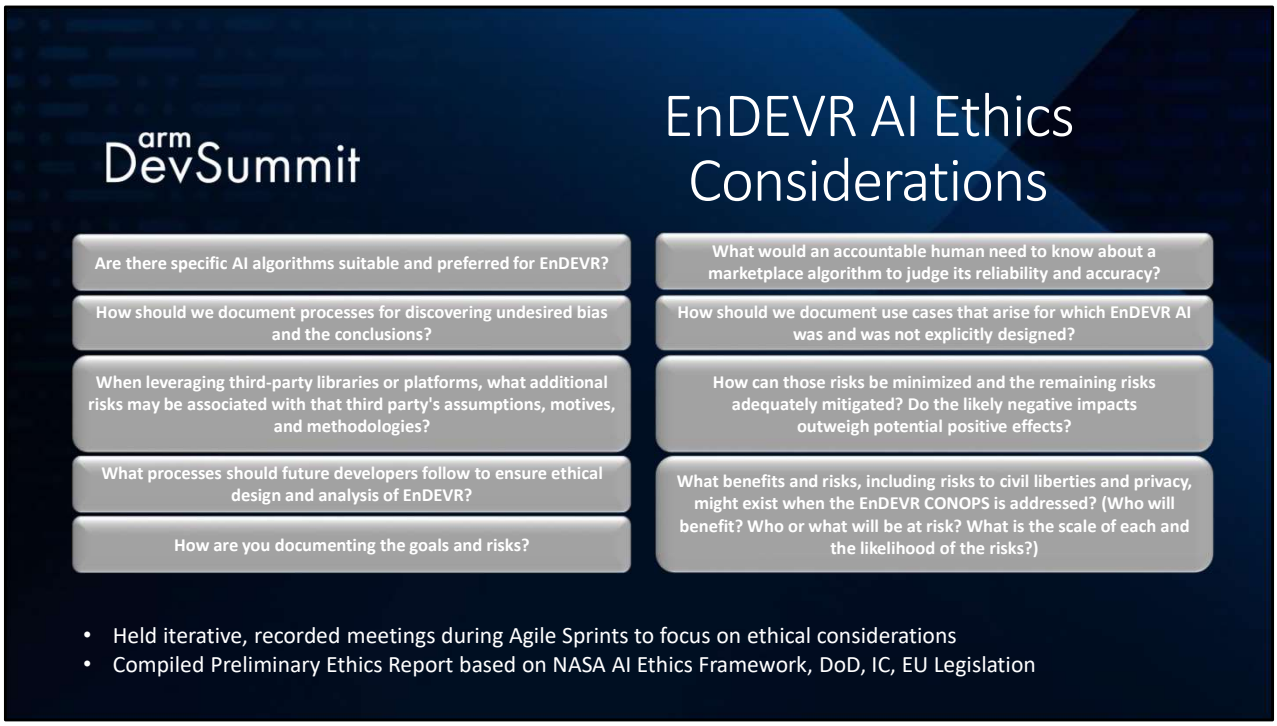


* NASA. NASA Framework for the Ethical Use of Artificial Intelligence (AI). Available from: https://ntrs.nasa.gov/api/citations/20210012886/downloads/NASA-TM- 20210012886. pdf

So just a quick aside on how this works at NASA. In 2019, a representative poll across NASA revealed over one hundred agency applications of AI in the previous three years, with hundreds of AI projects planned across various missions, centers, and mission support activities from 2020 to 2022 and beyond. In November and December of 2020, the White House and Office of Management and Budget (OMB) published guidance3 regarding AI principles, policy, and governance. As an enthusiastic and forward-leaning AI adopter, NASA created and has begun to apply an evolving, living set of AI policies, principles, and guidelines to provide AI practitioners an ethical framework for their work. They just released their Framework for the Ethical Use of Artificial Intelligence across the Agency this past April. My team's leadership was heavily involved in the creation of this framework and I can tell you, they pulled from a large number of studies and applications when putting this together.

They officially established key principles for projects that are implementing AI to follow. The idea is for these to serve as a guide. A project should demonstrate an assessment of these principles before, during, and after implementation. When developing this kind of framework for an organization (as I've seen in other frameworks like DoD's or the Intelligent Communities'), you have to have some set of

high-level themes for projects to work towards. Ideally, any given project would develop quantitative metrics for those themes and the project's actual concept-of-operations (CONOPS).



Throughout the summer, we started to compile a preliminary ethics report. To do so, we had the development team (mostly comprised of PhD interns) go through iterative iterations of reading and discussion over several months to give appropriate considerations to the subject. Their reading list included sources from the Department of Defense, U.S. Government's Intelligence Community, EU legislation, a range of ethics and security experts, and within NASA itself. At the beginning of the summer, the interns were guided by a set of discussion questions. The key thing to remember is that there were several goals associated with this exercise: for the developers to learn more about ethical concepts and frameworks surrounding AI and its related systems, applying those concepts to our own system (with particular attention paid to the EnDEVR Intelligent Assistant and storage and execution of arbitrary code), and presenting a set of recommendations for future ethical development of the EnDEVR system.

Based on the findings these discussions, the development team decided on the following recommendations going forward for EnDEVR:

First, keep an explicit list of the parties within the EnDEVR ecosystem, their assigned ethics and security responsibilities, and latest contact information. This one is pretty straightforward and allows us to evaluate accountability when incidences regarding ethics arise.

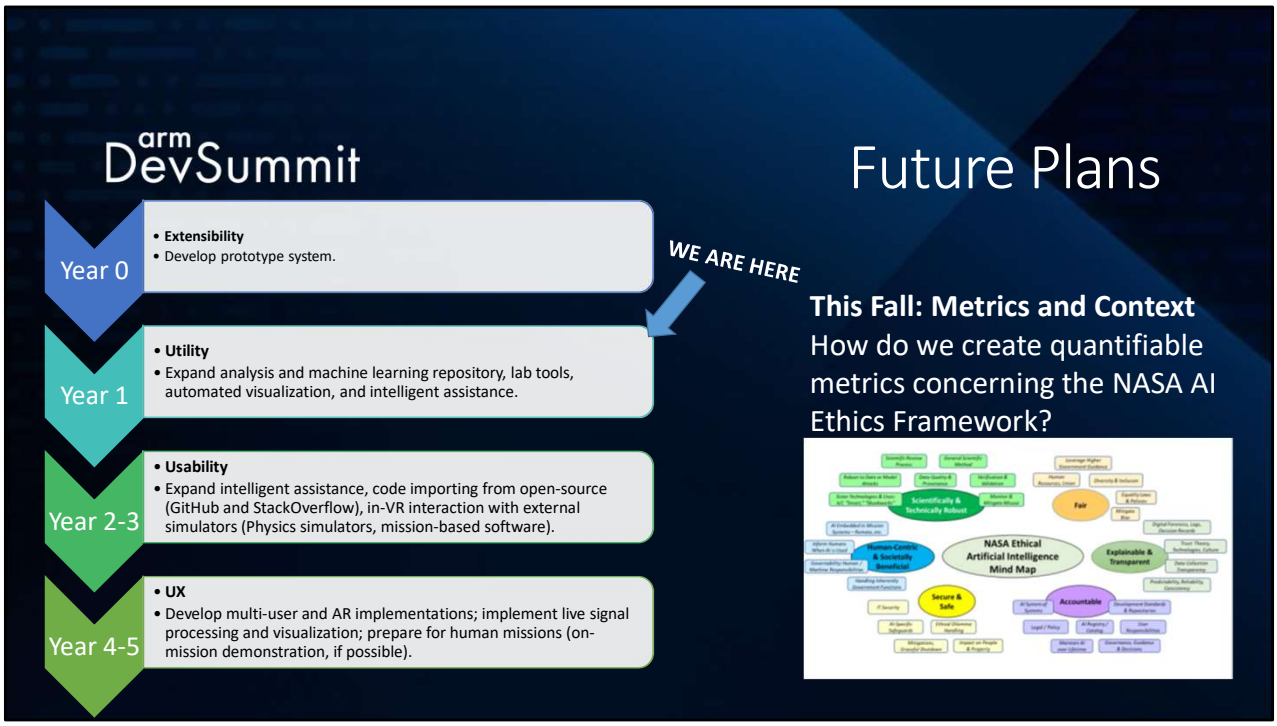
Second, Maintain a list of vetted third-party libraries and APIs that are approved for use within EnDEVR. The use of third-party libraries or platforms brings inherent risk to the system and its users. These potential risks include but are not limited to: malintent, system sustainability, and security risks. Malintent implies that the library was created with the intention of harming the systems of those who use it. Perhaps most important to EnDEVR is system sustainability: the long-term health of the system. In this context, security risks include anything that puts NASA at risk for system compromise or attack. Maintaining a list of these libraries is straightforward for us and will allow our team to continuously evaluate the health of our system when new vulnerabilities arise.

Third, Enforce requirements that users attempting to submit new algorithms to the EnDEVR marketplace must also submit detailed documentation regarding the internal operations of that algorithm. This should be supplemented with a review and vetting process for the algorithm, done by EnDEVR testers. For system availability, the algorithm can be immediately available to the user. But it will be labeled as Verified only once it passes the review process.

Fourth, Maintain a repository of living ethics and security documentation, with routine meetings for the team to review and update these documents as necessary. We are already doing this. The repository includes risk assessment models (e.g. DREAD, STRIDE, LINDDUN), descriptions of AI-based system components (e.g. Intelligent Assistant) and how they map to governmental and professional ethical guidelines, and the algorithmic descriptions submitted by users.

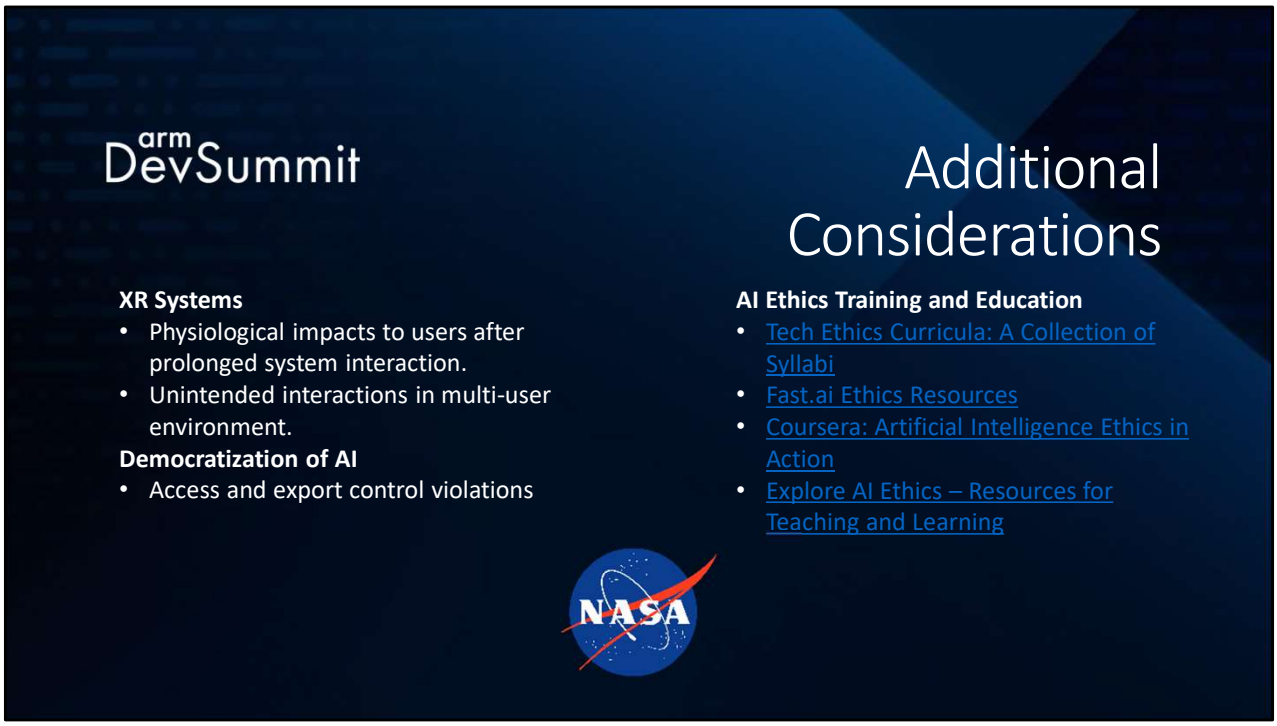
Fifth, Leverage the existing services, resources, and subject-matter expertise already available within NASA as much as possible. This will enable EnDEVR developers to address the standard system security concerns using solutions already present within the agency, and to take advantage of the breadth of knowledge embodied by NASA community members to address ethical and personal risks within the system.

Finally, and probably most importantly, Pursue continuous AI ethics education. It has become very clear that this area has NOT BEEN a part of the Computer Science or Data Analytics curriculum. And in fact, I'm writing an opinion piece on that right now for the Atlantic Council. NASA, at large, should work towards our workforce, particularly our technologists, becoming well-versed in the basics of AI and ethical ramifications of its use. The benefits of that for a workforce that will be working side-by-side with AI are immeasurable.



We have a longer-term roadmap for EnDEVR. The key isn't to just develop the capabilities that you see in this roadmap, but also to develop new techniques, some of them AI-based, as well as to understand the ethical ramifications thereof. And we do want to open-source certain capabilities, such as some of the fundamental technologies behind the intelligent assistant and algorithm sharing capabilities.

In addition, this Fall, we want to work on mapping the ethical priorities listed in the NASA AI Ethics Framework to quantifiable metrics concerning the EnDEVR system. How do we quantify transparency with respect to the algorithms submitted to the EnDEVR system? What novel approaches to quantifying Explainability in AI can we use to align with the framework? Instead of high-level themes and recommendations, we plan to study and publish results about HOW to quantify these things in a full Ethics Report for EnDEVR.



Finally, we recently discussed how some of our considerations apply in the context of XR systems, in general. These are elements that we will have to consider with respect to our system. But also, with respect to XR systems.

The first is looking at Physiological impacts to users after prolonged system interaction. The physiological effects of engaging with VR systems for extended periods of time are still under study, but can include such symptoms as nausea, headaches, and disorientation. As the EnDEVR system becomes more widely adopted, the likelihood of users experiencing adverse physiological effects after engaging with the system for too long is low, but the impact if it occurs is high.

The second is unintended interactions in multi-user environments. Studies have shown that users in a multi-user mixed-reality environment have trouble grasping concepts of ownership of virtual content, and have a tendency to leverage this virtual content to \mess with their fellow users. Users have also been known to suer sociological impacts such as exclusion during multi-user mixed-reality experiences. While the likelihood of this risk occurring is low, its potential impact is high due to the immersive nature of VR and the interactions of mis-intentioned fellow users.

The final one is looking at democratization of AI, which heavily applies to mobile systems like the Oculus (that we're using for this): Access and export control violations. This risk centers around the ability for users to access data sets or algorithms that they normally would not have access to, either because of flawed sharing authorization or because data and algorithms in aggregate leak more information than those entities normally would on their own. The likelihood of this happening and associated impact can be high depending on the organization.

Finally, just to remark on AI Ethics training. NASA and the larger scientific community must consider how to embed a code of ethics into emergent AI systems starting now. Fundamentally, we need to ensure that when AI systems reach near-human capabilities, ethical algorithms are already at the core of how they operate. Bolting on ethical behaviors once AI is well advanced would be a high-risk undertaking. That's because it is impossible to predict exactly when AGI and ASI thresholds will be achieved. If society builds ethics in early, there will be a better chance of ethical AI partners, no matter how, if, or when AI systems become self-aware. The more AI becomes integral to the organization, the more people will need to be trained to have a basic understanding of it. We leave you with a few resources to get you and your organization started on this endeavor. Thank you.