DECEMBER 2022 | Volume 43, Number 4
Published quarterly by the International Test and Evaluation Association

The ITEA Journal OF TEST AND EVALUATION

Cyber T&E – Testing for Resilience

- **Chaos Engineering**
- **Testing of Research and Development Projects**
- **Measuring the Measurers: Using Test to Validate Cyber Risk Assessments**

Connect with ITEA to Learn, Share, and Advance

About ITEA

For more than 40 years the International Test and Evaluation Association (ITEA), a 501(c)(3) not-for-profit education organization, has been advancing the exchange of technical, programmatic, and acquisition information among the test and evaluation community. ITEA members come together to learn and share with others from industry, government, and academia who are involved with the development and application of the policies and techniques used to assess effectiveness, reliability, interoperability, and safety of existing, legacy, and future technology-based weapon and non-weapon systems and products throughout their life cycle.

ITEA members embody a broad and diverse set of knowledge, skills, and abilities that span the full spectrum of the test and evaluation profession. All of which is shared with others through *The ITEA Journal of Test and Evaluation*—the industry's premier technical publication for the professional tester—and at ITEA's Annu-al Symposium, regional workshops, education courses, and local Chapter events. Join ITEA members—your peers in the industry—in contributing to *The ITEA Journal of Test and Evaluation* and participating at ITEA events so that you also can benefit from the opportunities to learn from others, share your knowledge, and help advance the T&E industry.

About the ITEA Journal

The ITEA Journal of Test and Evaluation (ISSN 1054-0229), published four times each year, is the premier publication for the Test and Evaluation industry. First published more than 40 years ago, *The ITEA Journal* quickly became and remains a leading journal in the field of test and evaluation, earning its stature as an authoritative international voice and eliciting ongoing acclaim for editorial excellence. Engineers, researchers, technicians, and academicians worldwide look to *The ITEA Journal* for the valuable information they need.

Each issue serves as a forum for authors of treatises on the cutting edge of testing science & technology and publishes papers on technical aspects ranging from basic research to applied research and development to operational testing. *The ITEA Journal* brings you the details of developments in this rapidly expanding area of technology long before they are commercial realities.

ANNUAL MEMBERSHIP FEES

- Industry Employees – $95 US
- Active-Duty Military or Full-Time Government Employee – $45
- Lifetime Memberships – prices vary
- Full Time Student – $25
- New T&E Professional – $25
- Large and Small Organizations – $800-$1500

JOIN ITEA

- WORKSHOPS
- EDUCATION COURSES

Annual ITEA Symposium

Test Instrumentation Workshop

Cyber T&E Workshop

Fundamentals of T&E Course

Operational Design of Experiments Course

Lunch & Learns

Professional Awards

Certification Program CTEP

ITEA Executive Office

11350 Random Hills Rd.
Suite 800
Fairfax, VA 22030
www.itea.org

Phone: 703.631.6220
Email: info@itea.org

Contents

The ITEA Journal
December 2022
Volume 43, Number 4

TECHNICAL ARTICLES

DEPARTMENTS

ON THE COVER:
This issue's cover represents the practices involved in cyber T&E and testing for resilience of systems and capabilities to evolving cyber threats.

The ITEA Journal of Test and Evaluation Themes for 2023-2024

The ITEA Publications Committee encourages you to write an article and share the upcoming themes with your colleagues.

Upcoming ITEA Journal Themes

Test and Evaluation Across the Acquisition Life Cycle

March 2023 (Issue 44 No 1)

For new and emerging technologies when should a T&E program start and end? How has integrated testing evolved over the years? What is the impact of the "Shift Left" initiative? Is there a role for T&E in fielded systems? Please consider submitting articles discussing T&E best practices and approaches across the full life cycle of a system from requirements development to deployment and beyond.

(Manuscript Deadline: December 1, 2022)

Test the Way We Operate – Best Practices and Lesson Learned

June 2023 (Issue 44 – Vol 2)

Testing in environments that reflect how systems will be used in operation provides an unbiased assessment of performance. This issue will showcase best practices and lessons learned from testing in realistic environments. Articles that include examples of how digital engineering can be used to create realistic environments and what realistic environments entail for machine learning and artificial intelligence enabled systems are also of interest.

(Manuscript Deadline: March 1, 2023)

DATAWorks Special Issue

September 2023 (Issue 44 – Vol 3)

The Defense and Aerospace Test and Analysis (DATA) Workshop is a workshop strategically designed to strengthen the community by applying rigorous statistical approaches to test design and data analysis in the fields of defense and aerospace. This issue will highlight technical articles from the 2023 issue of DATAWorks. Articles that showcase rigorous statistical approaches to test design and data analysis are also welcome in this issue.

(Manuscript Deadline: June 1, 2023)

T&E for Emerging Technologies

December 2023 (Issue 44 – Vol 4)

Technologies are rapidly evolving and software is providing novel capabilities in everything from radar to future communication systems. Artificial Intelligence and machine learning provide new capabilities to reprogrammable technologies. In this issue we will highlight technical articles for systems designed to change over time including Future G, reprogrammable radars, and all ML/AI enabled sensor technologies.

(Manuscript Deadline: September 1, 2023)

T&E for AI Enabled Systems

March 2024 (Issue 45 – Vol 1)

Test and evaluation provides important information at all stages of system development, production, and fielding. As systems change overtime, T&E provides information as to how performance might change as a function of system changes and operational environment changes. This issue will focus on different purposes of T&E programs and how information can be leveraged across the continuum of T&E.

(Manuscript Deadline: December 1, 2023)

CALL FOR ABSTRACTS

MARK YOUR CALENDAR

3rd Annual Multi-Domain Operations (MDO) Workshop

Multi-Domain Operations in an Extended Range Environment

Summer | 2023

Point Mugu Area | California

REQUESTED TOPICS:

- Rapid Experimentation
- Cross Domain / Multi-Level Security
- Technology Transition
- M&S to Support MDO in an Extended Range Environment
- Digital Engineering
- Methodologies for MDO in an Extended Range Environment
- MDO Tools in an Extended Range Environment
- Cyber / EW in Support of MDO
- Challenges & Solutions for MDO Test/ Training Planning and Execution in an Extended Range Environment
- AI & ML Battle Management

Abstracts Due March 1

Program Chair: Gil Torres, Naval Air Systems Command

Technical Chair: Kenny Sanchez, TRMC

Hosted by Channel Islands Chapter

26th Annual Test Instrumentation Workshop (TIW)

Instrumentation in a Constrained Environment

May 22-26 | 2023

Tuscany Suites | Las Vegas, NV

REQUESTED TOPICS:

- Cybersecurity
- Electronic Warfare
- Spectrum Limitations
- Hypersonics
- Testing in a Broad Open Area (e.g., Land, Sea and Space)
- GPS Denied
- Range Safety and Use of Autonomous Flight Termination
- Artificial Intelligence and Machine Learning

Abstracts Due January 13

Hands-On Lab Back by popular demand!

Program Chair: Policarpio Soberanis, Northrop Grumman

Technical Chair: Harry Cooper, JT4

Hosted by Southern Nevada and Antelope Valley Chapter

www.itea.org

Exhibit & Sponsorships Available. Contact Lena@itea.org

Become a Corporate Member

INDUSTRY VISIBILITY Unrivaled exposure for your company to the decision-makers in test and evaluation through economical advertising, exhibiting, and sponsorship opportunities

BUSINESS CONNECTIONS Interaction and expansive networking opportunities with test and evaluation professionals in government and industry

HIGH ROI ITEA's economical membership dues and exhibiting, advertising, and sponsorship rate deliver the highest return on your investment in the industry

BELONG TO YOUR PROFESSIONAL COMMUNITY ITEA provides a professional corporate home for your company and instantly affiliates you with the largest non-profit organization dedicated to the test and evaluation profession

Benefits

- Five individual memberships for organizations with 50 or more employees
- Three individual memberships for organizations with less than 50 employees
- Corporate Capabilities Directory
- Complimentary web link from the ITEA website to yours
- Discounts on exhibit booths, sponsorships, and registration fees at ITEA events
- Use of Corporate News to gain visibility in the Journal
- Discount on CareerConnections
- Discounted Sponsorship Rates
- Product Showcase visibility on Social Media
- Highlighted on Social Media/LinkedIn on a rotational basis

Membership Dues

Large Business - $1500

Small Business - $800

President's Corner

The ITEA Journal of Test and Evaluation 2022; 43: 193-194


Bruce Einfalt

Welcome ITEA family and friends! I hope you are healthy and enjoying the holiday season. We just completed a full year of outstanding ITEA Workshops, including an excellent annual Symposium. I could not be more pleased with the excitement and attendance at these events. I am extremely grateful to ELK Management Group and all our volunteers for making this year so successful!

Since the writing of September's President's corner, we held our 39th International Test and Evaluation Symposium in Virginia Beach, Virginia on September 27th-30th with a theme of "Forging the 21st Century T&E Tools in an Era of Great Power Competition." This event was hosted by the ITEA Hampton Roads Chapter. The Symposium was highlighted by many prominent speakers in the Test and Evaluation (T&E) community. The audience was thrilled to have a joint panel between Test Resource Management Center (TRMC) and Department of Homeland Security (DHS) on Critical and Emerging Technologies that was co-facilitated by George Rumford (SES), Director (acting), TRMC and Daniel Cotter (SES), Executive Director, Office of Science and Engineering, Science and Technology Directorate, DHS. The Symposium concluded with a US Coast Guard Station & Cutter tour hosted by DHS. If you were able to brave the hurricane, it was well worth attending.

I especially want to thank the Symposium Chair, Erwin Sabile, the committee chairs, and all the volunteers who helped make this year's Symposium successful. I was particularly impressed by how Erwin treated all the presenters and T&E audience as family and friends. His warm introductions and his summaries of speaker comments throughout the Symposium were done to perfection.

Our final ITEA event for 2022 was the 8th Cybersecurity Workshop in Destin, Florida held on October 18th-20th with a theme of "New Domains in Cybersecurity T&E." This event was hosted by the ITEA Emerald Coast Chapter which has successfully hosted past Cybersecurity workshops. The Program Chair was Shelby Pearce, who was the recipient of the 2022 ITEA Energizer award; and together with the Technical Chair, Andy Overbay, made a great team resulting in another outstanding workshop. The seven-phased, hands-on Red Team and Blue Team laboratory exercise was especially popular with the workshop attendees. I would like to thank the chairs, all committee members, and the numerous volunteers who, once again, made this event so successful!

ITEA continues to provide outstanding educational courses to Test Ranges and T&E organizations, such as The Fundamentals of T&E Course offered by Matt Reynolds and Larry Damman. These two gentlemen have been teaching this class for many years each time, they receive applause from their students — most recently from Naval Surface Warfare Center, Pt Hueneme. These students were lucky to hear such expertise from these two masters.

In addition, the Lunch and Learn Series schedule grew monthly and we are most thankful to our speakers and their diverse topics. These presentations greatly contribute to our overall education mission. Please check the ITEA website for details.

The ITEA 2023 calendar will begin in late January with a Systems Engineering joint conference with the Directed Energy Professional Society in Albuquerque, New Mexico. The spring and summer calendar will include our annual Test Instrumentation and Multi-Domain Operations workshops. We will close out the year with a celebration at our 40th ITEA Annual T&E Symposium in December. More information can be found on the ITEA website.

We just completed elections for new and renewing Directors on the ITEA Board of Directors. This election included three-year elected positions and a one-year appointed position. It is with great honor and appreciation that we welcome our newly elected and appointed ITEA Board of Directors for 2023: Van Sullivan, Robin Poston, Ph.D., Malcomn Tutty, Ph.D., and Keith Joiner, Ph.D. We also welcome Mark Neice to the Senior Advisory Committee.

As we welcome our newest Board and Senior Advisory Committee members, we must sadly say farewell to three departing elected board members who have made exceptional contributions to the ITEA organization — Peter Crump (Past President), Mark Brown (Ways and Mean Chair), and Brian Moore (Corporate Membership Chair).

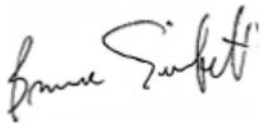
Our association had a fantastic year, and we remain fiscally healthy! We are finalizing our strategic plan and

have established a grants committee to seek future funding to support T&E interns. The Board also approved several bylaw changes in September, one of which was to rename the President's title to Chairman and Vice President to Vice Chairman. This means I was your last ITEA President and first Chairman of the Board. What a great honor!

Finally, my term as your Chairman for the ITEA Board is ending. I have enjoyed my time as your Board leader and I am forever grateful for the support of the Board, the individual members, and the corporate sponsors who have made our association so strong. I also want to thank Eileen Redd, Lena Moran, and Kathi Swagerty of ELK Management for their strong leadership and unwavering support. I am so glad they are part of our association.

I wish you a great Christmas and holiday season!

Sincerely,



Bruce Einfalt

Issue at a Glance

The ITEA Journal of Test and Evaluation 2022; 43: 195

To close out 2022, we have two new additions for our *Workforce of the Future* column. The first article comes from a Navreeta Singh, Princeton University, who interned with Jeremy Werner, Ph.D., at Director Operational Test and Evaluation (DOT&E). In their article, "Uncertainty Analysis Demonstration: A Missile Case Study," they highlight that validation is a key component of modeling and simulation (M&S) and role uncertainty quantification plays in the validation using sample missile test. The second student article comes from a team at the Virginia Tech National Security Institute. In her internship Cassidy Honea, Virginia Tech worked with researchers at VTNSI to develop the article, "Synthetic Aperture Radar (SAR): Utilizing Radon Image Transformations to Enhance the Detection of Ship Wakes." Erik Higgins, Daniel Sobien, and Justin Kauffman, Ph.D., worked with Ms. Honea to conduct the research, where they show that using simulated SAR that Radon Transformation improves the probability of ship wake detection.

Laura Freeman, Ph.D.

The first technical article on "Chaos Engineering" is authored by Jenn Bergstrom, Fellow at Parsons.

- Abstract: Chaos engineering, developed at Netflix, is a new approach for software testing at scale in an operational environment. The adoption of chaos engineering follows the pattern of other now popular public domain applications and tools, like Hadoop, TensorFlow, and PyTorch. Large tech companies dependent on web traffic and distributed software applications for their success have developed chaos experiment tools and released them to the open-source community. Early adopters have created as-a-Service chaos tools that are used by companies across multiple domains to validate the resiliency of their systems. The background and principles of chaos engineering are introduced and demonstrated for distributed systems operating in public cloud environments. Application to other domains, such as security, personnel, and to organizations, is also described.

In the second technical article, Michael Leite, Test Area Manager, R&D Programs at Department of Homeland Security, Science and Technology Directorate discusses, "Testing of Research and Development Projects."

- Abstract: The conduct and management of testing for research and development (R&D) programs involves different goals from the test and evaluation (T&E) efforts typically applied within acquisition programs. The Department of Homeland Security (DHS) has established a separate testing group to work with R&D program managers to ensure that testing of their projects is adequate to support the transition to acquisition, procurement, or commercialization.

Our final technical article comes from William "Data" Bryant, Ph.D. at Modern Technology Solutions, Inc. and discusses, "Measuring the Measurers: Using Test to Validate Cyber Risk Assessments."

- Cyber risk assessment processes claim to be able to predict the success of cyber-attacks against weapon systems and platforms in a range of highly contested cyber environments, but what evidence do we have that they are more accurate than random guesses or tea leaves? Fortunately, on systems that are also undergoing cyber test, we can leverage this test to validate that the risk process is effective by having risk assessors also provide predictions on what test cases will be successful and comparing the predictions and test results.

Enjoy the issue!

2023 Joint Conference:

T&E Support to Prototyping and Experimentation

Jan 30–Feb 2 | 2023

Marriott Pyramid North
Albuquerque, New Mexico

A collaboration between the 23rd Annual ITEA Systems of Systems Workshop and the 21st Annual Directed Energy T&E Conference

Call for Technical Papers, Tutorials, Poster Papers

> Abstracts Due by December 5, 2022

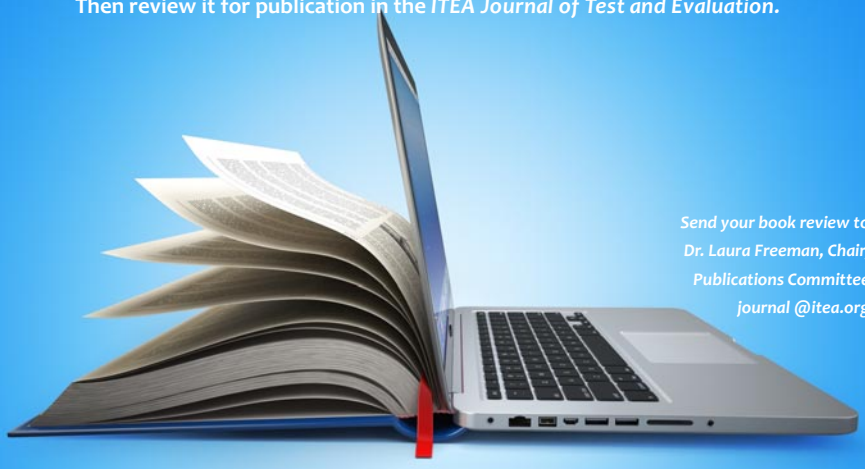
Go to www.itea.org for suggested topics and to register.

Co-produced by the International Test and Evaluation Association and the Directed Energy Professional Society

Hosted by the ITEA Roadrunner and Valley of the Sun Chapters

Read a book on one of the 2021—2022 theme topics.

Then review it for publication in the *ITEA Journal of Test and Evaluation*.



Send your book review to Dr. Laura Freeman, Chair, Publications Committee journal @itea.org

The ITEA Journal of Test and Evaluation 2022; 43: 197-201

~ WORKFORCE OF THE FUTURE: STUDENT ARTICLE ~

Synthetic Aperture Radar: Utilizing Radon Image Transformations to Enhance the Detection of Ship Wakes

Cassidy Honea
Erik Higgins
Virginia Tech, Blacksburg, VA

Daniel Sobien
Justin Kauffman, Ph.D.
National Security Institute, Virginia Tech, Arlington, VA

The output images from Synthetic Aperture Radar (SAR) simulations, highlighting surface ship wakes, are fed into a common image transformation to detect and enhance wakes that are present in the images. The image processing technique utilized in this work is the Radon Transform, commonly used in the medical field. Its corresponding inverse are used to obtain enhanced images. The goal of the Radon Transform is to detect "pixels of importance" and enhance the images overall. Results from using the Radon Transform and Inverse Radon Transform show that on simulated L-Band SAR images the enhanced images enable easier detection of a wake within the image compared to those without this processing step.

Keywords: Radon Transforms, Ship wakes, Synthetic Aperture Radar

Introduction

Satellite imaging platforms like the TerraSAR-X [1] and Sentinel-1a/1b [1] are examples of operating satellites that can offer real-time and large-scale observations in maritime environments. Our research focus is on the detection of wakes produced by surface ships to aid in applications such as ways to combat smuggling and sanction avoidance. It is extremely beneficial if the satellite is also SAR capable, meaning that it can filter through frequencies to pick up on ship wakes that are impossible to observe with the naked eye.

According to *Capella Space*, Synthetic Aperture Radar (SAR) is a way of creating an image using radio waves [2]. SAR is a candidate for remote sensing since the environmental conditions of the atmosphere do not affect the data return. Conditions such as time of day, clouds, or even smoke do not impact images generated from SAR. There are different bands of SAR that are dependent on the wavelengths of the waves sent from the sensor. The primary bands utilized in SAR are in the radio and microwave portions of the electromagnetic spectrum (see Figure 1). SAR bands from lowest frequency to highest respectfully include P-Band, L-Band, S-Band, C-Band, and X-Band [3]. Table 1 shows the range of frequencies and wavelengths for each band.

Most wakes that are easily detectable in SAR images are made by larger boats traveling on a linear path. Having a linear trajectory allows the characteristics of a wake to be more distinguishable in SAR imagery by highlighting the crests and troughs of waves through the expression of grayscale lines on the image.

Ship wakes on the sea surface can be easily obscured by wind waves on the surface, tidal currents, bottom topography, and other phenomena [4]. These obstacles can make it difficult to determine whether a wake is

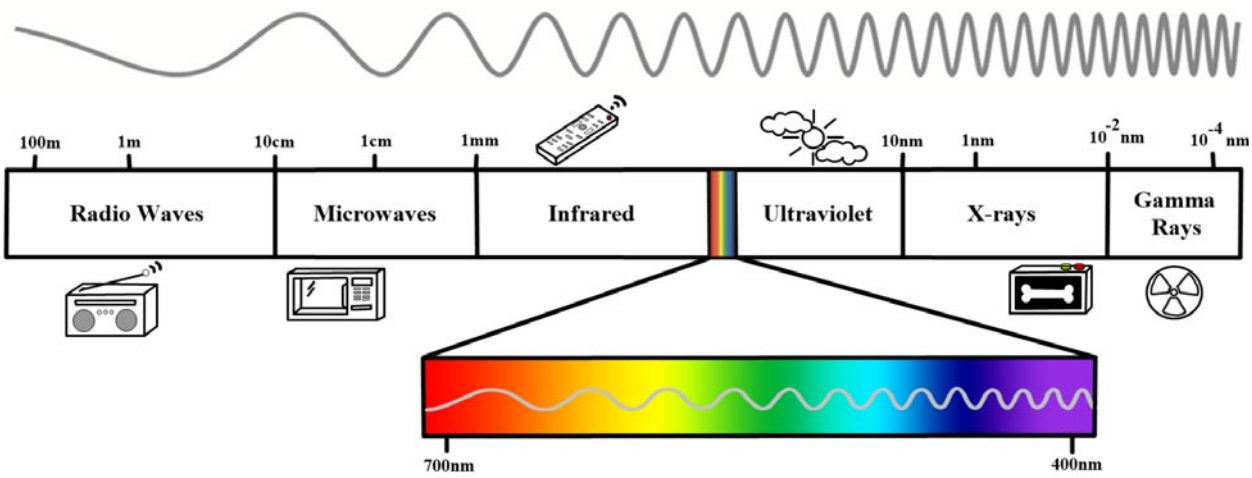


Figure 1: Electromagnetic Spectrum showing their respective wavelengths with examples of each.

Table 1: SAR bands with their corresponding frequencies and wavelengths [3].

Frequency SAR Band:	Frequency [GHz]	Wavelength [cm]
P-Band	0.25 - 0.5	60 - 120
L-Band	1 - 2	15 - 30
S-Band	2 - 3.75	8 - 15
C-Band	3.75 - 7.5	4 - 8
X-Band	7.5 - 12	2.5 - 4

present in the collected images from these real-world datasets provided by SAR capable satellites. This makes the problem more challenging, partly because of the great variety of appearances that wakes can have, and partly because of the presence of wake-like linear structures in the image and different kinds of noise. This introduces the question: How can we make these determinations easier and is there a way to "clean up" the image by highlighting the necessary information?

Image transformations can be used to modify an original image by converting it from one domain to another. Image transformations can be used for sparse object detection, clutter reduction, target classification, etc. Specifically, in this article, we propose the Radon Transform [5] and its inverse to tackle the aforementioned problems, where hidden data can be recovered to reevaluate and obtain previously obstructed data. Several different techniques were tested with the purpose of increasing probability in wake detection and consequently reducing the probability of missed or false detections.

This article is organized as follows: The Background introduces the Radon Transform followed by its inverse. The methodology of our analysis is presented with detailed steps. The imaging results on the provided synthetic L-Band Case images are reviewed in Results section, and finally conclusions are drawn.

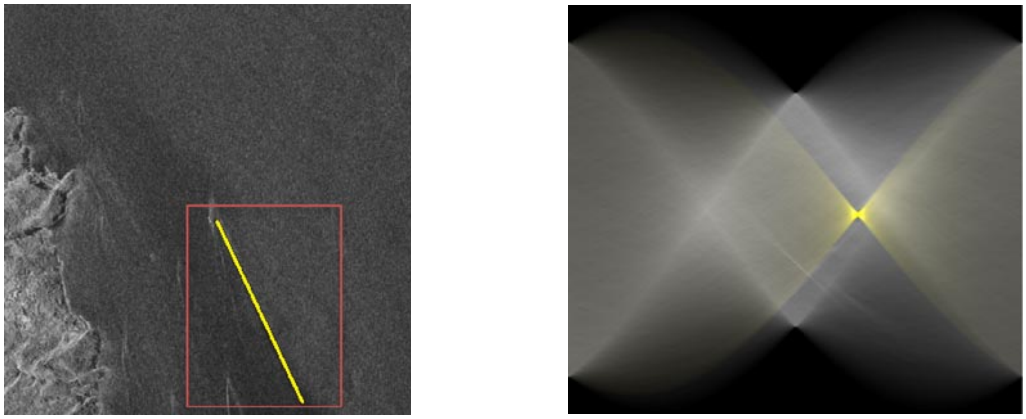
Background

The Radon Transform, defined in Equation 1, is an image processing technique (transform) that utilizes integration to enhance the quality of an image by filtering through some of the image noise.

$$R(\rho,\theta) = \int\int_D f(x,y)\delta(\rho - x\cos\theta - y\sin\theta)dxdy \quad (1)$$

$R(\rho,\theta)$ is the radon transform where ρ is the normal distance between the origin and a line, θ is the angle between the normal and x-axis, D represents the entire image domain, $f(x, y)$ the function value at position (x, y), and δ the Dirac delta function.

Used in most medical practices, ranging from X-Ray computed tomography (CT) to emission tomography [6], the Radon Transform integrates the values of the pixels within the image, along every line while each integral becomes a single point in the transform space. While looking into this transform, we asked the question: How can this transformation help in the detection of wakes created by maritime vessels? Looking into the different SAR bands, each image is processed at a different frequency. It is known that the Radon Transform can reduce the noise within an image, making the resulting image have a higher resolution than the input quality. Making it a candidate to declutter an image and become a valuable resource for precise wake estimation [7]. Figure 2 looks at highlighting a ship wake from the COSMO-SkyMed SAR data set, where the red box outlines the



(a) (b)

Figure 2: COSMO-SkyMed SAR image (a) where the Radon Transform is preformed on the section outlined in red. The wake is lined in yellow which can be seen in the sinogram (b). [8].

portion of the image where we performed the radon transform to produce a sinogram as shown in Figure 2b.

Methodology

We demonstrate the Radon Transform and its inverse on synthetically generated L-band images, details on how these images are generated can be found in [9] and [10]. Passing these images through these two transforms enable us to filter the image and reevaluate the resulting image in terms of a wake being present or not. We first took a synthetic image with a "clear wake" and performed the Radon Transform to analyze the sinogram, the output image from the Radon Transform [11], and then identify how the ship wake is represented within the sinogram. The detection of a ship wake within the sinogram could reduce the processing time in half if we are able to accurately identify a wake. We then performed the Inverse Radon Transform, which includes filtering, to demonstrate the noise reduction have a sense of the resulting output image through the entire image processing pipeline. Once we were familiar with the pipeline and how to identify ship wakes in the sinogram we repeated the procedure with images that were originally labeled as "no wake" or "half wake." Once we obtained all the resulting images, we compared the original images to the enhanced images to see if any of the results changed and if they need to be moved to a different category (i.e., move an image labeled "half wake" to "clear wake").

Results

Figure 3 shows all three steps of our methodology used on the image subsets within the L-Band Head Seas Case, where the waves are moving against the direction of motion for the ship. From each group, you can see a difference from the far-left columns to the right columns. This is caused by Forward Back Projection (FBP) within the Inverse Radon Transform where the original image undergoes some noise reduction, resulting in the same image but with a higher resolution and enhanced features. Descriptions of the SAR platform parameters that were varied are defined on the left most portion of Figure 3.

Figure 3a, labeled "clear wake," shows the synthetically produced full turbulent wake. Turbulent wakes appear on the surface of the water for a longer period due to their effect on surface roughness. Figures 3b and 3c are the resulting sinogram and denoised image from the Radon Transform and Inverse Radon Transform, respectively. The wake appears smoother for a longer period compared to shorter waves propagating because their phase velocity is slower than that of long (turbulent) waves [12]. Figure 3d, labeled "half wake," shows a similar synthetically produced turbulent wake with some energy loss after some time. Figures 3e and 3f are the resulting sinogram and denoised image from the Radon Transform and Inverse Radon Transform, respectively. Over time, the energy in the waves of water settle down, making the amplitude of the wave decrease. The water is only stagnant again once all the induced turbulence from the ship has died down, leading us into Figure 3g: "no wake." Figures 3h and 3i are the resulting sinogram and denoised image from the Radon Transform and Inverse Radon Transform, respectively.

It can be observed that as the parameters of the SAR platform change, the resulting synthetic wake changes as well. Noise in the original image can impact positive identification of a ship wake, leading to a designation

of "half wake." Using the Radon Transformation and its inverse, we can provide more confidence of an identification of a ship wake in resulting images. However, these particular conditions for the presented "half wake" in Figure 3 would result in a label of "no wake."

Conclusions

Radon Transforms have been used to enhance simulated SAR images to better classify if those images contain a ship wake. Even for the simulated SAR images, it can be difficult to determine if a ship wake is present given the amount of noise in some of the data, let alone if we are trying to identify ship wakes from real-world SAR. The approach utilizing Radon Transforms is demonstrated on synthetically generated data where the enhancements enable us to classify ship wakes more clearly in an image. This approach could be used on real-world collected SAR data to help detect ship wakes without the input from subject matter experts. ❑

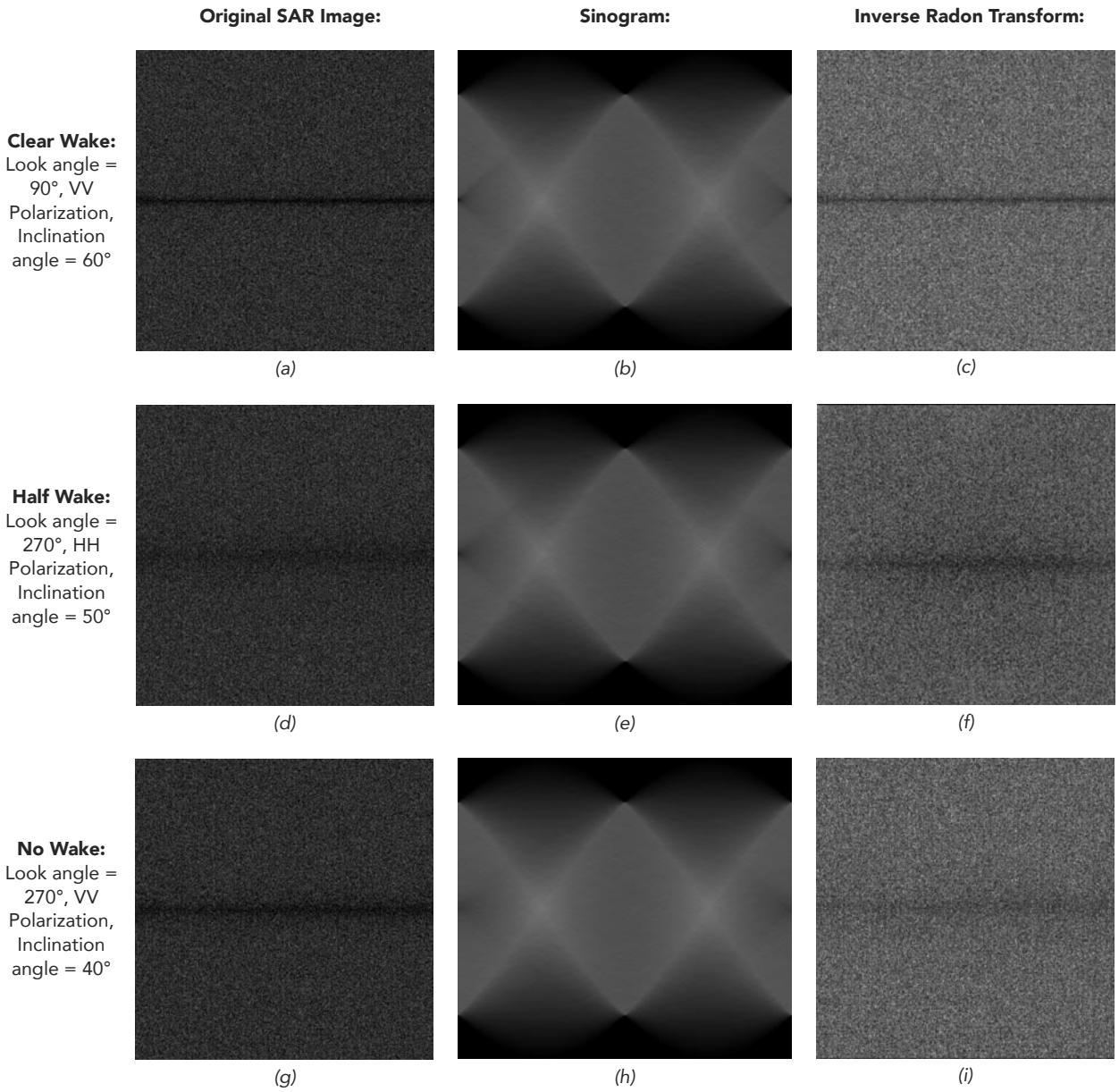


Figure 3: Each row has an assigned synthetically produced SAR image from the same hydrodynamic case (outline in [9]) and band: Head Seas Case: L-Band. Head Seas can be defined where the waves are moving directly against the course of the ship. Note VV/HH Polarization is vertical/horizontal transmit and vertical/horizontal receive.

CASSIDY HONEA is a junior at Virginia Polytechnic Institute and State University pursuing a degree in Ocean Engineering with two minors: Naval and Nuclear Engineering. She began research in this area during her sophomore year with more of a focus on hardware: making a low-cost custom drone, Synthetic Aperture Radar (SAR), and Remote Control (RC) boat. Last summer, Ms. Honea was selected for a Hume summer internship, where her work focused on utilizing Radon Transforms to assist in the classification of ship wake detection through a Naval Engineering Education Consortium (NEEC) grant through the Naval Surface Warfare Center (NSWC) Dahlgren.

ERIK HIGGINS is a 5th year Ph.D. student at Virginia Tech whose work focuses on computational mechanics, simulated remote sensing, machine learning, and data fusion. Mr. Higgins received a B.S. in Aerospace Engineering in 2018 and an M.S. in Aerospace Engineering in 2020, both from Virginia Tech.

DANIEL SOBIEN is a Research Associate at Virginia Tech's National Security Institute in Arlington, VA. His relevant research experience includes image classification and segmentation, data augmentations, testing and evaluation of computer vision models, and analysis of AI and human performers for detection and tracking objects in full motion video. His other research interests include AI assurance and causal machine learning. Mr. Sobien has a B.S. and M.S. in Aerospace Engineering, both from Virginia Tech.

JUSTIN KAUFFMAN, Ph.D. is a Research Assistant Professor of the Virginia Tech National Security Institute. His research interests include development of high-fidelity computational models and integrating machine learning and artificial intelligence into physics-based models to better predict phenomena of complex physical systems. Dr. Kauffman has a B.S. in Engineering Science, a B.S. in Mathematics, and a M.S. and Ph.D. in Engineering Science and Mechanics, all from The Pennsylvania State University.

Acknowledgments

This work relates to Department of Navy award N00174-22-1-0028 issued by the Office of Naval Research.

References

[1] C. Albinet, "SAR missions," *SAR Missions - Earth Online*, 2022. [Online]. Available: https://earth.esa.int/eogateway/activities/edap/sar-missions. [Accessed: 15-Aug-2022].

[2] D. Hogan and J. Brown, "Sar 101: An introduction to synthetic aperture radar," *Capella Space,* 10-Feb-2020. [Online]. Available: https://www.capellaspace.com/sar-101-an-introduction-to-synthetic-aperture-radar/ [Accessed: 18-Jul-2022].

[3] A. Moreira, P. Prats-Iraola, M. Younis, G. Krieger, I. Hajnsek and K. P. Papathanassiou, "A tutorial on synthetic aperture radar," in *IEEE Geoscience and Remote Sensing Magazine*, vol. 1, no. 1, pp. 6-43, March 2013, doi: 10.1109/MGRS.2013.2248301.

[4] Y. Liu and R. Deng, "Ship wakes in optical images," *Journal of Atmospheric and Oceanic Technology*, vol. 35, no. 8, pp. 1633–1648, 2018.

[5] M. D. Graziano, M. Grasso, and M. D'Errico, "Performance analysis of ship wake detection on sentinel-1 SAR images," *MDPI*, 2017.

[6] P. Kuchment, "The Radon Transform and medical imaging," *Society for Industrial and Applied Mathematics*, 2013.

[7] K. Landmark, A. H. Schistad Solberg, F. Albregtsen, A. Austeng and R. E. Hansen, "A Radon-Transform-Based Image Noise Filter—With Applications to Multibeam Bathymetry," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 11, pp. 6252-6273, Nov. 2015.

[8] E. S. Agency, "COSMO-Skymed," *Earth Online*. [Online]. Available: https://earth.esa.int/eogateway/missions/cosmo-skymed. [Accessed: 21-Sep-2022].

[9] E. Higgins, D. Sobien, L. Freeman, J. S. Pitt, "Data Fusion for Combining Information from Disparate Data Sources for Maritime Remote Sensing," AIAA SciTech, Virtual, 11-15 January 2021.

[10] E. Higgins, D. Sobien, L. Freeman, J. S. Pitt, "Ship Wake Detection Using Data Fusion in Multi-Sensor Remote Sensing Applications," AIAA SciTech, San Diego, CA, 3-7 January, 2022.

[11] J. Beatty, "The Radon Transform and the Mathematics of Medical Imaging," Honors Theses. 2012.

[12] A. T. Nylund, L. Arneborg, A. Tengberg, U. Mallast, and I.-M. Hassellöv, "In situ observations of turbulent ship wakes and their spatiotemporal extent," *Ocean Science*, vol. 17, no. 5, pp. 1285–1302, 2021.

The ITEA Journal of Test and Evaluation 2022; 43: 202-207

~ WORKFORCE OF THE FUTURE: STUDENT ARTICLE ~

Uncertainty Analysis Demonstration: A Missile Case Study

Navreeta Singh
Jeremy Werner, Ph.D.
DOT&E, Washington, DC

Background

Validation is a key component of modeling and simulation (M&S). Rigorous validation requires quantification of the uncertainty between live data and simulation output. The following demonstration supposes that we have data from a live-fire missile explosion test and simulated explosions [1]. A statistical analysis determines the extent to which the data and simulation agree.

Primer on Uncertainty

Uncertainty quantification estimates the extent to which a quantity, as measured, may differ from its actual value. The uncertainties themselves arise from limitations in measurements or M&S and can be categorized as *statistical* or *systematic*. Figure 1 below highlights the differences between the two.

Statistical uncertainty arises from stochastic effects (probabilistic effects that occur by chance) in a measurement process and is an estimate of imprecision. As the cause is random, statistical uncertainty may be reduced by accumulating more samples, and it approaches zero as the number of samples goes to infinity. Take determining the mean weight of a basketball approved for NBA games as an example. Weighing 1,000 different balls and then calculating the mean would yield a measurement with a much smaller statistical uncertainty than weighing only 10 balls.

On the other hand, systematic uncertainty is due to unknown but constant errors in measurement or M&S, which makes it independent of the number of samples. A systematic uncertainty estimates inaccuracy, and calibration error is a common source. If the scale used to measure the weight of the basketballs was calibrated only to a tolerance of 10 grams, then the systematic error in the mean weight of the basketballs will always be 10 grams, no matter the number of samples.

A Case Study

A model that characterizes a missile's impact on the target can significantly aid design and testing of that missile [2]. By depicting the number of fragments that perforate the target at a given distance from missile burst, the model can predict the amount of damage caused. For example, model output can then be used to inform a proximity sensor on the missile and help maximize area coverage on its target.

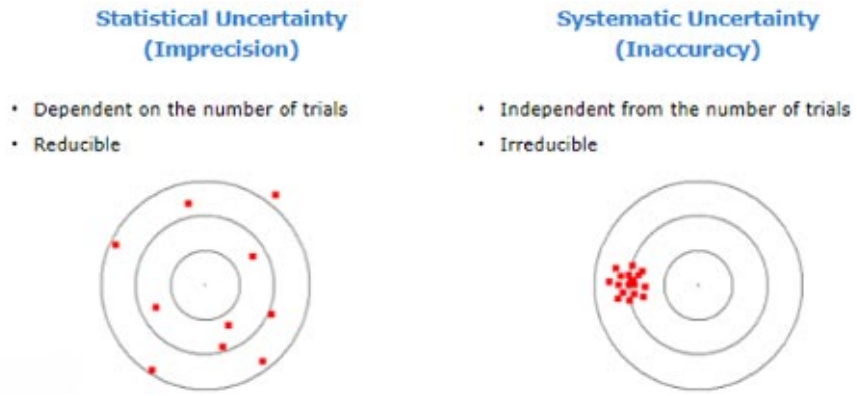


Figure 1: Visualization of the Difference Between Statistical and Systematic Uncertainty

Figure 2 below shows an experimental set-up for testing these parameters. The missile explodes in the center of the range, surrounded by "witness panels" at various distances from it. Few fragments impact the panels closest to the burst, as the explosion has not yet spread out. The panels farthest away also register few fragments, as the force of the explosion has dissipated. This leaves a mid-range "sweet spot," where the number of perforations is highest.

Method

The live data are first fit to a regression model. Because the perforation data are count data (i.e., discrete rather than continuous), we considered the following two main model options: Poisson and Negative Binomial. The Poisson model typically is used to describe random events that occur over time or space, such as the number of car accidents per month or the number of pieces of gum on a sidewalk square. Poisson models assume that each event is independent and that the mean of the data equals the variance. However, in practice, events often are correlated, which typically causes the variance to become larger than the mean. This larger variance, or overdispersion, is indeed observed in the missile data, because the events — the fragment bursts from a single mission explosion — are not independent.

The following two models can account for overdispersion: the quasi-Poisson and the Negative Binomial. The quasi-Poisson model includes an extra dispersion parameter to estimate how many times larger the data variance is than the mean. The Negative Binomial model considers the distribution parameter itself as a random variable whose variation accounts for the overdispersion. Although these two models were considered in addition to Poisson, the quasi-Poisson model (as implemented in the base library of the R Project statistical computing language [3]) is functionally identical to the Poisson model in terms of the regression it produces. This analysis discusses only the Poisson and Negative Binomial results [4].

After fitting the live data with Poisson and Negative Binomial distributions, the team computed the R^2 values (the proportions of variance in the data that are captured by the fitted regressions, also known as the coefficient of determination) and the probability values (p-values) of the χ^2 "goodness of fit" statistics, which estimate how well the models characterize the underlying data. Figure 3 below shows values computed from the average of 100 simulation runs, as well as from live data. The R^2 values

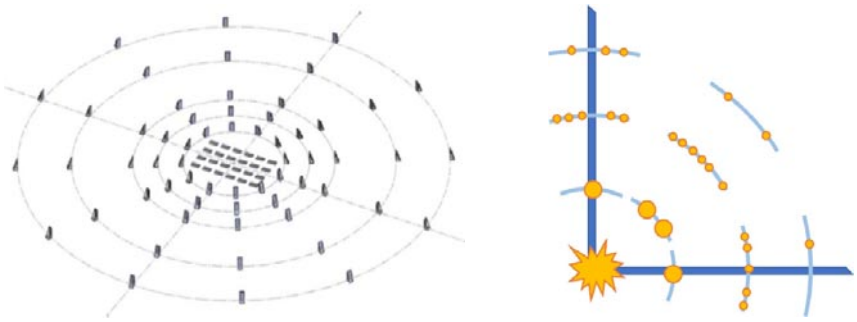


Figure 2: Experimental Set-up and Perforation Dependency Tied to Radius

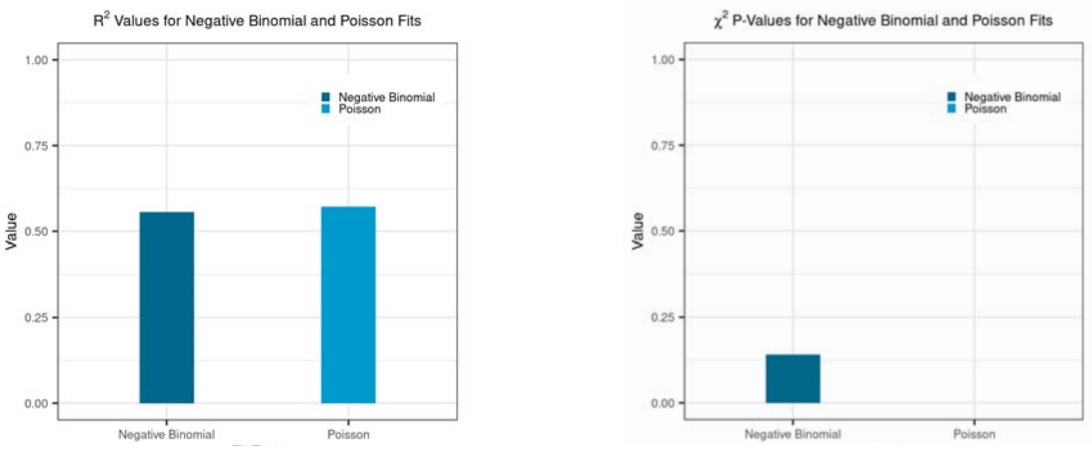


Figure 3: R^2 and χ^2 P-Values for Negative Binomial and Poisson Fits

for the Negative Binomial and Poisson fits are comparable, as the regression curves capture a similar proportion of the variance present in the data.

However, the fits' χ^2 p-values reveal the difference between the two.

The standard deviation of the Negative Binomial distribution is much wider, as this model accounts for overdispersion and, accordingly, yields a higher χ^2 p-value. Conversely, the Poisson fit doesn't account for overdispersion; thus, its χ^2 p-value is almost negligible. The Negative Binomial fit therefore is the better choice and is used for the remainder of this analysis.

Next, we compared the live data and simulation (see Figure 4 below). The 95% confidence band on the fit to the data, which is shown in gray, indicates that, for any given value on the horizontal axis, we are 95% sure that the mean of the parent distribution from which the data were sampled falls within these bounds. Shifting this interpretation, we can claim that to be considered consistent with the data, the mean of the simulation itself (teal curve) must fall within the confidence band — and so we correspondingly changed the color of the gray confidence band to teal.

We then drew points from the simulation, while assigning each point an error bar that represents the width of the confidence band, as any given simulated point could be drawn from the lowest or highest end of the band and remain within bounds. In doing so, we reinterpreted the *statistical uncertainty* that is latent in the fit to the data (due to the limited number of samples from which it was generated) as a *systematic uncertainty* in the simulation. The mean of the simulation could fall anywhere within the gray band and still be considered consistent with the data.

Analysis

Choosing the right hypothesis test is key. The left graph in Figure 5 below more clearly visualizes the extent of any disagreement between the two distributions, plotting the difference between the live data and simulation (black curve) with the 95% confidence band of that difference (gray band). The confidence band encompasses the horizontal axis (i.e., Live – Simulation = 0) for the entire range, demonstrating that the distributions are consistent.

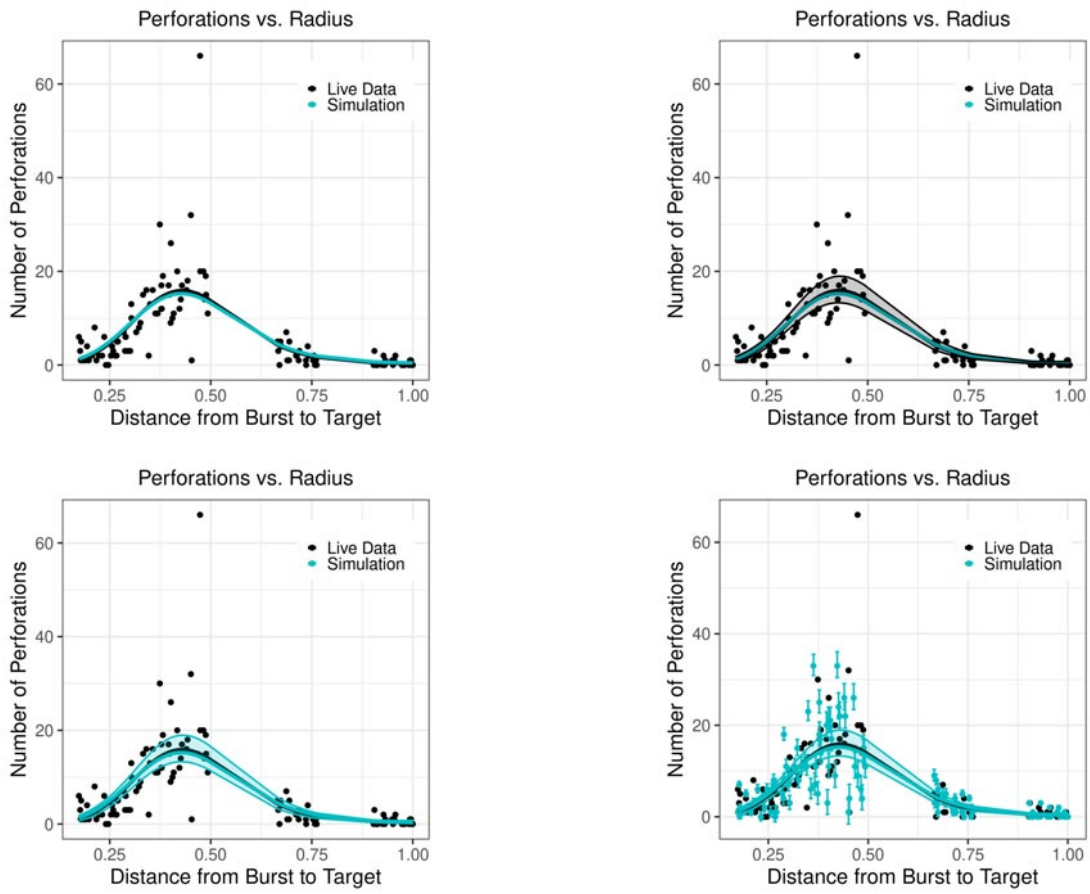


Figure 4: Live Data and Model Curve, with 95% Confidence Bands and Simulated Values

Hypothesis testing allows us to quantify the extent to which the live data and simulation agree or disagree. Gaussian distributed data can be analyzed with a student's t-test, which compares the mean of the simulation to that of the live data. But, because the data aren't Gaussian distributed, we instead used nonparametric tests that don't presume any distribution: the Kolmogorov-Smirnov (KS) test and the Mann-Whitney U (MW) test. The KS test compares the shapes of distributions through their empirical cumulative distribution functions (ECDFs), a plot of which is included on the right in Figure 5 below. The Mann-Whitney test compares the medians of the distributions in either the horizontal or vertical axes. By using a KS test, along with horizontal and vertical MW tests, we obtained a holistic comparison of the distributions.

Figure 6 exemplifies the importance of using multiple hypothesis tests. Each simulation was run 100 times,

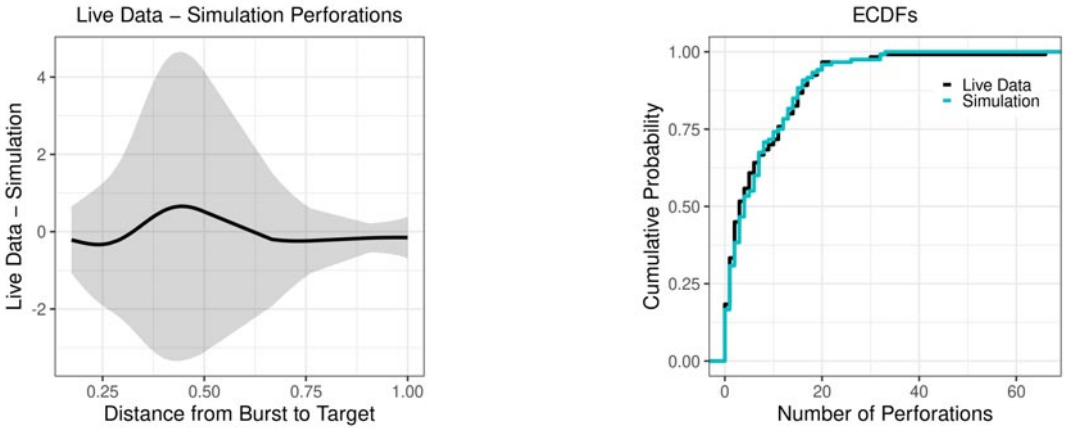


Figure 5: Live – Simulation Plot (Left) and ECDFs (Right)

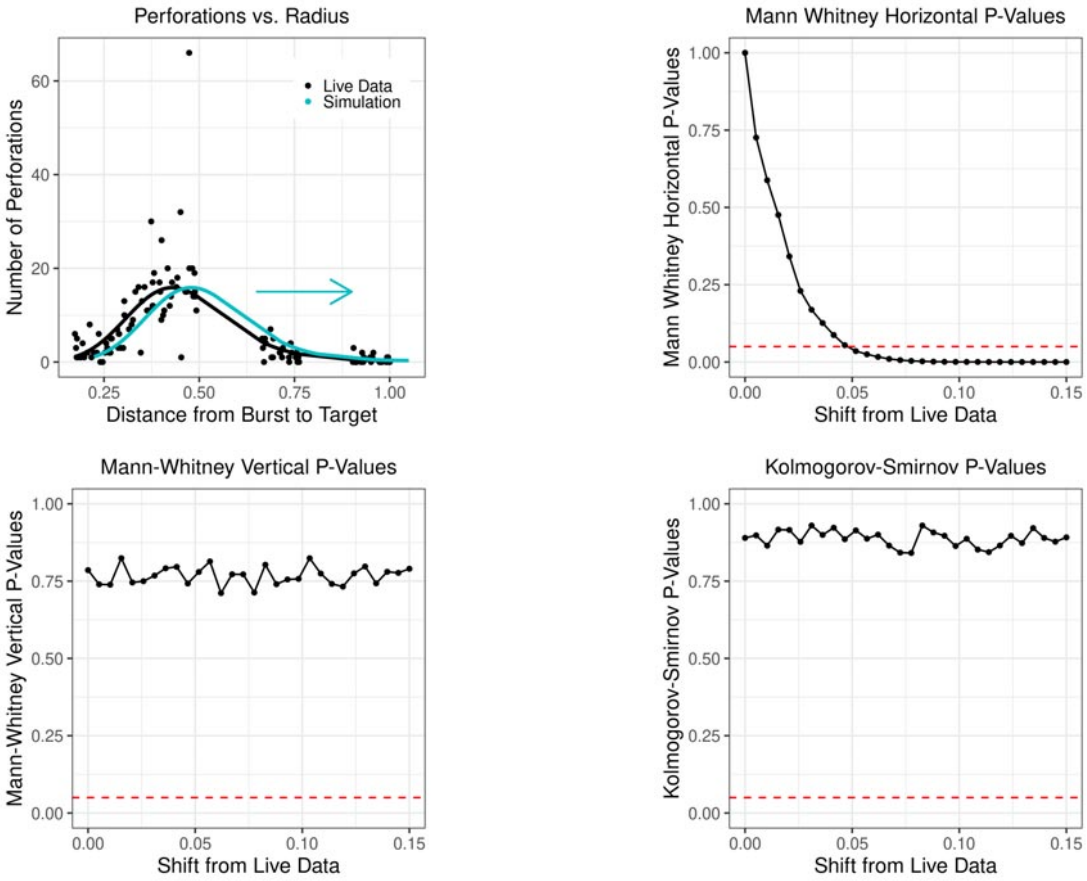


Figure 6: Horizontal and Vertical Mann-Whitney and Kolmogorov-Smirnov P-Values

and the means of the resulting p-values plotted. When we shifted the simulation to the right, both the KS and vertical MW test p-values remain unchanged. An analyst looking at these values might presume — incorrectly — that the live data and model agree, when they clearly do not. However, the horizontal MW test, which accounts for the direction in which the simulation is shifting, exhibits a steep downward trend as the simulation moves farther away from the live data. It intersects the horizontal p = 0.05 line at a shift of about 0.05 (arbitrary units), at which the probability of the data and model agreeing is 5%. The corresponding shift is exhibited in the Perforations vs. Radius plot shown in Figure 6. This is the point at which the data and model no longer exhibit the same distribution at the 95% confidence level.

Similarly, the KS test is sensitive to vertical shifts and changes in the shape of the distribution. We demonstrate this in Figure 7 below by comparing the live data against its regression line, shifted vertically by some value between 0 and 1. The KS p-values exhibit a sharp drop around a vertical shift of 0.4 (number of perforations). They cross the p = 0.05 line at a shift of only 0.55, demonstrating the sensitivity of the KS test to changes in the distribution.

What Does It Mean?

This analysis centered around uncertainty quantification, which moves beyond asking whether the data and simulation agree to determining the extent to which they may, in fact, vary but still be considered consistent. We established that the simulation could fall anywhere within the 95% confidence band and still be considered consistent with live data. This allowed us to reinterpret the statistical uncertainty latent in the fit to the data — due to the limited number of samples from which it was generated — as a systematic uncertainty in the simulation.

As US capabilities, the operating environment, and potential threats become more complex and challenging, operational testing and evaluation will rely more and more heavily on M&S. Just as with live data, quantifying the uncertainties that occur in M&S assessments of system performance is critical. Uncertainty quantification conveys the accuracy and precision of M&S results, helps to ensure those results' reliability and reproducibility, and allows testers and the intended user to have greater confidence in the predicted outcome. That, in turn, is critical to executing credible and adequate operational test and evaluation that provides decision makers and warfighters information they can trust. The straightforward method presented here for deriving systematic uncertainty will serve as a crucial tool in validating M&S venues — and setting the foundation to earn that trust. ❑

NAVREETA SINGH is a graduate student at Princeton University, pursuing a Master's degree in mechanical and aerospace engineering. She earned a Bachelor of Science in the same field from Princeton, along with certificates in materials science engineering and history, and the practice of diplomacy. Navreeta most recently worked in the Office of the Director, Operational Test and Evaluation (DOT&E) at the Pentagon, which sparked her interest in T&E. She previously conducted research at the Air Force Research Laboratory as an air armament scholar.

JEREMY WERNER, Ph.D., Senior Scientific Professional (ST) was appointed DOT&E's chief science advisor / chief scientist in December 2021 after starting at DOT&E as an action officer in the Naval Warfare Division in August 2021. Before

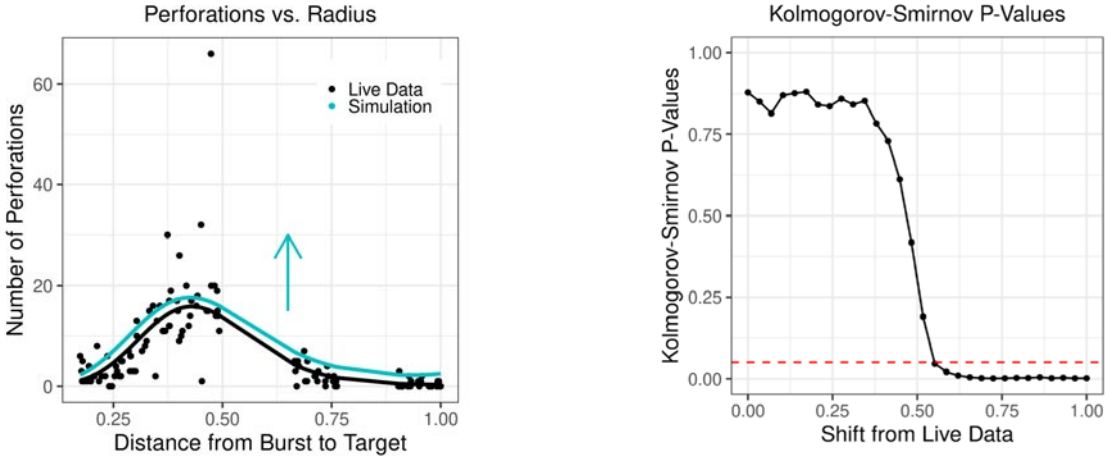


Figure 7: KS Response to Vertical Shift

then, Jeremy was at Johns Hopkins University Applied Physics Laboratory, where he founded a data science-oriented military operations research team that transformed the analytics of an ongoing military mission. Jeremy previously served as a research staff member at the Institute for Defense Analyses, where he supported DOT&E in the rigorous assessment of a variety of systems and platforms. Jeremy received a Ph.D. in physics from Princeton University, where he was an integral contributor to the Compact Muon Solenoid collaboration's experimental discovery of the Higgs boson at the Large Hadron Collider at CERN, the European Organization for Nuclear Research in Geneva, Switzerland. Jeremy is a native Californian and received a Bachelor's degree in physics from the University of California, Los Angeles, where he received the E. Lee Kinsey Prize (most outstanding graduating senior in physics).

Endnotes

[1] This study builds upon and uses input from IDA's 2018 "Comparing M&S Output to Live Test Data: A Missile System Case Study" by D. Thomas and K. Avery, to include the visualization shown in the left side of Figure 2.

[2] The analysis software needed to automatically reproduce all the findings in this study is publicly available at http:FIXME.it.

[3] https://www.r-project.org/

[4] When using statistical software, quasi models like quasi-Poisson also have limitations because they do not produce exact likelihood. Several statistical tests and fit measures are unavailable.

The ITEA Journal of Test and Evaluation 2022; 43: 208-218

Chaos Engineering

Jenn Bergstrom
Parsons Corporation, Denver, CO

Chaos engineering, developed at Netflix, is a new approach for software testing at scale in an operational environment. The adoption of chaos engineering follows the pattern of other now-popular public domain applications and tools, like Hadoop, TensorFlow, and PyTorch. Large tech companies dependent on web traffic and distributed software applications for their success have developed chaos experiment tools and released them to the open-source community. Early adopters have created as-a-Service chaos tools that are used by companies across multiple domains to validate the resiliency of their systems. The background and principles of chaos engineering are introduced and demonstrated for distributed systems operating in public cloud environments. Application to other domains, such as security, personnel, and to organizations, is also described.

Keywords: Software Testing, System Testing, Chaos Engineering, Resilient Systems, Fault Injection

Introduction

As an introduction, it might be insightful for the reader to know that this article could have been titled "Embrace the Chaos! Or, why I like breaking stuff on purpose." I have been breaking stuff on purpose for several years now and haven't been fired yet. How did this come about? The concept of chaos has a history that is now a few hundred years old, and chaos engineering is an outgrowth of that history, in synergy with the growth of cloud computing and widely distributed software microservice applications.

Chaos theory as a branch of physics dates to Henri Poincare in the late 19th century in his study of the three-body problem, but became formalized in the mid-20th century. Edward Lorenz, an American mathematician and meteorologist, is a name prominently associated with chaos theory and nonlinear dynamical systems. He showed that though chaotic systems appear to be random they are anything but, exhibiting underlying patterns and structure, emergent behavior, complexity, and self-organization. The study of chaos was greatly accelerated due to the digital computer, allowing detailed analysis of systems of nonlinear partial differential equations, establishment of the minimum requirements for chaos to exist, and providing dramatic visualization of related structures, such as attractors and trajectories. Later, chaos was observed in diverse systems beyond climate and weather, including networks, such as biological, chemical, social, and communications [1]. Edward Lorenz [2] wrote an accessible introduction to chaos, and science author James Gleick [3] wrote a popular book on the subject. Chaos theory has also been a factor in modern fiction, with a starring role in a Michael Crichton novel [4] and movie.

Chaos engineering is the discipline of experimenting on a system to build confidence in the system's capability to withstand turbulent conditions in production [5]. The purpose is to address the most significant system weaknesses before they affect users or customers, and to ensure that failures experienced by the system are not damaging to the capabilities delivered to the users and customers. Chaos engineering provides a way to manage inherent chaos and build confidence in complex production environments.

At first sight, chaos engineering may appear to demand mutually exclusive needs: intentionally introducing faults that cause system failures and testing at scale. A production network is characterized by an ever-changing mix of protocols, services, number of users, and connected devices; in fact, most networks today are networks of networks. Even with modeling and simulation on high-performance computers, it's nearly impossible to represent large scale network behavior. The only recourse is accepting the unknown and unpredictable and testing in a production environment to

understand how the systems will fail, as well as where, when, and how best to respond to minimize disruption to users and customers.

Background

When considering the practice of deliberately injecting failure into a production system, one may find oneself instinctively asking "Who came up with this idea?" But the approach, developed by Netflix when the company made a full migration of their software systems into Amazon Web Services (AWS), was driven by the application of the scientific method to complex system testing. Chaos engineering has five commonly identified steps for implementation, while the scientific method is described as having between five and eight steps. The steps are compared briefly for commonality in Table 1 and will be looked at more thoroughly in the "Nuts and Bolts" section of this article.

Netflix realized as they were completing their migration that the scale of their production system rendered the common practice of testing in staging or test environments ineffective. Because public cloud environments provide resources that are commoditized, there is a reasonable expectation that resources will experience failure. And because Netflix is a global streaming service, system service level agreements (SLAs) are aggressive and intolerant of downtime. Netflix needed a different way to test. This need led to the creation of chaos experiments, which are well-defined, well-planned, and well-observed tests completed by deliberately injecting failure states into the production system.

Initially, the chaos experiments Netflix implemented were small. Netflix created the Chaos Monkey in 2010, which was an automated test tool that randomly chose a production server and disabled it. This experiment validated that the Netflix application could gracefully handle failure of single servers in their highly distributed environment. However, Netflix quickly realized that testing the failure of a single server in their highly distributed environment only assessed the system against one small potential failure, and they began to develop a suite of experiments to apply to their systems. These experiments tested the failure of an entire data center, of multiple data centers, of network connectivity failure, of data store failure, and more. The experiments created to test these failure states were called the Netflix Simian Army [6], and the army grew from both observed failures in the production system and from identification of potential failures within the production system. Netflix further encouraged the growth of the Army by soliciting their users for recommendations for new experiments.

Table 1: General Comparison of the Scientific Method and Chaos Engineering

Step	Scientific Method	Chaos Engineering	Comparison
1	Observe the System	Validate System	For both the scientific method and chaos engineering, this initial step focuses on defining, understanding, and observing the characteristic behaviors of the system under investigation.
2	Define a Question to Investigate Gather Information	Form Hypothesis	In the second step of chaos engineering, the chaos engineer forms a hypothesis of potential system failures that could occur. In the second and third steps of the scientific method, the scientist defines a question to be investigated and gathers information related to the question.
3	Form Hypothesis	Plan Experiment	In chaos engineering, the third step involves planning out the chaos experiment that will be used to test the hypothesis defined in the third step. The fourth step of the scientific method has the scientist forming a hypothesis based upon the identified question and observations of the system.
4	Test the Hypothesis Gather and Analyze Data	Run Experiment	The fourth step of chaos engineering involves running the experiment that was planned out in the third step and in capturing the system data generated during the running of the experiment. The fifth and sixth steps of the scientific method involve testing the hypothesis, collecting data about the experiment, and analyzing the data.
5	Draw Conclusions Report Results Evaluate Results	Monitor and Repeat	In chaos engineering, the fifth step involves monitoring the experiment while it runs, drawing conclusions about the system based upon the data, and then running the experiment again. The seventh and eighth steps of the scientific method involve drawing conclusions from the experiment completed in steps five and six, reporting the results of the experiment, and evaluating those results against the initial hypothesis to either validate or invalidate its premise.

The code for Chaos Monkey was open-sourced in 2012. Netflix created the official Chaos Engineer job title in 2014, and in 2015 created an official Chaos Engineering Team [7]. In the same year, Netflix launched Chaos Community Day as a tactic to encourage adoption of chaos engineering by the broader technology sector. Companies across multiple industries began adopting chaos engineering for testing their distributed system deployments. The rate of adoption is increasing, and companies using chaos engineering are no longer constrained to those known for their technical innovation. Gartner estimates that up to 40% of organizations will incorporate chaos engineering into their DevOps disciplines by 2023 [8].

Chaotic Examples

Chaos engineering can be incorporated into programs, products, and projects across a range of disciplines. The techniques for implementing chaos vary depending upon the system being put under test. At Parsons, we run chaos experiments in multiple programs against targets ranging from IT infrastructure to distributed software applications and into our organizations themselves, including the people, processes, and tools used by our markets.

As an example of IT infrastructure and distributed software chaos engineering, we examine the experiments that are implemented in Parsons' Screaming Aardvark program. Screaming Aardvark processes commercial satellite data through a software system that is hosted in a Kubernetes cluster deployed on compute resources across on-premises, AWS, and Azure cloud environments. This system design provides resiliency at depth, protecting the program against individual software component failure, individual server failure, server rack failure, and data center failure. Because of the widely distributed and hybrid cloud nature of the deployment, traditional failure testing on the application and its infrastructure proved insufficient for assuring the resiliency of the system. In response to this challenge the Screaming Aardvark team chose to incorporate chaos engineering.

The team implemented three specific chaos experiments to test the IT infrastructure and distributed software of the system. The first test, which the team calls Khaos Monkey, terminates single pods running within the Kubernetes cluster. When the experiment is run, the automated script assesses the system to identify the currently running pods across the entire distributed system, randomly selects a subset of pods from each of the nodes, and terminates the selected pods sequentially, one at a time, while capturing log messages and recording the status of the system as the experiment is executing. This test validates the resiliency of each major component of the software system to individual pod failure.

The second test, called Khaos Gorilla, chooses a single node, either worker or control plane, and terminates it. For the on-premises environment, this means restarting the VM the node is running on. For the public cloud, this means issuing a terminate command on the instance or VM the node is running on. This test validates the resiliency of the Kubernetes cluster deployment, demonstrating that the infrastructure can recover from the loss of an entire server. This test also validates the resiliency of the software system by demonstrating that the software continues processing the satellite data without interruption or data loss even when multiple components of the solution become inaccessible concurrently.

The third test implemented for Screaming Aardvark is called Khaos Kong. This test emulates the loss of an entire data center and runs by terminating all resources operating within an AWS availability zone or an Azure region. Depending on which is chosen, this experiment can terminate up to half of the entire Kubernetes cluster. Khaos Kong tests the ability of the IT infrastructure and software system to mend itself and continue processing satellite data in the case of catastrophic failure.

The initial implementation of each of the Khaos experiments described above was manual, with an engineer triggering each experiment and controlling it. However, as Screaming Aardvark progressed in its implementation maturity, the experiments were similarly matured, automated, and incorporated into the development and staging delivery pipelines. By taking this step, chaos engineering is codified into the system, run with every code merge, and used to continually validate the resiliency of the system.

As mentioned previously, the tenets of chaos engineering can be applied in scenarios and against systems that are outside of software and IT infrastructure. At Parsons, the Parsons X team applies chaos engineering to its organizational structure, which includes the people, processes, and tools that contribute to the organization's day-to-day work behaviors, to ensure the organization is designed and operated in a way that builds resiliency and prevents single points of failure. Examples of chaos experiments that are applied to the Parsons X organization include Lying Lemur, Latency Loris, and Absentee Aye-Aye.

The Lying Lemur chaos experiment designates one or more individuals within the team as a liar for the day. Each designated liar is instructed to answer some percentage of all questions they are asked during the experiment incorrectly. The specific percentage varies,

typically ranging between ten and twenty-five percent, and the liar documents each instance where they lied in their response so that correct responses can be sent to all affected individuals after the conclusion of the experiment. The purpose of this experiment is to determine if the team knows enough about the program to detect when an answer seems likely to be incorrect, and to demonstrate the team's resiliency and ability to seek out the correct answer elsewhere.

Latency Loris is a chaos experiment that slows down responsiveness. In this experiment, one or more individuals are designated as slow lorises, which means that they are not allowed to respond to any requests for information via any method until a specified amount of time (30 minutes, 1 hour, 2 hours) has passed. The purpose of this experiment is like that of Lying Lemur, to demonstrate the team's resiliency and ability to seek out the needed answer elsewhere when the primary point of contact for it is delayed in their response. This experiment also tests tasking for the team: Does the questioner have productive tasking they can shift to while waiting for a response, or are they dependent upon the answer coming quickly to be productive?

Finally, the Absentee Aye-Aye experiment is like Khaos Gorilla, in that it removes a key piece from the organizational puzzle. In Absentee Aye-Aye, at the beginning of an individual's workday they are removed from their usual tasking and assigned a different, nonrelated task to focus on. The individual is not allowed to do anything that is related to their usual tasking for the duration of the experiment, including answering questions about it, doing any work on it, or interacting with the team involved in that task execution. This experiment is explicitly defined to identify any single points of failure in the processes and practices of the team tied to the redirected individual, as well as helping to identify any less recognized tasking that the individual does for the team.

In the Parsons X group, chaos experiments like the ones described above are run on a once-quarterly rhythm. As with regression testing in IT and software, this rhythm ensures that the organization protects its resiliency because the failure states induced as part of the chaos experiments recur.

Nuts and Bolts

The experiments summarized in the previous examples are effective tools. However, they are effective because the organizations and programs running them prepared and planned for the injection of chaos experiments into their operations. What does that preparation look like? How can an organization that is not currently using chaos engineering begin to inject it into their software, systems, and organizations?

There are five main enabling principles for chaos engineering. Any organization that intends to introduce chaos into their practice must understand the five principles, and their organization must be able to integrate them. These five principles are *validation, hypothesis, planning, execution, and repeatability*, and they align with the five steps described in Table 1.

Validation is using the process of observation to understand the way the system under test currently operates. For software and IT systems, this step includes ensuring that appropriate logging of the system state is captured to provide visibility into the steady state behavior of the system and confirming that system documentation is aligned with the actual deployment. For organizational systems, this step includes ensuring that processes and policies including tools and established team rhythms are understood and documented as executed.

Hypothesis incorporates the scientific method steps of defining a question to investigate and gathering information, and the output of this step in chaos engineering is a hypothesis that a chaos experiment will be designed to test.

The *planning* step of chaos engineering covers the scientific method's step of forming a hypothesis and involves defining the scope of the experiment, the aspects of the system that will be impacted by the experiment, the criteria that will be used to trigger a stop if the impact of the experiment is larger than intended, the duration of the experiment, and the team support requirements during the running of the experiment.

The *execution* step in chaos engineering encompasses the scientific method's steps of testing the hypothesis and gathering and analyzing the data. This step involves running the experiment, capturing the results of the experiment, and examining those results to understand what the impact of the experiment on the system was. In this step of chaos engineering, the team assesses the steps of the experiment and the results in the system of each step. The team does this by looking at logs, interviewing involved personnel, and reviewing the response of the team itself to the experiment. Existing remediation/failover plans are assessed for sufficiency in the induced failure state. Unintended and unexpected ripple effects are identified during the examination as well and can be used to determine additional experiments for future chaos engineering activities.

The final step in chaos engineering, *repeatability*, aligns with the scientific method's draw conclusions, report results, and evaluate results steps. This step includes reporting of the experiment results to stakeholders and

planning for new experiments identified in the execution step. This final step also includes activities between execution of chaos experiments as well, as the system is monitored for any unintended failures.

There are conditions that must be considered prior to incorporation of chaos engineering into a system to decrease the likelihood of failed experiment execution. Chaos experiments should not be run if the team already knows that the test will cause significant failure. If the system is opaque, meaning that it is either not well understood or not well monitored, that system is not a good candidate for chaos engineering. Visibility into the system's functionality and an understanding of the steady state operation is required for effective chaos engineering to be applied. Similarly, if there is not a well-defined recovery plan for failure in the system, chaos experiments should not be applied. Because chaos experiments involve injecting real failure states into a system, they should not be implemented without strong stakeholder support. Finally, if the chaos engineering team is not able to clearly define the expected area of impact for the test, criteria to stop the test, and the mechanism that will be used to halt the test in the case of a breach of the intended area of impact, chaos engineering should not be applied. A few examples of poor candidates for chaos experimentation:

1. A legacy software system that is manually deployed, with significant customization, and depended upon by multiple divisions within the organization.
2. An executive position within the company or key customer interface during essential negotiations.

If the system has been assessed and does not have any of the failure conditions identified in the previous paragraph, it may be a candidate for chaos engineering. However, before starting to experiment, ensure the following conditions are met:

1. The experiment can be scheduled to run during normal operating hours, when the team is present to support.
2. Stakeholders have been notified of the intended experimentation and provided their approval.
3. The environment the experiment will be run within is well understood and defined.
4. The experiment(s) to be run are clearly documented, including the intent of the experiment, the mechanism for injecting chaos, the steps that will be taken to restore the system after completion of the experiment, and the criteria and mechanism for stopping the experiment early if unintended effects are detected.

An additional consideration in implementation of chaos engineering on a system is whether the needed tools are available to the team running the test. For software and IT chaos engineering, there are a variety of commercial-off-the-shelf (COTS), as-a-service (aaS), and open-source (OS) tools available to use. The open-source tools are generally tailored to specific use cases or technology stacks and will not be applicable for every type of testing; the aaS and COTS tools may not meet all the requirements for chaos engineering in a specific system, so sometimes building your own tool is the best choice. However, there are risks with building your own tool, such as ineffective or insufficient mechanisms to halt an experiment that is spinning out of control, that are most often well covered in the aaS, COTS, and OS tools.

Choose the tools to use deliberately!

Once the tools are in place, the system is assessed for readiness, the stakeholders are on board, and the team is on hand, it is time to inject chaos! Let's get to it. When beginning to incorporate chaos into a system, start by choosing the target. Start small, as Netflix did with their Chaos Monkey. Test failure of a single pod, instance, VM, or designate a single Lying Lemur for the day.

Don't start with a Khaos Kong!

In addition to limiting the scope of the experiment, don't invest a lot of time into automating the experiment for the initial runs. If the experiment provides valuable insight into the system and is chosen to be run repeatedly, then take steps to automate. But for the initial few runs, manual implementation is an effective strategy.

Plan the experiment out. Consider where the experiment will be run. For an IT system or software application, will the experiment be run in the production environment? In staging or development?

Decide when the experiment will run. What day? What time? How much notice will impacted teams be given prior to the experiment?

Identify who needs to be present and dedicated to the experiment. Are all the support staff, the engineers who can repair the system, the staff that will need to deal with the impacts of the failures, and any other participants available and ready to support? How long will the experiment run? Will it take a few minutes? Hours? A full day? Longer?

Clearly identify the start and stop times and conditions that would cause an early halt to the experiment. Identify and document the steps that will need to be completed to restore the system to its original state once the experiment has concluded.

Once the plan has been created, it is time to run the experiment. Notify stakeholders of the start time,

planned duration, and expected impact. Start the experiment according to the defined plan. Pay attention while the experiment is running. Take notes, watch what happens, and observe how the involved teams react and respond to the experiment. Watch for unexpected effects related to the actions taken in the experiment and document the symptoms of those effects. If the identified stop conditions occur, stop the experiment early and initiate reconstitution of the original state of the system.

After the experiment has concluded and the system has been restored to its initial state, analyze the data collected during the run. Assess the effectiveness of the experiment. Did it reveal something new about the system? Identify a failure point that was unrecognized before? Did it validate the resiliency of the component of the system that was put under test? Was the hypothesis valid, and did the experiment reveal information that should be monitored and assessed regularly? If the experiment was valuable, consider ways to automate the experiment execution. If unexpected impacts were identified, consider creating new experiments to test those impacts in the future.

Thus far, we have examined the purpose, strategy, and implementation tactics that can be used to inject chaos experiments into systems of software, IT infrastructure, and organizations. We have walked through a few examples of chaos engineering as it has been implemented at Parsons, and we have discussed the steps involved in beginning to incorporate chaos into systems. The following section will walk through execution of a Chaos Monkey experiment in a simple distributed software application.

Demonstration

For the demonstration that will be shown in this section, a simple three-tier web application has been deployed in AWS' public cloud. The application is hosted on EC2 instances within an autoscaling group, is backed by DynamoDB and S3 datastores, and is fronted by an Application Load Balancer. Route53 has been used to customize the URL for the application. Because the application is hosted in AWS, the chaos experiment will be orchestrated using AWS Fault Injection Simulator (FIS), AWS's Chaos-as-a-Service offering that was released in general availability in March 2021. A simple Chaos Monkey experiment will be shown twice, once with the application deployed in a non-resilient, single availability zone manner and once with the application deployed with multi-availability zone resiliency. All images shown are provided with approval from AWS and show the system as viewed through the AWS console.

The Chaos Monkey experiment shown has two defined stages, as shown at the bottom of Figure 1. First, 50% of all running instances are terminated. Second, the experiment waits for 5 minutes to allow the autoscaling group health checks to recognize the loss of the EC2 instances and to automatically launch replacement instances.

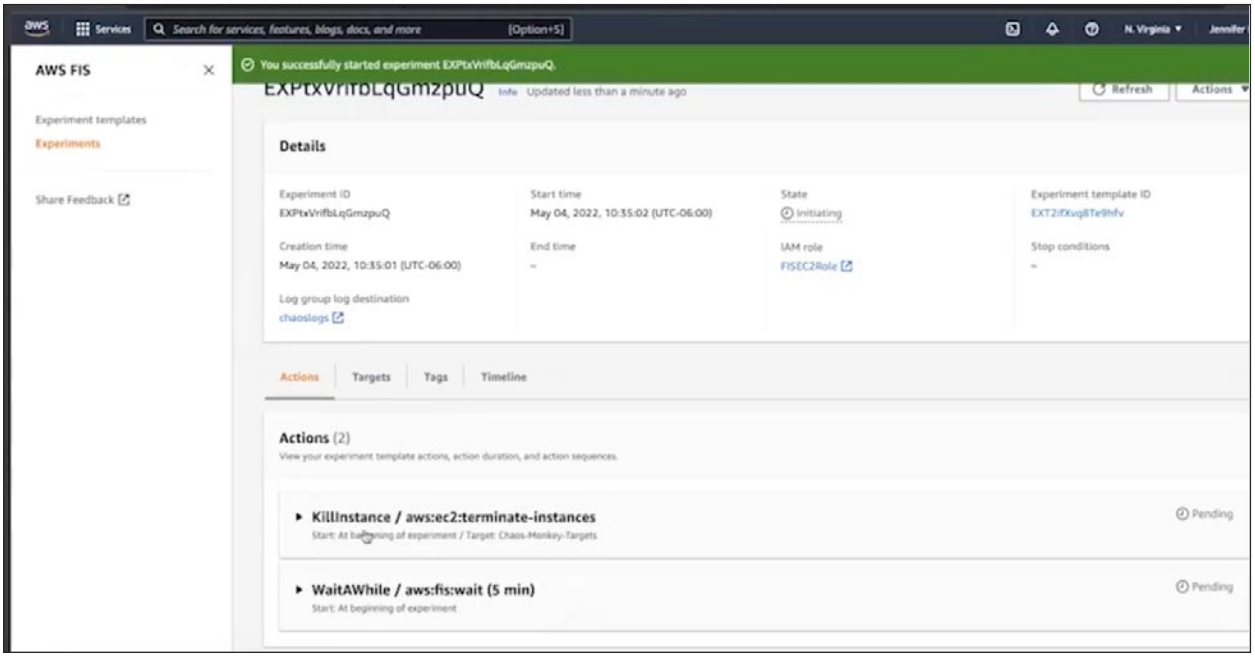


Figure 1: AWS Fault Injection Simulator Chaos Monkey Experiment Definition

For the initial deployment of the system, the autoscaling group is set to only want one instance, as shown in Figure 2.

With a single running instance, the website is active as shown in Figure 3.

When the FIS experiment is started, the first stage runs, selects the single running instance for termination, and issues the command to terminate the instance. Figure 4 shows the target group within the Chaos Monkey experiment, and Figure 5 shows the instances all terminated.

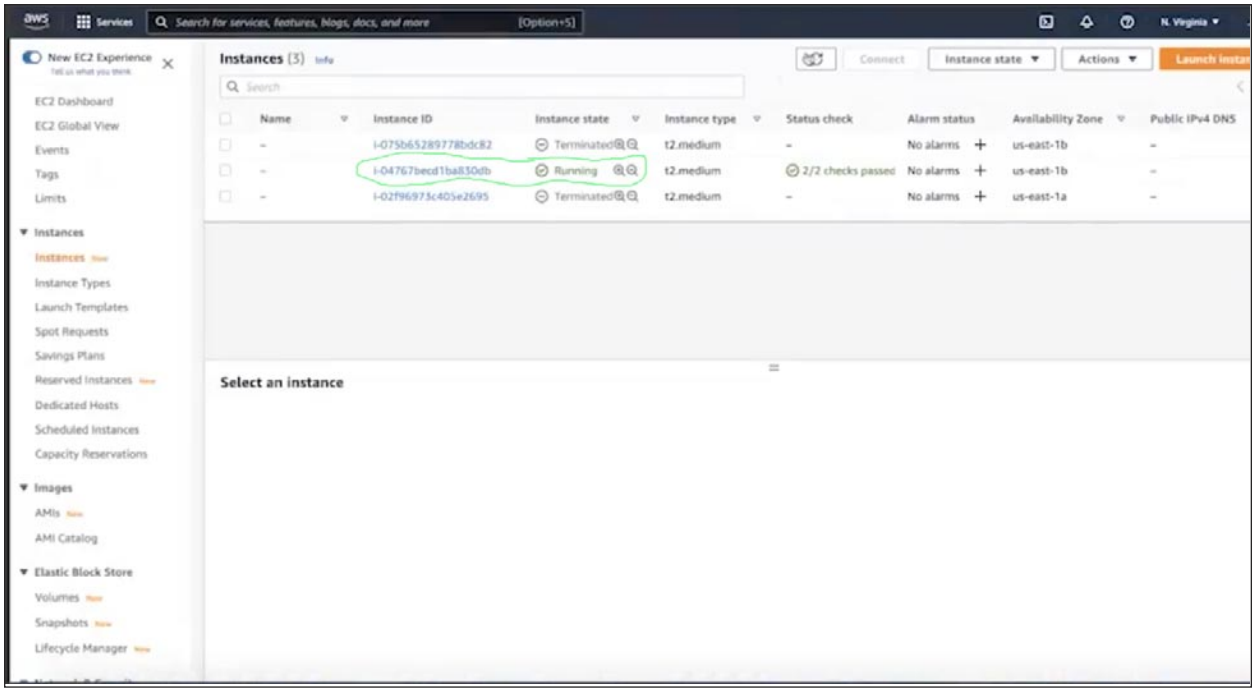


Figure 2: Single running instance shown in EC2, circled in green. The other two instances shown are in "Terminated" state, not running.

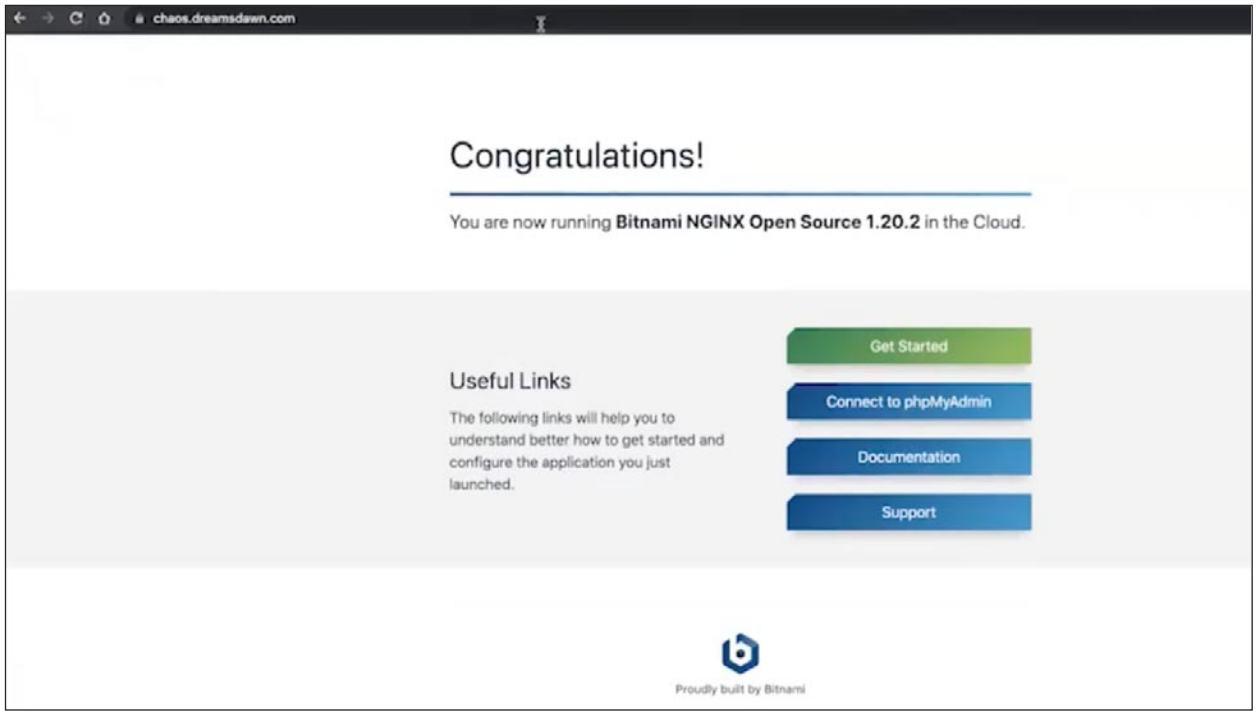


Figure 3: Basic Bitnami NGINX Website Running on the EC2 Instance.

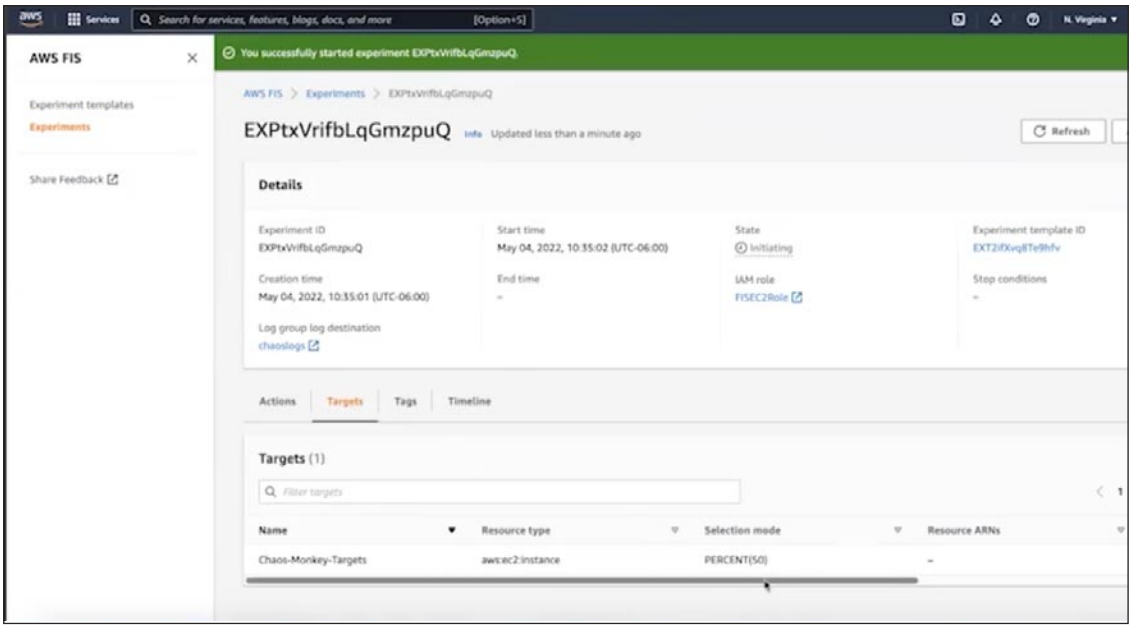


Figure 4: The Chaos Monkey Target Group Definition

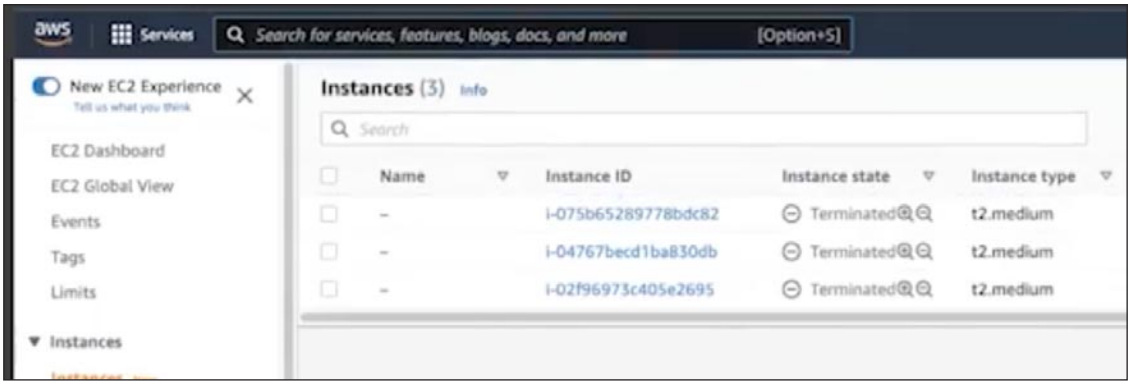


Figure 5: The Single Running Instance Has Been Terminated

And as shown in Figure 6, the website is now out of service. Once the loss of the instance is detected by the autoscaling group, a new instance is initialized as shown in Figure 7 and the website is restored as shown in Figure 8. The system experienced approximately three minutes of downtime from when the chaos experiment terminated the original instance to when the new instance was active and the website restored.

As a result of the initial chaos experiment, the autoscaling group was updated to provide two EC2 instances as shown in Figure 9. When the Chaos Monkey experiment is run again, one of the two EC2 instances is selected for termination as designed (shown in Figure 10). But this time because of the change made to the autoscaling group, the website does not experience any downtime due to the loss of the instance, as shown in Figure 11.

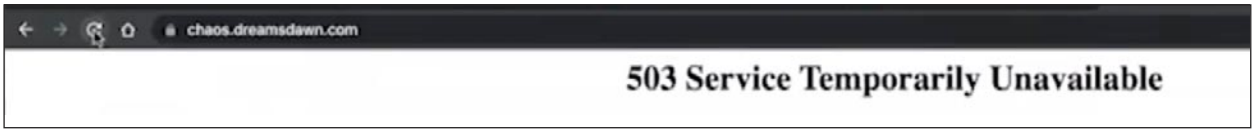


Figure 6: Website is Out of Service with the Loss of the Single EC2 Instance Due to the Chaos Monkey Experiment

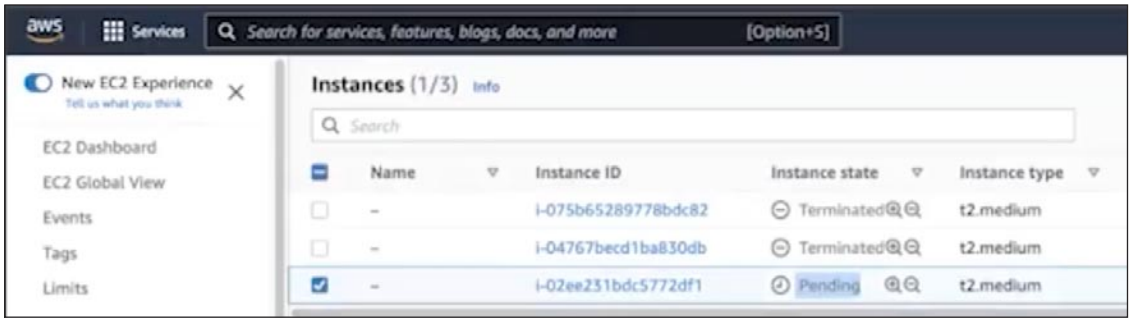


Figure 7: Replacement EC2 Instance (Selected in the Image) Initializing

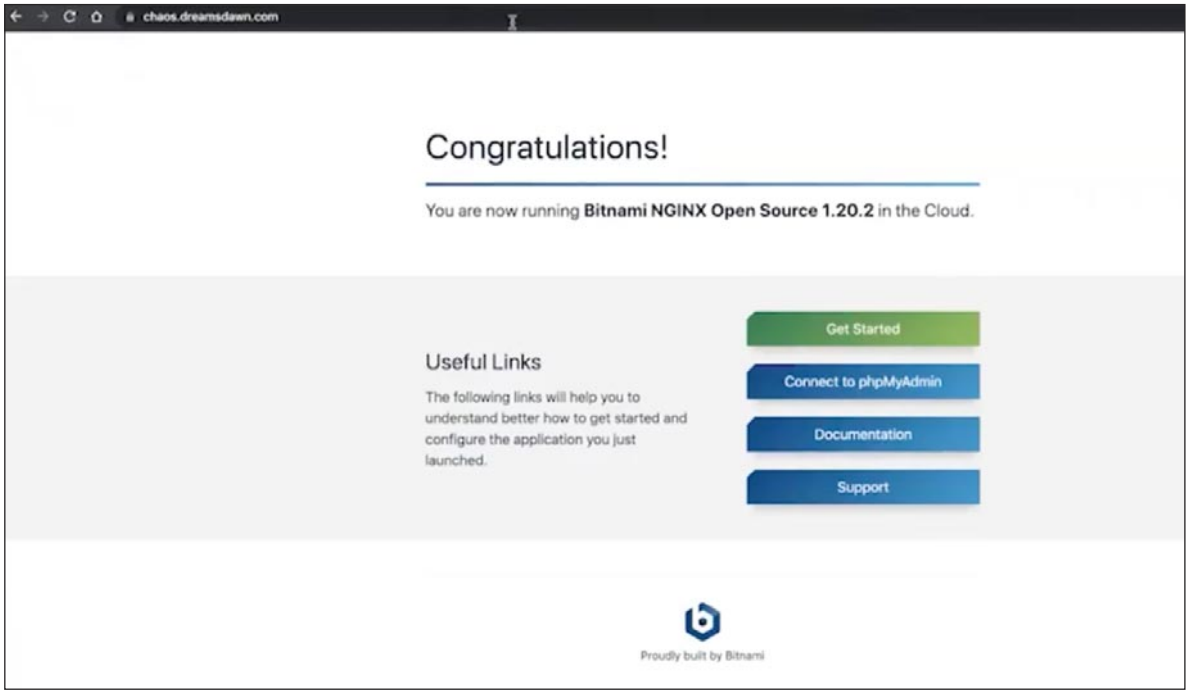


Figure 8: Website is Restored after Approximately 3 Minutes of Downtime

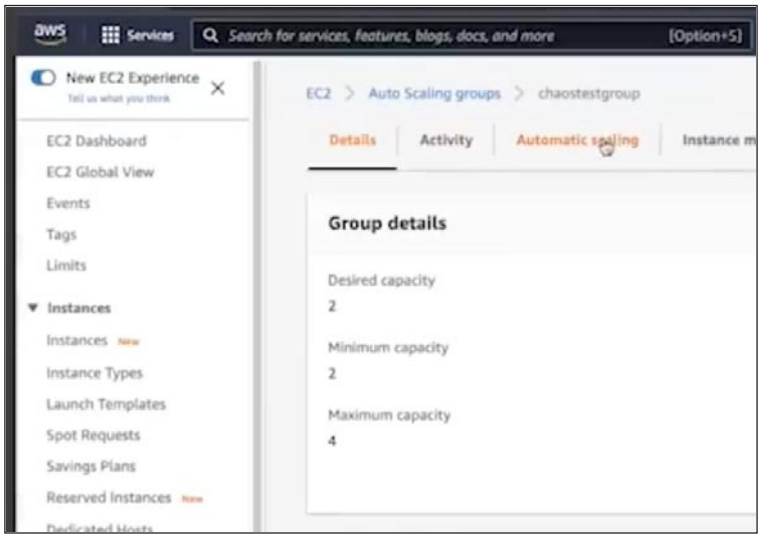


Figure 9: Updated Auto Scaling Group, Now with 2 Desired and 2 Minimum EC2 Instances

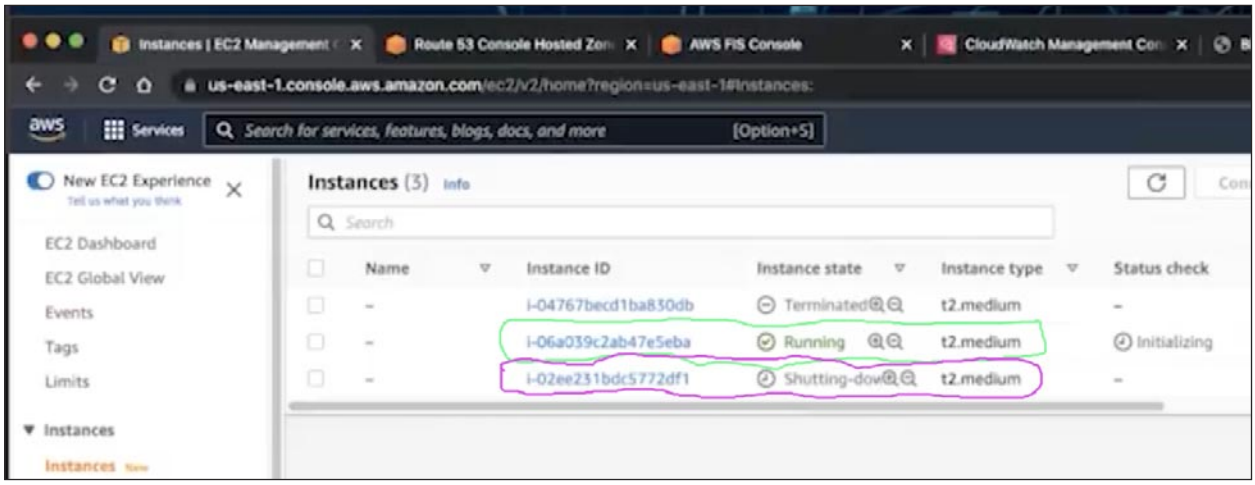


Figure 10: FIS shutting down one of the two running EC2 instances. Note the instance at the bottom of the list and circled in purple is shutting down, while the instance circled in green is still running.

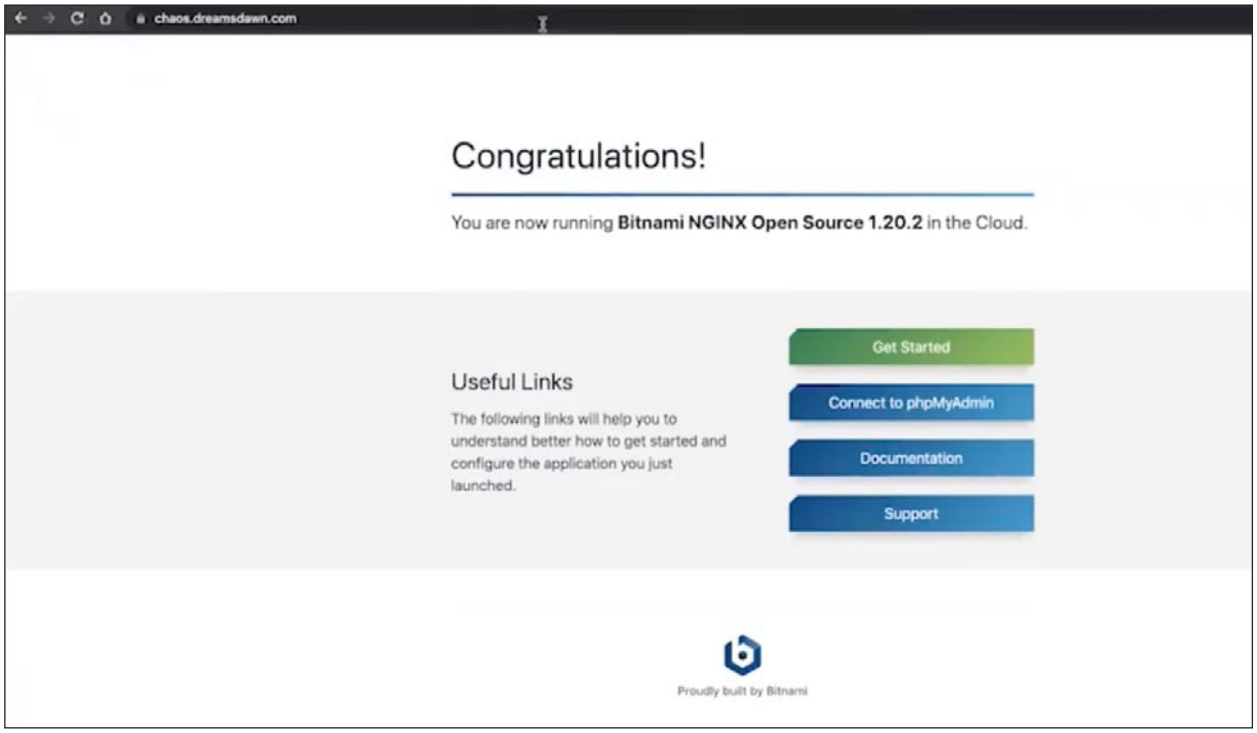


Figure 11: The Website is Active Throughout the Experiment

Conclusions and Cautions

For the widely distributed software architectures and organizational structures that are common in industry today, traditional methods for testing are insufficient. Test environments cannot provide the resolution and verisimilitude to the operational environment that is required for effective testing. A new method was required. Chaos engineering attempts to address the shortcomings of traditional methods by strategically injecting failure into the operational system to validate the scalability and resiliency of the system under test. While chaos engineering was originated to test IT infrastructures and widely distributed software applications in public cloud environments, the disciplined principles applied, like the scientific method, can be applied in additional circumstances. Security and organizational chaos engineering take the five tenets of chaos engineering and apply them to test the resiliency of security tooling, the security of IT infrastructures and software applications, and the people, processes, and tools relied upon by the organization.

Caution must be taken when incorporating chaos engineering into systems. Because a key tenet of chaos experiments is injection of actual failure states into the system under test, unexpected consequences and effects can and do occur. For this reason, rigorous preparation must be undertaken before experiments are run. Stakeholder knowledge and buy-in is essential. When run well, chaos experiments improve knowledge of a system and the way it functions, can provide evidence of resiliency, and increase confidence in the implementation of the system. However, if run ineffectively, chaos experiments can damage customer relationships, cost the organization revenue, and negatively impact the teams running and experiencing the tests and their effects on the system.

For many companies across multiple sectors of industry, the benefits outweigh the risks. Disciplined chaos engineering has been adopted by companies ranging in size from less than 100 employees up to large enterprises with well over 10,000 employees, and Gartner anticipates that greater than 40% of organizations using DevOps will include chaos engineering practices by 2023 [8]. ❑

JENN BERGSTROM is a Parsons Fellow with nearly 20 years of software and cloud engineering, business development, and leadership experience. Her continuous curiosity has led her to develop expertise in subjects ranging from multi-cloud solutions architecture, DevSecOps, and Chaos Engineering to Organizational Change and Project Management. Jenn graduated from Utah State University with a B.S. in Electrical Engineering and a minor in Mathematics and added an M.S. in Computer Science from Colorado Technical University shortly after. In her role as Senior Technical Director for Parsons X, Jenn designs and implements innovative solutions to complex problems and uses innovative techniques such as Chaos Engineering to implement highly resilient and secure systems and teams.

References

[1] National Research Council 2005. *Network Science*. Washington, DC: The National Academies Press. https://doi.org/10.17226/11516.

[2] Edward Lorenz, *The Essence of Chaos*, University of Washington Press, 1996.

[3] James Gleick, *Chaos: Making a New Science*, New York: Penguin, 1988.

[4] Michael Crichton, *Jurassic Park*, Ballantine Books, 2012.

[5] "Principles of Chaos Engineering," 2019. http://www.principlesofchaos.org.

[6] Izrailevsky and Tseitlin, "The Netflix Simian Army." Netflix Technology Blog, 2011. https://netflixtechblog.com/the-netflix-simian-army-16e57fbab116.

[7] Rosenthal and Jones, *Chaos Engineering: System Resiliency in Practice*. O'Reilly Media, 2020.

[8] Katie Costello, "The I&O Leader's Guide to Chaos Engineering," 2021. https://www.gartner.com/smarterwithgartner/the-io-leaders-guide-to-chaos-engineering.

The ITEA Journal of Test and Evaluation 2022; 43: 219-223

Testing of Research and Development Projects

Michael J. Leite
ManTech International Corporation, Herndon, VA

The conduct and management of testing for research and development (R&D) programs involves different goals from the test and evaluation (T&E) efforts typically applied within acquisition programs. The Department of Homeland Security (DHS) has established a separate testing group to work with R&D program managers to ensure that testing of their projects is adequate to support the transition to acquisition, procurement, or commercialization.

Introduction

In 2018, the Science and Technology Directorate (S&T) of the Department of Homeland Security (DHS) conducted a study to improve its support to DHS Components and become more responsive to their capability gaps. The recommendations of the study were implemented in fiscal year 2019 with the reorganization of the directorate and the establishment of a matrix support model. Key elements of the matrix model were the employment of systems engineering (SE) processes and techniques from test and evaluation (T&E) to support the research and development (R&D) programs.

The Problem

While testing is an essential element of R&D projects, the test objectives and methods are often different from those for acquisition and procurement. Testing in support of acquisition programs is typically divided into developmental and operational testing. Developmental testing ensures that a product is built in conformance with the contractual specifications: "Did we build it right?" Operational testing establishes the effectiveness, suitability, and resilience of the product: "Did we build the right thing?"

Testing to support acquisition decisions often involves long test periods and complex data collection. This complexity is due to the need for life cycle performance predictions, which include reliability, maintainability, logistic support, and cost projections. Accordingly, the power, confidence, and statistical significance of the results become key drivers in the amount and type of testing.

Similarly, the testing of procurement items is centered upon quality and performance. Consequently, first-article testing and the testing of samples from production runs or delivery lots are common means of testing procurement articles. As with acquisition testing, statistical and modeling methods are used to predict quality and performance of the delivered items and to confirm conformance to the published specifications or advertised performance criteria.

R&D testing, while using similar methods and techniques, often does not have a production-ready product to test. Rather, the testing is measuring a phenomenon, characterizing a material property, or assessing multiple solution options. R&D testing is often characterized into three types.

The first type of R&D test is the Critical Experiment. A critical experiment establishes a key capability or property of the research product. It may be a detection range, size, or physical property.

The second type is the Characterization Experiment. In this case, the program has established the fundamental property or capability; the next step is to fully characterize it. This is a controlled experiment in which the range or extent of the phenomenon is determined. In many cases, this takes the form of multiple tests in which the input or environmental conditions are varied to establish the performance envelope.

The third type is the Demonstration. Once the capability has been characterized, the next step is to show its utility. This may be accomplished by a demonstration in a laboratory environment or in a controlled field environment. Demonstrations can involve individual facilities or be scaled up to entire cities. The demonstration system is typically a prototype incorporating the essential system elements. It may also include design variants that are being considered for inclusion in the production system.

The end state of the demonstrations for an R&D project is a decision to move forward with a procurement or an acquisition.

A principal element of the transition from R&D to acquisition or procurement is the transition of testing results, data, and analysis to the acquisition or procurement team. For example, the data form the basis for the reliability, availability, and maintainability (RAM) assessment that is conducted as the operational product is developed. It is also the basis for the logistics and supportability planning.

The Process

The S&T matrix model was instantiated in an Operating Model Blueprint (OMBP) with three phases and a Business Process Flow (BPF) with nine sequential processes. The OMBP Phases and their associated BPF processes are shown in Figure 1.

Testing plays a key role in support of each of the three phases of the OMBP. The Test and Evaluation Division (TED) of the S&T Directorate analyzed each of the processes of the BPF and identified nine touchpoints in which T&E professionals from TED could support the R&D program manager.

It should be noted at this point most R&D programs are not large enough to have a dedicated T&E Manager assigned. Accordingly, selected members of the TED staff are matrixed into the R&D program organization. Also, they are separated from the TED staff that supports the leadership role of the Director of Test and Evaluation for acquisition programs. This precludes any conflict of interest for staff members responsible for the implementation of acquisition T&E. When an R&D program moves forward to an acquisition program, it is handed off from the R&D team to the appropriate acquisition Test Area Manager or Deputy Director.

TEA Touchpoints

To differentiate the testing support for R&D programs from the T&E role for acquisition programs, the R&D support effort was designated Testing, Experimentation, and Assessment (TEA). Nine TEA Touchpoints were identified where specific TEA activities can support the program manager. Figure 2 overlays the nine TEA touchpoints onto the nine BPF processes of the OMBP.

The nine TEA touchpoints were identified as a consequence of analyzing each of the business processes and determining how a testing subject matter expert could support the program or project manager. The initial analysis was conducted in 2020 by MITRE under contract to the DHS Director, Office of Test and Evaluation, and resulted in a draft *DHS Supplemental Guidance, Testing, Experimentation, and Assessment for Research and Development*. This document was subsequently revised in 2021 and updated to conform to version two of the BPF.

To avoid conflict with the numbering system for BPF processes and decisions, the touchpoints were given alpha designations. The following paragraphs describe the nine touchpoints and their application.

TEA A: Review Requirements

Requirements are reviewed to ensure they are measurable, testable, and achievable. This touchpoint occurs initially during Business Process 2, Operational Analysis and Gap Decomposition. It is also repeated during

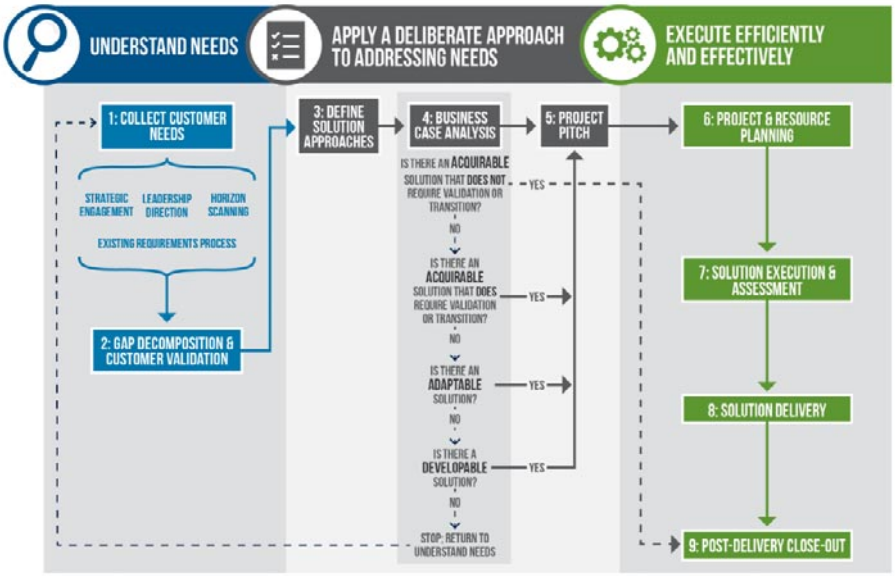


Figure 1: S&T Operating Model Blueprint and Business Process Flow

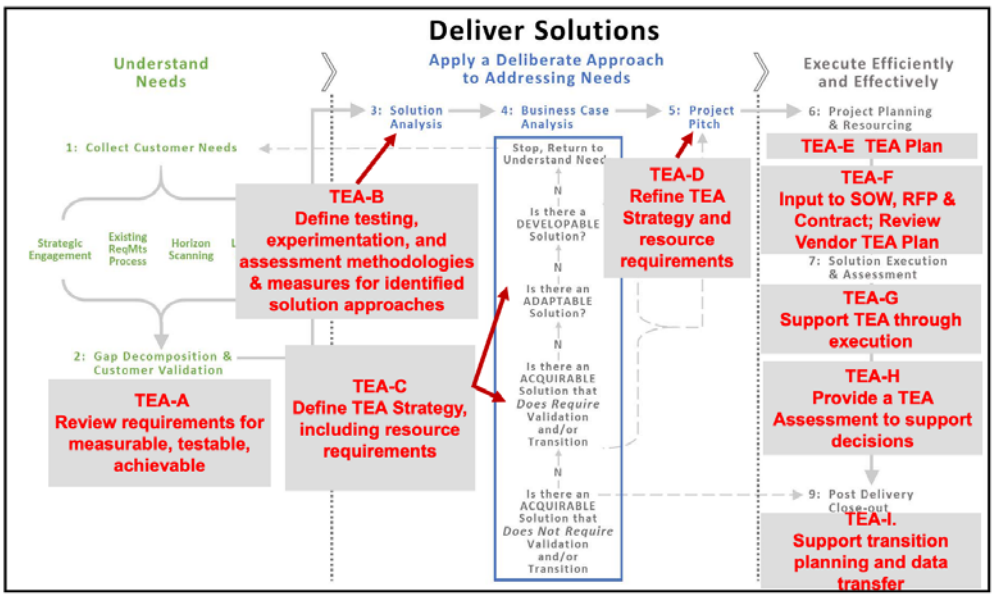


Figure 2: TEA Touchpoints Overlayed onto BPF Processes

Business Process 3, Define Solution Approaches. As the gap is decomposed into its constituent needs for the user, it is essential that the requirements are validated before proceeding further. Similarly, as the solution options are being defined, the requirements should again be validated with an emphasis upon achievability and testability. Testers can provide an independent assessment to the program manager.

TEA B: Define TEA Methodologies and Measures

The Define Testing, Experimentation, Assessment Methodologies, and Measures touchpoint occurs during Business Process 3, Solution Analysis. The testers assist the program manager with analysis and prioritization of TEA options and provide information on Modeling and Simulation (M&S) capabilities and test facilities. As noted above, the analysis of Touchpoint A may be repeated to account for new information that may have been developed. The advantage to identifying the options for testing while solution options are being examined is that the testability and achievability are established at the outset. This analysis ensures that the solution options for the gap are viable.

TEA-C: Define TEA Strategy

The TEA Strategy is initially defined during Business Process 4, Business Case Analysis. It provides a description of the testing activities planned for the R&D program to verify that its objectives have been met. The TEA Strategy also provides the basis for the assessment that justifies its transition to an acquisition program.

The TEA Strategy may take many forms. It could be a briefing, a standalone document, or embedded in the Program Management Plan. Accordingly, the content outline addresses the topics and content to be addressed in the TEA Strategy; however, the format of the strategy should conform to the template for the selected medium of delivery.

The content of the TEA Strategy is driven by the type of solution being developed and the transition that is envisioned for the solution. Three types of solutions are considered: those that are being developed from scratch, those that are adaptations of existing products, and those that are procurements of available commercial or governmental products. For solutions that are being developed from scratch, there are three possible outcomes: transition to an acquisition program, initiation of a procurement, or transfer to a commercial vendor for production and general distribution. Each of the foregoing imposes specific requirements on the TEA Strategy and its execution.

TEA D: Refine TEA Strategy

The TEA Strategy is refined in Business Process 5, the Project Pitch. At this point, the preferred solution is presented to the customer. The TEA D activity consists of refining the TEA Strategy with more specificity, particularly in the areas of resourcing and costing for TEA activities. With the completion of this Business Process, the program transitions to execution; and the TEA strategy serves as the basis for the program and/or vendor development of a detailed TEA Plan.

TEA-E: Document the TEA Plan

Documentation of the TEA plan starts at Business Process 6, Project Planning and Resourcing, and continues through the execution phase of the Business Process. The TEA Plan may be written by the government or a contractor (or both collaboratively). If the plan is written by a contractor, it must be approved by the government. The TEA plan provides a detailed blueprint for acquiring, testing, and assessing the solution; it also provides input to the SOW, RFP, and contract, if applicable. Additionally, the plan defines how and when TEA monitoring will be performed throughout the development process. Testers on the program staff ensure that the TEA Plan is complete and executable; they also ensure that the costing, resourcing, and scheduling are reasonable.

Test plans for individual events are typically prepared by the activity performing the testing and are approved by the program manager. As part of the TEA process, the assigned TEA Test Area Manager reviews the test plan and provides recommendations to the program manager regarding proper test procedures, data collection, and reporting.

TEA-F: Support Execution of the TEA Strategy / TEA Plan and Input to SOW/RFP/Contract

TEA F is the second touchpoint in Business Process 6. This activity is the implementation of the TEA Strategy and/or government-developed TEA Plan. It may include assisting in the development of the Statement of Work (SOW), Request for Proposals (RFP), or contract to specify the TEA requirements. A requirement for a vendor TEA Plan and Strategy should be included as a requirement in the SOW or RFP. As requested, the testers may provide specifications and guidance for development of the TEA Plan. TEA personnel may also review contractor submissions for compliance with the SOW or RFP.

TEA-G: Support TEA Through Execution

TEA-G is the first TEA touchpoint in BPF 7, Solution Execution and Assessment. Periodic program-level and portfolio-level reviews are performed to assess progress, and TEA data on solution capabilities and performance are gathered and reported. Testers participate in reviews and provide feedback related to TEA and will also collect TEA data as the project progresses. The testers also provide independent assessments of TEA activities and progress along with feedback to the PM.

TEA-H: Develop a TEA Assessment to Support Decisions

TEA-H is the second TEA touchpoint in BPF 7. The results of the process form the input to the decision at which the customer accepts the solution. Testing is conducted to collect data for analysis to validate that the product meets the customer requirements in the intended environment.

Testers assist the PM in the development of the TEA Summary. The summary provides a concise overview of the testing, experimentation, and assessment data collected throughout the development/adaptation process. It concludes with an assessment of the solution's capabilities and limitations for its proposed use.

TEA-I: Support Transition Planning and Data Transfer

Following customer acceptance of a solution, the product may be transitioned or commercialized to industry. Testers support the transition planning and transfer of test and assessment data. The TEA-I touchpoint is part of BPF 9, Post-delivery Close-out.

All relevant test, experimentation, and assessment data should be transitioned to the industry with the solution product to ensure successful turnover and continued maintenance, and to facilitate further assessment including RAM projections. The data also support the analysis associated with manufacturing processes.

TEA Support Examples

Both the Business Process Flow (BPF) and the Testing, Experimentation, and Assessment (TEA) initiative are new processes within the DHS Science and Technology Directorate. This period of adaptation has permitted the testing support to be integrated into the matrix organization for project execution.

TEA provides two principal areas of support to program managers. The first is the review of test plans and programmatic documents. Testers were able to provide guidance in the areas of objectives, test procedures, data collection, and support requirements. Testers also review the requirements for proposed projects and recommend clarifications to make them more testable.

The second area of support is the development of TEA Strategies. For R&D projects, the TEA Strategy takes the place of the T&E Master Plan (TEMP) and provides an outline of the testing to be conducted to demonstrate the requirements of the R&D project. The strategy is the basis of the test plans developed for the project. They may include plans for Critical Experiments, Characterization Experiments, and Demonstrations.

As an example, the Department, in response to national policy, has been asked to investigate the use of electric vehicles for law-enforcement work. The task was passed to the Federal Protective Service. While they have experience in vehicle acquisition and maintenance, the introduction of electric vehicles was a new challenge. Participating in their integrated product team (IPT), the Test Area Managers developed the TEA strategy for the testing of the prototype vehicles, which formed the basis for the program test plans. They were also able to establish liaisons with other DHS Components and the Science and Technology Directorate that were undertaking related R&D initiatives. These included vehicle cybersecurity and environmental hazards of lithium-ion batteries. By exploiting the synergies of the projects, the overall R&D effort is more effective, and duplication of efforts is avoided.

Since most R&D projects do not have a T&E Manager, the experienced testers from the T&E Division provide valuable support to the R&D program managers. In many cases, program managers only need to consult with a T&E subject matter expert regarding their test program; test area managers have been available to provide this service.

On several occasions, program managers have requested the names and contact information for facilities that could support their test programs. In providing recommendations of available facilities, the test area managers also can facilitate access, through the T&E Division's inter-service agreements, to government labs that can provide T&E services to R&D programs.

Similarly, program managers want someone to work with their support contractors and help develop the test documentation. In those cases, test area managers review documentation and provide guidance regarding test methods, test conduct, safety considerations, permitting, and coordination of test activities with outside agencies.

Moreover, a TEA workshop has been developed to provide an introduction to the TEA process and show how it can support the program manager and staff. The pilot offering was in February 2021. There have been two subsequent offerings with a total of sixty-seven graduates. Going forward, the workshop will be offered twice each year.

Future Development

At the present time, the TEA team has been included as a member of all the R&D matrix teams. This means that, going forward, testing considerations can be addressed in the early stages of program planning and execution. Many program managers do not have experience in test and evaluation; thus, they are glad to have experienced testers available to augment their program staffs. Through the S&T matrix organization, testers can be made available when needed to support the programs as they pass through each process of the BPF.

While functionally separate from the T&E staff supporting acquisition programs, the TEA staff can provide liaison for the PM and facilitate the transition of R&D programs to acquisition.

A short course is being prepared in addition to the TEA workshop to provide R&D managers and staff an introduction to the testing considerations for their programs. The course will introduce the analyses and findings that the testing should provide for the successful demonstration of an R&D project. Additionally, the course will illustrate the testing results that are necessary to transition from R&D into an acquisition program.

Conclusion

The Department of Homeland Security has recognized the need to provide testing support to research and development programs. A team has been established in the Test and Evaluation Division to provide testing support for R&D program managers. The test team has been integrated into the matrix teams for the R&D programs. Support requests have included document reviews, requirements reviews, and the development of TEA strategies. The integration of expanded R&D testing techniques is well underway and promising results are visible on the horizon. ❑

MICHAEL LEITE is a Test Area Manager for ManTech where he supports Testing, Experimentation, and Assessment of research and development programs of the Science and Technology Division of the DHS. Previously, he was a Training Instructor Supervisor teaching the test and evaluation courses for DHS. In previous assignments, he was a T&E Subject Matter Expert supporting the Coalition Branch of Joint Interoperability Test Command (JITC). He has provided T&E support to the DoD Director of Test and Evaluation and the Army Test and Evaluation Command. His T&E experience also includes developmental testing for the Naval Sea Systems Command and the AEGIS Shipbuilding Program and operational testing for the PATRIOT Missile Program. A retired Navy Captain and a Registered Professional Engineer, he has previously supported modeling and simulation efforts of the Defense Modeling and Simulation Coordination Office, Missile Defense Agency, DD-21 Gold Team, Cooperative Engagement Capability Program, and the Joint Theater Air and Missile Defense Organization.

The ITEA Journal of Test and Evaluation 2022; 43: 224-230

Measuring the Measurers: Using Test to Validate Cyber Risk Assessments

William D. "Data" Bryant, Ph.D.
Modern Technology Solutions, Inc. (MTSI), Alexandria, VA

Cyber risk assessment processes claim to be able to predict the success of cyber attacks against weapon systems and platforms in a range of highly contested cyber environments, but what evidence do we have that they are more accurate than random guesses or tea leaves? Fortunately, on systems that are also undergoing cyber testing, we can leverage this test to validate that the risk process is effective by having risk assessors also provide predictions on what test cases will be successful and comparing the predictions and test results.

Introduction

The threat of cyber attacks on our critical weapon systems and platforms continues to increase, and there is still no consensus on which risk assessment methods are most effective at predicting the performance of a weapon system or platform in its expected cyber-contested operating environment. Juvenal, a Roman poet, famously asked, "*quis custodiet ipsos custodes?*" or "who will watch the watchmen?" If we think we have a way to measure the risk of cyber attack on an aircraft, platform, or other Cyber-Physical System (CPS), our similar question should be, "who will measure the measurers?" or how do we know the risk assessment is correct? Without validation that a risk assessment approach works, we may be making critical programmatic and design decisions based on something with no more validity than reading crystal balls or sheep entrails.

In some other types of risk assessment, historical data is plentiful, and we can use it to determine a reasonable set of probabilities for both likelihood and impact. Floods along the Mississippi, tornadoes in Oklahoma, and hurricanes in Florida are all extremely difficult to predict far in advance. However, the probability over time of them striking a particular area and the amount of damage done when they do can be predicted accurately from historical data. There is enough historical data that we are confident enough in our probabilities to use them to inform major decisions, such as facility construction.

So far, cyber attacks on CPSs have been rare and rarely reported if they do happen. In addition, cyber attacks are a modern phenomenon, and we have widely understood them as a potential problem for CPS for less than a decade, so there is little historical data from which to draw. Making things even worse, the typical design of cyber attacks keeps them hidden and their effects active for as long as possible [1].

Despite these issues, most cyber risk assessment methodologies promise to provide risk measurements accurate enough to be used as the basis for programmatic and design decisions. But how do we know a particular risk assessment process works without an extensive historical record to compare it against? Fortunately, there is a way to "measure the measurers" for those programs that also accomplish cyber testing. If a risk assessment process claims to predict the success of unconstrained nation-state cyber attackers, it should also be able to predict the success of constrained test teams running known test cases. Following the completion of the cyber testing, we can compare the cyber analysts' predicted results to the actual results of the testing. If there is a high correlation between the predictions and the results, that should add confidence that the risk assessment process is producing actionable results; if there is a low correlation, it should reduce confidence in that process and might suggest the need for adjustments in the risk scoring for the broader risk scenarios.

Mission Based Cyber Risk Assessments (MBCRA)

MBCRA's are defined as a "process of identifying, estimating, assessing, and prioritizing risks based on impacts to DoD operational missions resulting from cyber effects on the system(s) employed [2]." Per DoDI

5000.89, DoD program managers will conduct MBCRAs as part of their cybersecurity test and evaluation [3]. There are a large number of MBCRA approaches available; one 2017 Institute for Defense Analysis study identified more than 20 methodologies and more have been developed since [4]. One of the most common approaches utilized across a wide range of DoD programs is the Cyber Table Top, the details of which can be found in the Cyber Table Top Guide [5].

Another MBCRA process that we have used successfully on a number of systems with small teams and limited resources is the Unified Risk Assessment and Measurement System or URAMS® [6]. It is a collaborative methodology that provides a diverse set of integrated qualitative and quantitative tools to provide risk management for weapon systems and aviation platforms throughout the development life cycle and across a range of contested cyberspace environments.

URAMS starts with an engineering analysis, and our preferred tool is Systems-Theoretic Process Analysis for Security (STPA-Sec). The development of this tool leveraged the safety analysis work done at MIT, and a range of military weapon systems and civilian aerospace systems use STPA-Sec to significant effect [7]. STPA-Sec is grounded in systems engineering and is focused on mission-level losses as the true drivers of relevant security design. STPA-Sec also enables analysis of a system's security posture early in the life cycle, enabling true "baking in" of security.

From the analysis, a set of risk scenarios are developed that are specific to the system under consideration and its expected operating environment. Then, those risk scenarios are scored using any of a wide range of available scoring tools. URAMS scoring tools are characterized first by the risk model and what factors are assumed to contribute to overall risk and second by the input type. Inputs can be provided as single-point values, single-point values with confidence, three-point estimates, or 90% confidence intervals. Selection of input type depends on the training and experience of the assessors, as well as how important uncertainty is to the decision makers. While human subject matter experts (SMEs) are utilized as the basis for scoring in URAMS, automated and algorithmic-based approaches can and should be used to inform those SMEs.

The risk scenarios can then be combined utilizing a simple Monte Carlo simulation to determine the overall risk for a system or portfolio of systems. Combining risk facilitates building a structured assurance case that includes the analyzed mission structure connected to the specific risk scenarios and their scores, which flow up through the mission elements to the overall system. Most importantly, specific evidence, such as testing results and design features, can also be added to the assurance case to validate the risk scores. The assurance case is presented in a format that allows decision makers to rapidly assess whether the scoring is reasonable, based upon their understanding of the mission and the evidence provided.

Cyber Test Plans

Whatever specific method is used, MBCRAs should be executed iteratively, and their results should inform test planning [8]. A list of scored mission-based risk scenarios can provide focus and prioritization for the test cases that are built. Testing the most mission-significant

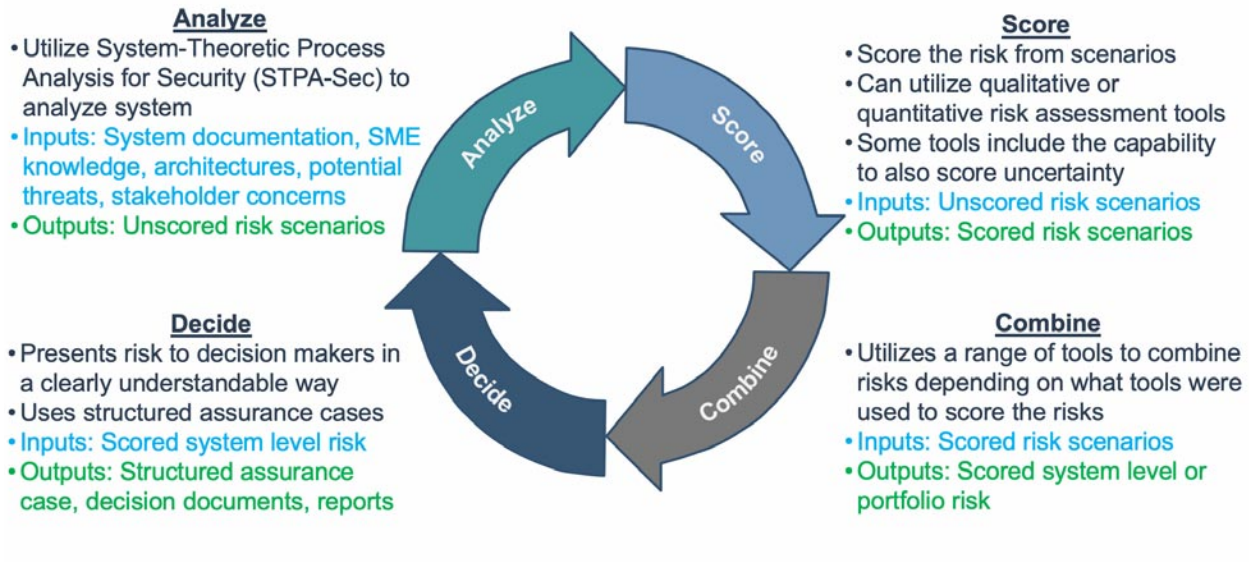


Figure 1: The Unified Risk Assessment and Measurement System (URAMS)

risk scenarios will help focus limited test resources on those attacks that may have the greatest mission impact and provides a clear linkage between cyber test results and mission.

When MBCRAs inform cyber test results, senior decision makers can be provided with the mission impacts associated with the various tests in place of vulnerabilities, as vulnerabilities are often hard to connect directly to mission impact. Using MBCRAs to help focus and improve cyber testing is a critical step forward and is one that forward-leaning programs are already executing. The next step forward is to close the feedback loop from test back to risk assessment and utilize that same testing to measure how well the MBCRA predicted risk.

Measuring the Measurers

MBCRAs, whether they generate their results from human assessors or risk algorithms, claim to be able to assess the level of risk that a cyber adversary will be able to affect a friendly mission. If these approaches can achieve some level of success against this very difficult and unconstrained problem set, we can assume that they should be able to achieve at least a similar, if not better, level of success predicting how successful a test team will be in affecting a friendly mission, given that the MBCRA analysts know exactly what attacks will be launched from where, and what resources will be put against the problem as given in the test plan.

While the details and milestones will vary depending on the program and acquisition pathway utilized, Figure 2 shows a general flow of the process below.

From early conceptual design to operations, mission and system design work starts the process wherever a system is on the design life cycle. The second step is to execute whatever risk analysis process the program has selected. To be able to derive the maximum benefit from prediction and prioritization of risk, the risk assessment process must score the risks in some way, whether that is through quantitative or qualitative scoring methods [9].

As noted above, test planners can utilize those scored risk scenarios in the fourth step to perform test planning and select which tests will be executed. However, not all the highest-ranked risk scenarios may be tested early in the program due to constraints such as test resource availability or safety. As a result, some high-risk scenarios may only be tested after resources such as a full Systems Integration Lab (SIL) are developed, or they may be tested in segments without demonstrating the full mission effect.

So far, nothing in the proposed flow is different from what forward-leaning programs do today. However, the fifth step of sending the test plan back to the risk assessors is new and is where the feedback loop is closed. The risk assessors then take the test cases and score them using their understanding of the system and whatever models or algorithms they already built as if the proposed tests were cyber attacks against the system under test. Since they will have far more detailed information and less uncertainty from test cases than potential adversary attacks, they should then be able to provide predictions on the results of the proposed testing in the test plan.

Understanding and Presenting Results

After completing the testing, the predictions can be compared to the actual test results to show how well the risk assessment process performed. Although there

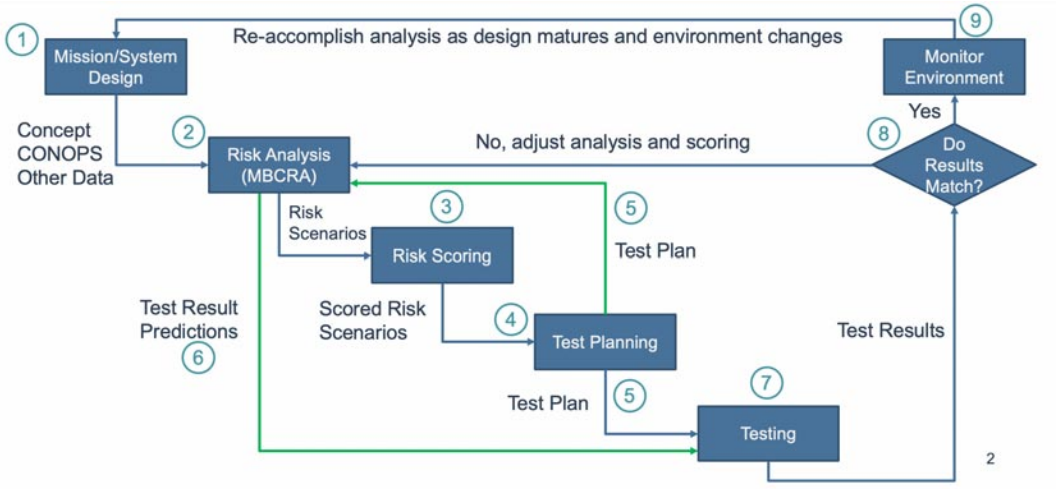


Figure 2: Draft Risk Assessment Process Flow

are many potential approaches to accomplish this, the method illustrated below requires that each test case prediction be scored as a binary "hit" when correctly predicted and a "miss" when not. Predictions should be simple one or two-sentence descriptions of the projected outcome. If desired, a single test case could be broken into several pieces with different predictions, such as whether testers can access a particular component and what effect they can generate. The testers or some other group of analysts should score whether the predictions were correct or not versus the risk analysts, who may be more likely to score their results successfully due to well-established confirmation bias that may influence their ratings.

With a set of test predictions scored as "hits " and "misses," utilizing cumulative probabilities from a binomial distribution in small segments from 0-100% a curve can be created that shows which potential values of accuracy for a risk assessment are more likely in a range from 0-100% [10]. Note that the total area under the curve is 1.0 or 100%. Figure 3 below shows an example assessment process that predicted ten test case results and got seven of them correct.

In this case, the highest point on the curve, or the mode, is 75%, not 70%, due to the shape of the binomial distribution as it becomes more skewed the further the mode is from 50%. However, after only one test, the uncertainty is high and the 90% Confidence Interval (90CI), within which we have a 90% confidence that the true value lies, is 49-91%.

If we then accomplish another trial of this risk assessment process on a different system, we will develop a separate set of risk scenarios based on the new system and mission. We can then utilize those risk scenarios to inform test on the second system and again predict what results are expected from that system's cyber testing. Using those results we can then update the probability curve in Figure 3 to incorporate the data on how well the risk assessment process performed in assessing this second system. We perform this update using Bayesian statistics, which provide a set of tools by which we can update an existing state of knowledge given additional data [11]. For our example, in the second trial, the risk assessment process correctly predicted the results in eight out of ten test cases, as shown in Figure 4.

The combined curve has a maximum at 79% and a 90CI of 59-89%. The uncertainty in the predicted success of our risk assessment process decreased significantly after the second trial, evidenced by the 90CI width narrowing from 42% to 30%. Further trials with comparable results would continue to decrease the uncertainty, while trials that generated significantly different results would increase the uncertainty.

While these results are useful to program managers and engineers, it is potentially even more useful to utilize

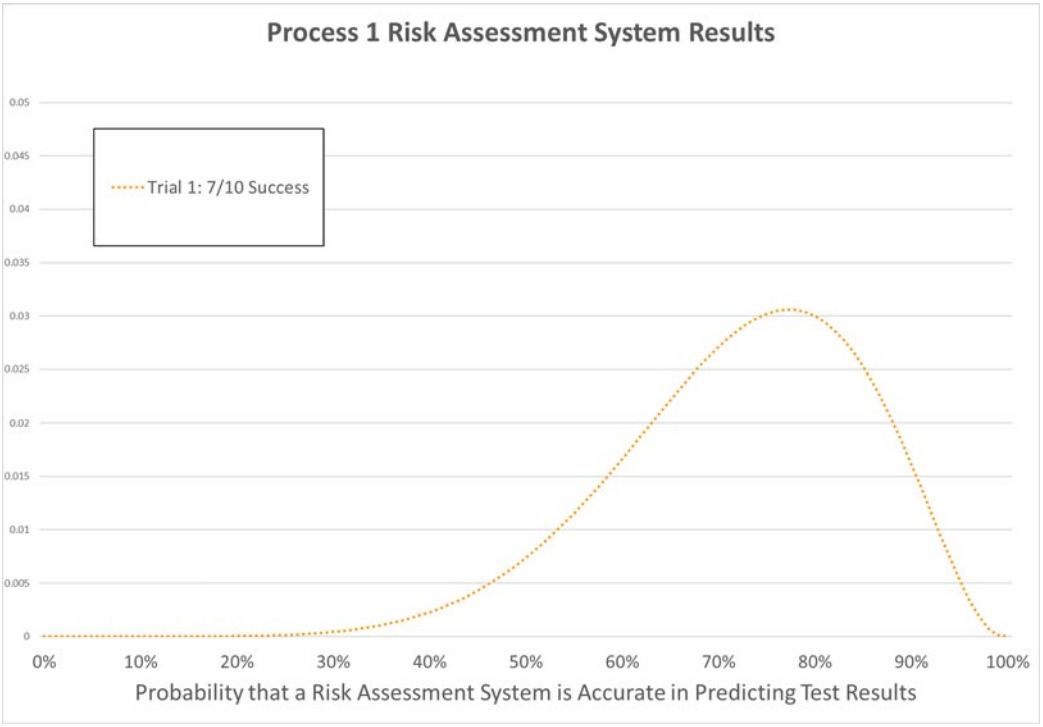


Figure 3: Process 1 Example Risk Assessment Process Results

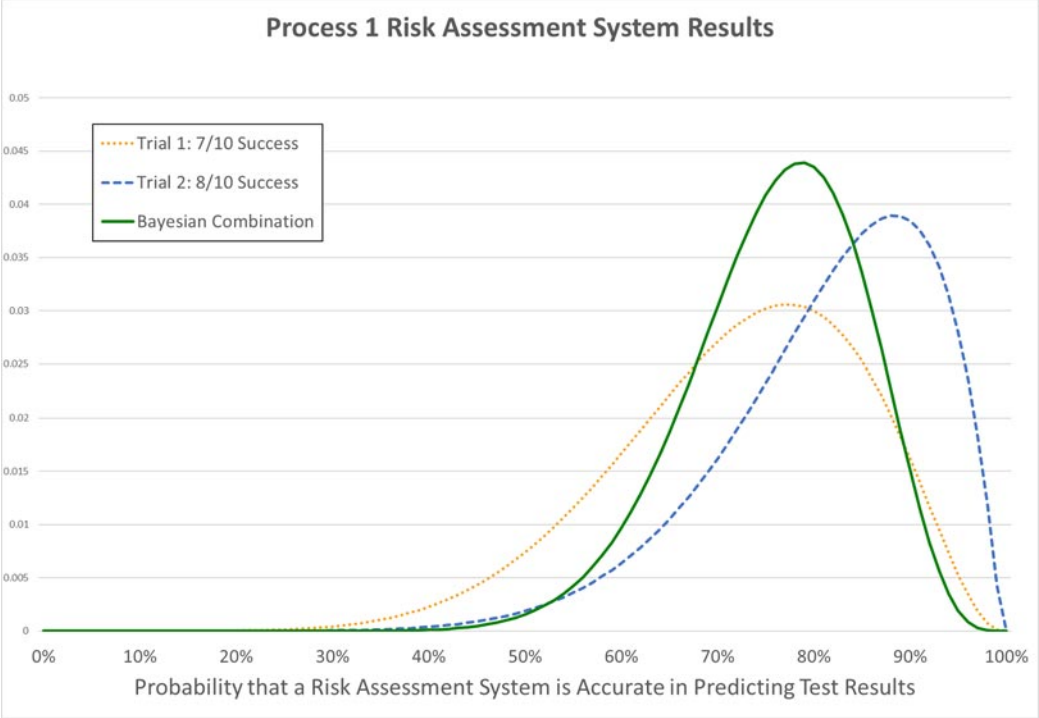


Figure 4: Process 1 Example Risk Assessment Updated with Second Trial

this approach to compare two different risk assessment processes under consideration. For example, if a second risk assessment process also did two trials with ten test cases, but only correctly predicted three results in the first trial and four results in the second trial, we could plot both combined curves together as in Figure 5.

There is only a small area where the two curves overlap and where it is possible that risk assessment process 2 performs better than risk assessment process 1. A quick Monte Carlo simulation showed less than a 1% probability of process 2 outperforming process 1 [12]. Thus, a program manager can be confident that it is highly unlikely that risk assessment process 2 will outperform risk assessment process 1. As a result, they should select risk assessment process 1 for future work, assuming other factors, like cost, are similar.

Another way to leverage this comparison approach is to utilize a "control group" and compare a proposed risk assessment approach to whatever group within, or outside, of the program that is working in this area. For example, if the program has a group executing the Risk Management Framework (RMF), those experts could be asked to score which test cases they think will be successful based upon their understanding of the system developed while executing the RMF. If their scoring is as good or better than a proposed risk assessment process, then utilizing those experts to score risk scenarios may make more sense than standing up a separate risk assessment group and process. If a separate risk assessment process is more effective, a chart like Figure 5 can show how much more effective and program managers can determine if the additional cost and resources provide enough additional value.

Conclusions

Cyber risk assessments and cyber testing should be closely linked. Cyber risk assessments can help focus and inform what test cases are selected for execution, while cyber testing provides an opportunity to validate or invalidate the effectiveness of cyber risk assessments by having the cyber assessors also predict the outcomes of the test cases before running the tests. These results can then be used to compare various risk assessment processes to determine which ones programs should utilize.

One of the major advantages of utilizing cyber test results to determine which risk assessment processes are more effective is that it leverages work already being performed in most cases instead of implementing something new. A program that will do both a cyber risk assessment and cyber testing needs to connect the two processes and form a feedback loop between them to help determine how well correlated they are.

Because risk assessors could get either "lucky" or "unlucky" if utilizing very few data points, the

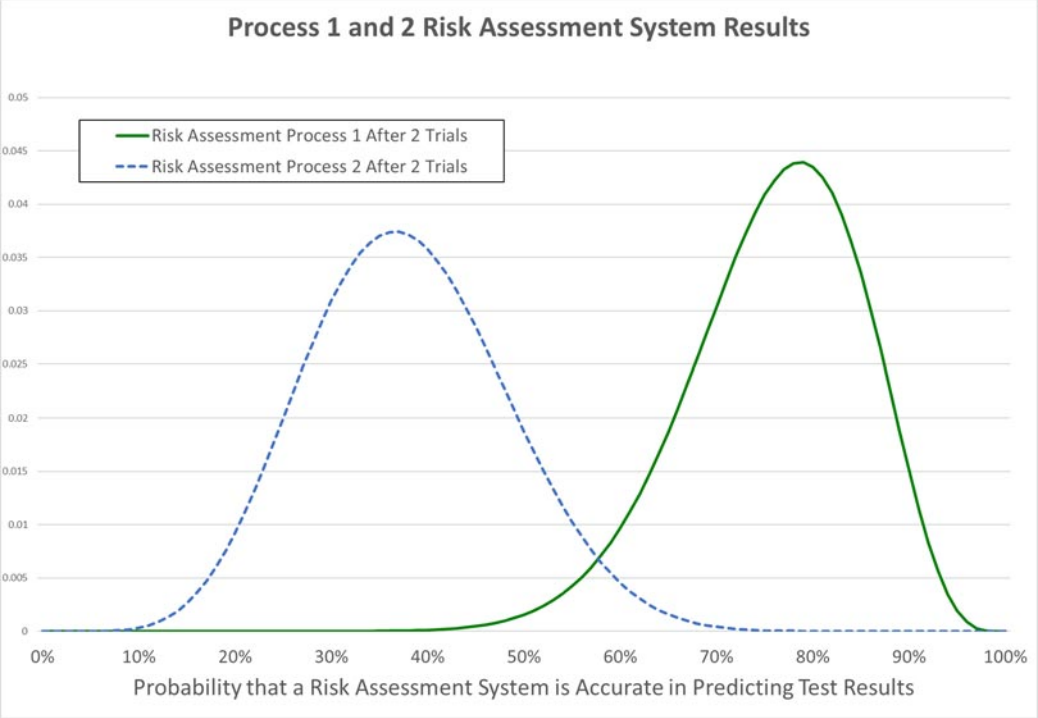


Figure 5: Risk Assessment Process 1 and 2 Compared

uncertainty of the results should be considered. Using a Bayesian statistical approach and updating the prior distributions with new data enables a small number of measurements, such as the twenty test case comparisons in the examples above, to clearly demonstrate which of the two processes was likely to be more accurate. Of course, note that what we are measuring is the ability of a risk assessment process to predict the outcome of test cases, and we are assuming that this will at least correlate with the risk assessment processes' ability to predict actual cyber attacks. This assumption seems reasonable, although we expect that the accuracy of predictions for actual cyber attacks will be lower than for test cases since there is less data available and more uncertainty. Even so, the key assumption is that a risk assessment process more successful at predicting test cases will also be more successful at predicting the outcome of cyber attacks.

Following this approach and leveraging ongoing cyber testing to measure the effectiveness of our risk assessment measurement tools would represent a significant step forward in integrating cyber risk assessment and cyber testing on cyber physical systems such as aircraft and weapon systems. This improvement would help bring risk assessment out of the purely theoretical academic world into the empirical world, where we can utilize the data to improve performance. ❑

WILLIAM D. "DATA" BRYANT, Ph.D. is a cyberspace defense and risk leader who works for Modern Technology Solutions, Incorporated (MTSI). His diverse background in operations, planning, and strategy includes more than 25 years of service in the Air Force, where he was a fighter pilot, planner, and strategist . He holds multiple degrees in aeronautical engineering, space systems, military strategy, and organizational management. He has also authored numerous works on various aspects of defending cyber physical systems and cyberspace superiority, including International Conflict and Cyberspace Superiority: Theory and Practice.

References

[1] William D. Bryant, *International Conflict and Cyberspace Superiority: Theory and Practice* (London, Routledge, 2016), pp. 171.

[2] Department of Defense, *Cybersecurity Test and Evaluation Guidebook*, Version 2.0, 25 April 2018, pp. 12.

[3] Department of Defense, DoD Instruction 5000.89, Test and Evaluation, 19 November 2020, pp. 16.

[4] M. A. Ambroso and R. T. Hutton, *Comparative Review of DoD Mission-Based Cyber Risk Assessment Methodologies*, Institute for Defense Analysis, December 2017.

[5] Department of Defense, *The Department of Defense Cyber Table Top Guidebook*, 16 September 2021, Version 2.0.

[6] William D. Bryant, *The Unified Risk Assessment and Measurement System (URAMS) for Weapon Systems and Platforms: Cutting the Gordian Knot*, version 2.0, Available: Modern Technology Solutions Inc., 2022, www.mtsi-va.com/weapon-systems-cybersecurity/.

[7] For a description of STPA, on which STPA-Sec is based, reference the STPA Handbook that can be found at http://psas.scripts.mit.edu/home/materials/. For a description of STPA-Sec created by Dr. William Young see http://psas.scripts.mit.edu/home/wp-content/uploads/2020/07/STPA-Sec-Tutorial.pdf.

[8] Department of Defense, *Cybersecurity Test and Evaluation Guidebook*, Version 2.0, 25 April 2018, 8.

[9] Note that many risk scoring processes, such as the typical risk cubes scored from 1-5, utilize ordinal scoring methods that cannot be legitimately combined via arithmetic operations such as multiplication.

[10] The method utilized in this section was adapted from Douglas W. Hubbard, *How to Measure Anything: Finding the Value of Intangibles in Business*, 3rd ed. (Hoboken, NJ: Wiley, 2014), 267-272. In his example, he utilized a normal distribution as the prior, in this case the prior was a binomial distribution. Reference Hubbard's chapter 10 for a fuller description of the theory behind the process provided here.

[11] For a description of Bayesian statistics see Douglas W. Hubbard, *How to Measure Anything: Finding the Value of Intangibles in Business*, 3rd ed. (Hoboken, NJ: Wiley, 2014), 247-284.

[12] A 10k element Monte Carlo simulation was run on the above distributions and 99.3% of the time process 1 outperformed process 2, so there is only approximately a 0.7% chance that process 2 will outperform process 1 on any analysis.

Northeast Region

Massachusetts
New England Chapter
Robert Reyling, President

New Jersey
South Jersey Chapter
John Frederick, President

Ohio
Miami Valley Chapter
President—Vacant

Pennsylvania
Penn State Chapter
Bruce Einfalt, President

Mid-Atlantic Region

Maryland
Francis Scott Key Chapter
Robert McKelvey, President

Southern Maryland Chapter
President—Vacant

DC/Northern Virginia
George Washington Chapter
Rick Bailer, President

Virginia
Hampton Roads Chapter
Erwin Sabile, CTEP, President

Southeast Region

Alabama
Rocket City Chapter
Lewis Hundley, President

Florida
Central Florida Chapter
President—Vacant

Emerald Coast Chapter
Kate Snow, President

Georgia
Atlanta Chapter
Joseph Hurst, President

South Carolina
Charleston Chapter
President—Vacant

Tennessee
Volunteer Chapter
President—Vacant

West Region

Regional Vice President
Terrance McKearney

California
Antelope Valley Chapter
Wendy Peterson, President

Channel Islands Chapter
Gil Torres, President

China Lake Chapter
Greater San Diego Chapter
Presidents—Vacant

Hawaii
Mid-Pacific Chapter
Shannon Wigent, President

Washington
Pacific Northwest Chapter
President—Vacant

Southwest Region

Regional Vice President
David Webb

Colorado
Rocky Mountain Chapter
President—Vacant

Arizona
Huachuca Chapter
President—Vacant

Valley of the Sun Chapter
Steve Woffinden, President

Nevada
Southern Nevada Chapter
Darryl Johnson, President

New Mexico
Roadrunner Chapter
Ralph Galetti, President

White Sands Chapter
Steve Aragon, President

Utah
Greater Salt Lake Chapter
President—Vacant

International

Regional Vice President
Peter Nikoloff

Australia
Southern Cross Chapter
Peter Nikoloff, President

Europe
European Chapter
Adrian Britton, President

Israel
Israel Chapter
President—Vacant

Corporate Member News

New Corporate Members

The ITEA leadership would like to recognize our new Corporate Members.

DigiFlight
We Accelerate Mission Success
DigiFlight applies an interdisciplinary approach to overcoming today's acquisition and technology management challenges. Using proven principles, our experienced professionals support all of the planning, development, testing and deployment tasks associated with building new systems, updating legacy technology, and integrating existing tools and equipment.

Radiance Technologies
Radiance is a proud employee-owned company!

Employee- ownership has benefits to our customers and workforce. We are singularly focused on providing our customers with game-changing solutions to their most challenging requirements.

Radiance Technologies develops innovative solutions for defense, intelligence, and civilian customers' advanced challenges. We bring together people with remarkable technical abilities and deep knowledge of customer missions, then equip them with modern laboratories, tools, and facilities. Our experience spans the directed energy, hypersonics, space, aviation, test and evaluation, intelligence, microelectronics, cyber, and critical infrastructure markets.

❑ ❑ ❑

ACRO – BOEING ST LOUIS

Acroamatics has delivered the next generation of lunchbox Ground Support Units (GSU) to Boeing St Louis in support of upcoming flight test projects. The Model 3022AP is an industry leading portable telemetry acquisition and processing solution.

Amongst new capabilities delivered in these updated MD3022AP units is an added standalone battery capable of 4+ hours of multi-stream system operation.

Other new capabilities include integrated dual and triple 24-in LCD displays, multiple RF receiver capacity, and turn-key enhanced Chapter 7 and enhanced networked (TMoIP) telemetry processing features.

❑ ❑ ❑

Wideband Systems, Inc
Advanced Signal Recording Systems Selected by The Navy

After years of research and evaluating numerous TM reorders the NAVY has chosen WSI Wideband DRS8500X Recorders for a major system upgrade, replacing several other previous generation recorders.

Wideband Systems, Inc., is the global leader in ground based advanced signal recording systems. WSI recorders are deployed on every major US range as well as abroad. Our Wideband systems provide flexible interfaces (direct IF and RF recording, pre-detect analog, post-detect digital PCM, as well as Video, Ethernet, Serial, and 1553 recording, …). WSI recorders employ the latest technology, including un-matched performance (1600Mb/sec to 9600+Mb/sec), simultaneous record, export (same mission or any other mission), CH10 Publish capability, Simultaneous Playback and Publish, Cued Trigger Playback, Dual RAID configuration (User configured for Mirrored or Striped Mode), Direct Record to any external attached target drive, Remote Control via GigE post and many other advanced field-requested features.

With strong commitment to our customers, WSI provides unmatched technical support and industry-leading two-year warranties. Additionally, all WSI recorders include lifetime no-cost software and firmware updates and technical support.

Wideband Systems joined the Delta Information Systems (DIS) family in 2021. The combination of the companies has allowed DIS to provide state of the end-to-end Telemetry systems. The relationship affords WSI a wealth of technical support and resources, thus allowing us to offer new superior features, and bring new products to market in short order.

Chapter News

Emerald Coast Chapter

The Emerald Coast Chapter hosted the 8th Annual Cybersecurity Workshop from Oct 18-20 in Miramar Beach, Florida. The theme of the event was "New Domains in Cybersecurity T&E." The event format included tutorials on Tuesday followed by the workshop on Wednesday and Thursday. With over 130 registrants, our attendees

had the unique opportunity to engage with many of the guest speakers during breaks and networking reception during the week. The event featured pre-workshop tutorials, guest speakers, 32 technical presentations, and a hands-on Red Team Blue Team Exercise that ran during three separate sessions to allow for maximum participation. When a Chapter hosts an event, it is the volunteers who step up to provide the planning and execution and are the ones behind the scenes making sure the event is seamless. We acknowledge those who worked the registration desk, the audio-visual support, and the photographer. Well done, friends, well done.

Registrants arrived early to attend the pre-workshop tutorials to receive educational credits for their professional portfolio. They spent 4 hours learning about Cybersecurity Assessment of MI-STD-1553 with **Adam McCorkle**, GTRI; A Process for Distributed LVC in T&E by **Michael O'Connor**, TRIDEUM Corporation; Intro to Cyber Resilience T&E and T&E in a Digital Engineering Environment by **Jean Petty**, Department of Homeland Security; Cybersecurity Solutions with JMETC, TENA, and TRMC BDA by **Gene Hudgins**, TRMC/KBR; and ARCUS Cloud – Cyber Tools and Ranges by **Peter Walsh**, Jackpine Technologies Corporation.

The Chairman of ITEA, **Bruce Einfalt**, kicked off the workshop offering a glimpse into what membership means to the T&E community. One of the benefits is receiving recognition from ITEA Test & Evaluation Professional Awards program and what better way to explain what this program means than to call up **Shelby Pierce**, the Chair of the Cyber workshop. In

Chairman, Bruce Einfalt presenting Shelby Peirce with Energizer Award

September at the ITEA Annual Symposium, Shelby was unable to accept the award in person due to a conflict in her schedule and was presented the Association's Energizer Award for her "behind the scenes" contributions over the past year orchestrating the Cybersecurity Workshop. Her leadership skills along with her positive energy inspired her team of volunteers to work alongside her as she cultivated a technical program that gained momentum and proved to be highly successful. Technical Chair and her partner in designing this program, **Andy Overbay**, was the emcee for the week and introduced our speakers.

One of the recurring themes from several of the speakers was how to test enough in a resource- and time-constrained environment. A few even discussed the fact that the education system will not produce cybersecurity experts fast enough to meet the growing demand. The test community is no exception to this. Automated testing and testing integrated with development were also discussed as ways to improve the amount of testing that is done without compromising schedules. Another recurring theme was data management. Several speakers discussed the fact that data is produced in huge volumes, but it isn't managed and exploited well.

The first speaker, **James Wells**, the Director of Test and Evaluation, Science and Technology Directorate for the DHS, discussed some of the challenges of Operational Test (OT) at DHS. He said most of the units "being deployed" (doing their primary mission) all the time made OT difficult and often OT had to be done with the unit while it was executing its mission. He advised the test community not to focus on the tasks for testing, but instead to focus on the effects on the mission [if something fails to perform]. He advised that testers should ensure systems and capabilities are secure and resilient as designed – built – and operated.

The Cybersecurity Technical Director, Office of the Undersecretary of Defense for Research and Engineering, Developmental Test, Evaluation, and Assessments, **Sarah Standard**, emphasized the need for modernized iterative T&E methodology. She discussed the need for DT and OT to collaborate to speed up testing and the importance of testing because all of it supports program decisions at many levels.

Dr. Mike Shields, the President of Vigilant Cyber Systems, Inc. and current Test Resource Management Center T&E Science and Technology Cyber Test Technology Chief Scientist, shared one of the goals of his organization to rapidly progress high-risk/high-reward technology from technology readiness level (TRL) 3 to TRL 6. He stressed that we need people that don't just check the boxes when it comes to cyber test. He said we need to figure out how to measure the efficacy of testing and closed his comments with, "he who uses his data best wins."

The last speaker of the day was **George Rumford**, the Acting Director for the Department of Defense Test Resource Management Center, which falls under the Under Secretary of Defense for Research and Engineering, who told us that we need more resources for testing, and that there is a push to put more money in the budget to support testing, especially cyber testing. He told us that he isn't a fan of the shift-left mantra, as it assumes our schedule wasn't moving fast enough already and it really hasn't changed anything in the last 20 years. He said we should "cyber early and cyber often." He asked, "Why bother building data lakes and data oceans if we're just going to drown?" The point was the data doesn't matter if we don't do something meaningful with it. He said he expects about $2.1 billion to be added to the budget for testing next year.

Day two of the conference started with **Maj Gen Evan Diertien**, Commander, Air Force Test Center, who also emphasized the need for everyone testing together. He discussed the challenges of sharing data between platforms. He said sharing data between a couple of platforms is fairly easy but gets very complex when more platforms are added. He went on to say that there were resiliency issues as they become loaded with more and more users. Further, he said to optimize the kill web, we need data from all the systems out there. Cyber is one of the big challenges here, as individual platform data is classified at different levels. It becomes hard to sort through, analyze, and understand.

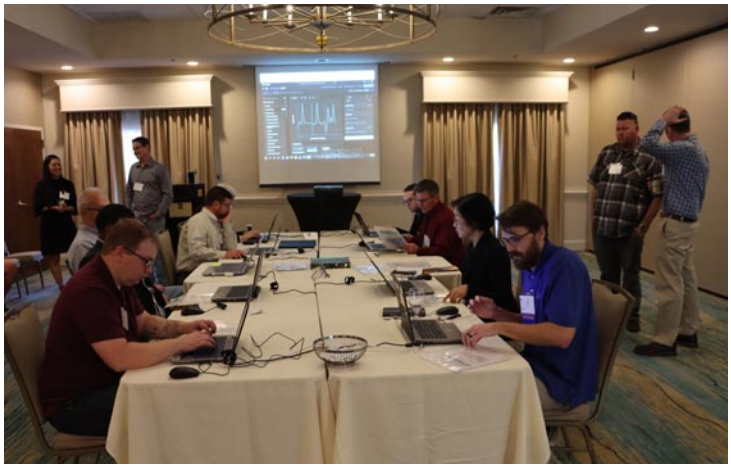
A.J. Pathmanathan, Director, National Cyber Range Complex reviewed the current and future state of Cyber Ranges, explaining that they are increasing the number of cyber ranges under the National Cyber Range Complex from five to nine. While there are several contracting efforts associated with running and expanding the cyber ranges, most users are repeat customers and there is no charge to use the ranges.

Echoing the sentiment from previous speakers that we can't produce cyber professionals fast enough, **Joe Bradley**, the Director of the Cyber Resiliency Office for Weapons Systems (CROWS) and the Director of Engineering and Technical Management for the Air Force Life Cycle Management Center explained that over the next three Future Years Defense Programs there will be a need to hire over 850 cyber professionals. He said we need to do cybersecurity based on intelligence data based on what the threats are and acknowledged that new vulnerabilities are discovered in systems frequently. The good news is that the CROWS has funding for prototype solutions.

Col William "Dollar" Young, former Commander of the 350th Spectrum Warfare Wing at Eglin AFB, shared his thoughts on cybersecurity, acknowledging the need to address cyber often and early to avoid the expense associated with adding it later in the development stages. He told us that we need to iterate until we have "good enough" security given the mission, platform, and the system of interest and the results need to be formally captured in a security concept of operations – "a hole in your policy is a hole in your architecture."

Col. David Hoffman, Commander, 96th Cyberspace Test Group, Eglin AFB, presented the missions of each squadron in the Group with emphasis on the 48th Cyberspace Test Squadron's capabilities and the Avionics Cyber Range (ACR). The ACR is composed of three labs (Avionics Cyber Test Lab, Cyber Threat Development Lab, and Cyber Electronic Warfare Lab) and National Cyber Range Instantiation. The ACR building facility construction is scheduled to complete in December 2022 and the integration of labs will begin in 2023.

Hands-on-Lab Experiment Team from left to right: Brandon Hyneman, Bill Beach, John Hardy, Sean Conway, and Annie Respigio



Hands-on-Lab Experiment during the workshop

One of the highlights of the workshop session was the Red Team Blue Team event put on by the 412th Range Squadron from Edwards, AFB. This highly attended session allowed participants to execute a Red Team cyber attack and a Blue Team defense in real time. Attendees were able to experience both sides of the attack by moving from one role to another.

A spur of the moment Special Technical Exchange Roundtable took place as attendees waited to take their turn in the exercise. This last-minute addition to the program was an open discussion on strategies, challenges, and solutions on interesting topics that brought about a lot of starter conversations and sidebar engagements.

The Chapter would like to extend a special thank you to our generous sponsors who helped defer the costs associated with hosting such an awesome event.

The workshop raised over $2000.00 for the Emerald Coast Chapter's STEM scholarship fund. The Chapter is pleased to be able to enhance the future T&E workforce one student at a time. The Emerald Coast Chapter is already in the planning stages as they prepare to host the ITEA Annual Test & Evaluation Symposium from December 4-7, 2023 at the Hilton Sandestin Beach Golf Resort and Spa in Miramar Beach, Florida.

Association News

By ELK Management Group

As we close out a great year with ITEA, let's take a walk down memory lane together. **Bruce Einfalt**, Applied Research Laboratory, Penn State, served as the Chairman of the Board of Directors for the past year and one of his crucial initiatives was to establish a committee for investigating Grant opportunities for ITEA. Bruce will continue to lead this effort as his term ends while serving as the Chair of the Senior Advisory Committee for the Board. Bruce encourages the membership to reach out to him if they would like to get involved. This effort has the potential to change the trajectory of ITEA and its goal of enhancing the future workforce for test and evaluation. **Tim Morey**, KBR, Vice Chairman of the Board of Directors led two efforts, one involved a re-work of the bylaws of ITEA and the other was establishing a team of leaders to develop a strategic plan for 2023 and beyond. This team continues to meet to regularly and will be defining the implementation plan to set the wheels in motion for an even stronger ITEA of the future. Keep an eye out for a membership survey geared toward international growth!

Our Board of Directors consists of 11 elected members and up to 4 appointed members and each election cycle we must tip our hats to the volunteer leadership who serves this association with dedication and enthusiasm. Several of our Directors served more than 1 term (3 years) and we want to acknowledge their service

to this organization. The following individuals will be leaving the Board as the year comes to an end: **Mark Brown**, Ph.D., Scientific Research Corporation, **Pete Crump**, CTEP, Georgia Tech Research Institute, **Brian Moore**, M6 Consulting Group, LLC, and our current chairman, **Bruce Einfalt**. In the 3rd Quarter, ITEA held its membership election campaign, and we would like to welcome our newest board members for the 2023-2025 term: **Steve Seiden**, Acquired Data Solutions, **Malcolm Tutty**, Ph.D., Royal Australian Air Force, **Van Sullivan**, Trideum Corporation, and **Robin Poston**, Ph.D., CTEP, Southern Methodist University. We also welcome **Keith Joiner**, University of New South Wales, Canberra as an appointed member to the Board for a 1-year term.

ELK Management Group has the distinct pleasure to serve the leadership of ITEA. This also includes the board members who are currently serving in their roles as Vice Chair, Tim Morey (mentioned above), Secretary, **Mark Phillips**, CTEP, Raytheon Missiles and Defense (2023), Treasurer, **Erwin Sabile**, CTEP, Booz Allen Hamilton (2023), Directors, **Joe Bullington**, CTEP, Jacobs Technology (2024), Cathy **O'Carroll**, CTEP, QinetiQ UK (2024), **D. Steve Woffinden**, CTEP, General Dynamics Mission Systems (2023), and **M. John Rafferty**, CTEP, US Air Force (2024). In addition to its Directors, ITEA also has government advisors who, while not voting members, do participate in achieving the goals of this non-profit educational association. We currently have representation from the Department of Homeland Security, Science and Technology, Test and Evaluation Directorate and the United States Army White Sands Missile Range.

In 2022 we held a variety of workshops and our annual symposia with great success. We owe a debt of gratitude to the volunteers who stepped up to plan and execute these events that not only tackled hot topics and challenges in our community but laid the groundwork for continued discussions and solutions that will lead to a safer and stronger environment for our Warfighters. The program chairs, technical chairs, instructors, panel leads, and technical session leads worked together to bring in the best presentations and guest speakers to enlighten their registrants on a variety of topics. We are so thankful to the behind-the-scenes volunteers working to make sure the presentations were collected, seen, or heard without a hitch; registering our attendees on-site as a seamless operation; taking photos; orchestrating STEAM competitions and Hand-on-Lab Exercises. Their time and energy was very much appreciated. At every event you will see banners with the company names of our generous sponsors who understand the importance of marketing, branding, and supporting the test community. You will also have the opportunity to meet with those companies exhibiting their capabilities and products — these companies see the value in getting in-front of the right audience. The year was filled with a lot of successes. We hope you'll join us next year.

Did you realize that in 2022 ITEA local Chapters provided scholarships that totaled nearly $15,000 to deserving students? These chapters are supporting the future T&E workforce one student at a time; congratulations! And thank you to those who sponsor, exhibit, and attend our workshops, conferences, and symposia – the Chapters earn scholarship dollars for hosting an event.

Our Professional Development Committee has been working hard in 2022 hosting educational courses through SAM.Gov and promoting their two-tiered certification program. Contact us. We'll give you all the details. This past year we hosted our monthly Lunch & Learn webinars — with people logging in from all over the world. You will want to check out what's on tap in 2023! We kick off January with a 45-minute session on the newly announced BAA for technologies required to test and evaluate future warfighting capabilities!

As we look out into 2023, we see so many more opportunities to connect with our membership and we look forward to getting to know you and to learn how we can make your Association work for you! Keep an eye out for our newsletter each quarter with even more news.

Have a wonderful holiday season and a successful New Year!

Certified Test and Evaluation Professionals

The following individuals have been awarded the **Certified Test and Evaluation Professionals** (CTEP) designation, which recognizes those individuals who demonstrate the following: They meet the minimum level of competency in the requisite Knowledge, Skills, and Abilities (KSA) that have been identified by T&E subject-matter experts (SMEs); their commitment to maintain currency in the field; and their dedication to advancing the profession.

Darryl Ahner, Ph.D., CTEP
Air Force Institute of Technology

Benjamin Andersen, CTEP
Modern Technology Solutions, Inc.

Sinisa Antoski, CTEP
General Dynamics Land Systems

W. Dave Bell, Ph.D., CTEP
The MITRE Corporation

Chris Bennett, CTEP
Bennett Consulting

Chad Boser, CTEP
Core Services Group

Melissa Keckley Bosshardt, CTEP
Naval Surface Warfare Center - PHD

Richard Boyer, CTEP
Scientific Research Corporation

Gary Brandstrom, CTEP
Raytheon Missile Systems Co.

Eric Breault, CTEP
Naval Surface Warfare Center - PHD

Brandon Brookins, CTEP
Naval Surface Warfare Center - PHD

C. David Brown, Ph.D., CTEP
Chesapeake Systems Engineering

Joe Bullington, CTEP
Jacobs

CAPT Caroline Goulart Campos, CTEP

Ismael Campos, CTEP
Brazilian Army Commission

Thomas Cash, CTEP
CGI Federal

Alice Cao, CTEP
Scientific Research Corporation

Peter H. Christensen, CTEP
Independent Consultant

Francis Xavier Costello, Jr., CTEP
AMERICAN SYSTEMS

Michael Cribbs, CTEP
NOVA Systems Australia

Peter G. Crump, CTEP
Georgia Tech Research Institute

Ramon Cuevas, CTEP
Northrup Grumman

1LT Rebeca Camurça Cunha, CTEP
Brazilian Army Commission

John Cunnick, CTEP
Naval Surface Warfare Center - PHD

Shane Deichman, CTEP
BAE Systems

Scott Dunphy, CTEP
IESE Solutions

Victoria Falkos, CTEP
Radiance Technologies

Michael Flynn, Ph.D., CTEP
Defense Acquisition University

Christine Fuentes, Ph.D., CTEP
The MITRE Corporation

Erik Glass, CTEP
Naval Surface Warfare Center - PHD

Phil Hallenbeck, CTEP
The MITRE Corporation

Johnathan Gettner, CTEP
Naval Surface Warfare Center - PHD

Mark Gillenson, Ph.D., CTEP
University of Memphis

John Jozef Hamann, CTEP
Booz Allen Hamilton

David Ronald Harrison, CTEP
Lockheed Martin Space

John Heavener, CTEP
Department of Homeland Security

Jessica Hodge, CTEP
Booz Allen Hamilton

Jerahmi Howard, CTEP
Booz Allen Hamilton

Darryl "Plug" Johnson, CTEP
JT4 LLC

Garfield S. Jones, CTEP
Department of Homeland Security

Mark Kiemele Ph.D., CTEP
Air Academy Associates

Joseph Lazarus, CTEP
STAT COE

Michael Lilienthal, Ph.D., CTEP
EWA Government Systems, Inc.

Eric Lowy, CTEP
WJ Hughes FAA Technical Center

Joseph Maki, CTEP
Naval Surface Warfare Center - PHD

Barrett McCann, CTEP
Booz Allen Hamilton

Charles McKee, CTEP
Taverene Analytics LLC

Martin "Marty" J. Mears, CTEP
Odyssey Systems Consulting Group

Henry C Merhoff, CTEP
Louis P. Solomon Consulting Group

Jason Morris, CTEP
Booz Allen Hamilton

Jeffrey Mueller, CTEP
Booz Allen Hamilton

Terry Murphy, CTEP
Department of Homeland Security

Renee Oats, CTEP
Naval Surface Warfare Center - PHD

Cathy O'Carroll, CTEP
Qinetiq

Martha O'Connor, CTEP
Department of Homeland Security

Steve Peduto, CTEP
Modern Technology Solutions, Inc. (MTSI)

Mark Phillips, CTEP
Raytheon

Robin Poston, Ph.D. CTEP
University of Memphis

Rafael Queiroz, CTEP
Brazil Army

Francisco Ramos, Jr., CTEP
Naval Surface Warfare Center - PHD

Robert Randolph, CTEP
Department of Defense

Kevin Lewis Renew, CTEP

Matthew Reynolds, CTEP
Independent Consultant

Anibal Torres-Rivera, CTEP
Naval Surface Warfare Center - PHD

Erwin Sabile, CTEP
Booz Allen Hamilton

Thomas Sachse, CTEP
US Navy, PEO SUB

Christopher James Sacra, CTEP
COMOPTEVFOR

Paul Shaw Ph.D., CTEP
Defense Acquisition University

Emmanuel (Mano) Skevofilax, CTEP
FAA William J. Hughes Technical Center

Charlie St. Clair, CTEP
Naval Surface Warfare Center - PHD

Chad Stevens, CTEP
KBR

Lowanda Studevent, PhD., CTEP
MITRE Corporation

Keith Sumner, CTEP
Booz Allen Hamilton

Stacie Taylor, CTEP
AFIT/LS School of Systems and Logistics

Miles Thompson, CTEP
The MITRE Corporation

Lesly Urbina, CTEP
Naval Surface Warfare Center - PHD

Jarrod Villanueva, CTEP
Naval Surface Warfare Center - PHD

Derick Wingler, CTEP
Booz Allen Hamilton

Steve Woffinden, CTEP
General Dynamics, Mission Systems

Johnathan Woodcock, CTEP
Naval Surface Warfare Center - PHD

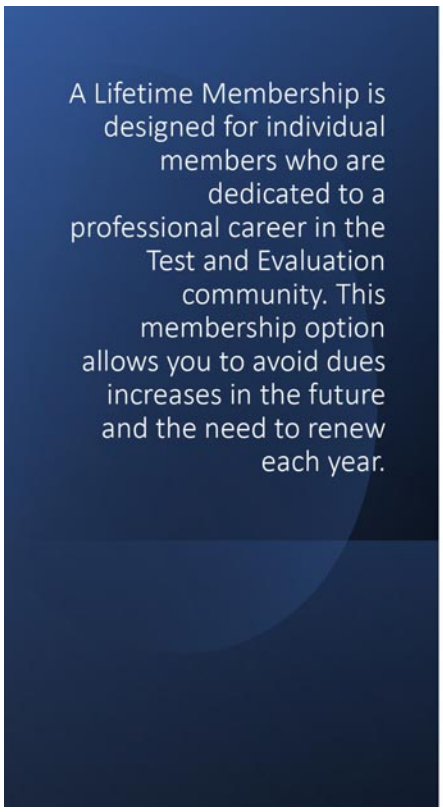
ITEA Corporate Members

Acquired Data Solutions, Inc.
Ad hoc Research
Advanced Test Equipment Rentals
Air Academy Associates
Amentum
AMERICAN SYSTEMS
Ansys, Inc.
Apogee Labs, Inc.
Applied Intuition Inc.
Applied Research Laboratory/PSU
Arcata Associates, Inc
Astro Haven Enterprises
Avionics Interface Technologies
Avionics Test & Analysis Corporation
BlueHalo, LLC
Booz Allen Hamilton
BrainGu
CALCULEX, Inc.
Calspan
Calypso AI Corp
Capability Analysis & Measurement Org., LLC
Collins Aerospace
Command Post Technologies, Inc.
Core Services Group
Creative Digital Systems Integration
Curtiss-Wright
Dell EMC Corporation
DEWESoft, LLC
Digiflight
DIS | Acroamatics
DIS | GDP Space Systems
DIS | Wideband
Dynetics, Inc.
ERC
EWA Government Systems, Inc.
Flight Test & Research Institute (Instituto de Pesquisas e Ensaios em Voo)
Frontier Technology Inc.
GaN Corporation
Garud Technology Services, Inc.
General Dynamics Mission Systems
Georgia Tech Research Institute
HII Mission Technologies Division
Jacobs Engineering
Joint Research and Development, Incorporated
JT4 LLC
KBR
Kistler Instruments
Lumistar, Inc.
ManTech
Maritime Warfare Centre (MWC)
NetAcquire Corporation
Nomad GCS
NovaSystems
ORCA Technologies
Parraid
Parsons
PeopleTec
Peraton Labs
Photo-Sonics, Inc.
QinetiQ LTD
Quintron Systems Inc.
Radiance Technologies
Raven Defense
Rotating Precision Mechanisms, Inc.
RoundTable Defense, LLC
SA-TECH, Inc.
Safran Data Systems, Inc.
Scientific Research Corporation
SemQuest Inc.
Summation Research, Inc.
SURVICE Engineering Company
SYMVIONICS Inc.
System Testing Excellence Program
Systems Engineering & Management Company
Technical Systems Integrators, Inc.
Tactical Data Link Bureau
Telspan Data
The MIL Corporation
TRAX International
Trenton Systems
TRIDEUM Corporation
WJ Hughes FAA Technical Center

As of December 2022

ITEA Lifetime Members

Charles E. Adolph
Rebecca Badgley
Damir Banjanovic
William B. Baker
Karen Barker
Suzanne M. Beers, Ph.D.
Eileen A. Bjorkman, Ph.D.
John V. Bolino
Katherine E. Bower
Nathan Grady Byrd
Duston Cline
Michael Cribbs
Dick H. Dickson
John W. Diem
Adam Emond
Anthony Fahy
Duane A. Goehring
Ed Greer
Dwayne T. Hill
Steven J. Hutchison, Ph.D.
Cathy Jaggard
Darwin L. Johnson
Dianne H. Jordan, Ph.D.
Mark London, Ph.D.
William M. Loughlin
Florestan Lucas
Hans Mair
Fred K. McCoy
Robert F. McKelvey
Christian Phillipp Merz
Douglas D. Messer
Hans Miller
James A. Neumeister
Peter G. Nikoloff
Parth Patel
Alan Plishker
Paul A. Polski
John Michael Rafferty
Matthew T. Reynolds
Chetan Sawyer
Jim Sebolka
William P. Singletary
Michael A. Spain
Richard M. Stuckey
Thomas Tarn
Miles Thompson
Scott Thompson
Robert Vargo
Dave Wellons
Chester Williams
Beth Wilson



CALL FOR PAPERS & TUTORIALS

40th International Test and Evaluation Symposium

NEXT-GENERATION TEST DOMINANCE

December 5-7, 2023

Hilton Sandestin Beach Golf Resort & Spa
Miramar Beach, FL

New technologies are coming fast to our workplaces, changing the way we accomplish tasks and prepare to fight future wars. Are these systems fully tested when they're being integrated into our practices? Are we leveraging the rapidly expanding toolkit to increase the level of complexity in our test and evaluation processes to keep up? Or are we letting the new technology pass us by in favor of what we already know how to execute? Are we learning from commercial partner, military, and academic channels or stove-piped in our immediate area?

Bring us your ideas, software tools, and hardware solutions to testing in contested EMS environments, modernizing the speed and depth of T&E, and making us ready for a near-peer adversary.

TECHNICAL SESSIONS

Topics for Consideration:

- Artificial Intelligence
- Autonomous Systems Testing
- Agile Testing
- 5 G
- Machine Learning
- Cyber Testing
- Digital Engineering
- Weapons
- Apps Deployment
- Electromagnetic Warfare
- Cognitive EW
- Software Programming
- Weapon System Survivability
- Sonar
- Signature Evaluation

Sessions will be unclassified and open to the general T&E community, however, restricted sessions (up to Controlled Unclassified Information (CUI)) may be made available. Abstracts will be reviewed for a presentation during a conference session or as a poster paper. Abstracts accepted for presentation, along with the slides, will be available online for attendees after the Symposium concludes.

Submit an abstract for a technical session or tutorial by May 1, 2023

TUTORIALS

Tutorials that address the theme of the Symposium are being requested. Pending final program decisions, the tutorials will be scheduled in two 4-hour blocks on Tuesday, December 5th. Tutorials must be strictly non-marketing in scope, unclassified, and publicly releasable. Please provide a synopsis of tutorial no later than May 1, 2023 to Symposium@itea.org.

For Abstract Submission Form or more information, visit:
www.itea.org/event/40th-te-symposium

Destin is close to:
Eglin AFB, Hurlburt Field, Duke Field
Army Ranger Camp in FWB, 7th Special Forces Grp
Coast Guard Station Destin
NAS Pensacola, Whiting Field, Panama City NSA
20th Space Surveillance Squadron

Hosted by the ITEA Emerald Coast Chapter

www.itea.org