# 38th International Test & Evaluation Symposium

*"Evolution of T&E in an Age of Rapid Technological Change"*

Hosted by ITEA Valley of the Sun Chapter

*9/2/21 Rev L*

*** Workshop will be held in Mountain Time**

*Tutorial Descriptions PDF*

## 13-Sept First Day – Tutorials

## 8:00 a.m. – 12:00 p.m. Morning Tutorials

**Building a Culture of Analytic Excellence**
*Mark Kiemele, PhD, Air Academy Associates*

**Ethically-aligned Experimentation and T&E: A human-centric view of complex military and safety critical intelligent and autonomous systems**
*Keith Joiner, PhD, & Malcolm G. Tutty, PhD, UNSW ADFA,ITEA Southern Cross Chapter - Australia*

**Incorporating T&E into Acquisition Contracts**
*Terry Murphy, Department of Homeland Security, T&E*

**Introduction to Cybersecurity Test and Evaluation**
*Pete Christensen, Consultant*

**Mind's Eye to Metal – T&E in the Acquisition and Development Team**
*David Brown, PhD & David Bell, PhD, MITRE Corporation*

## 1:00 p.m. – 5:00 p.m. Afternoon Tutorials

**Air Force's New MBCRA (Mission Based Cyber Risk Assessment) and Integrated Engineering Approach**
*Kevin McGowan, COLSA*

**Laser System Propagation T&E Challenges**
*Douglas H. Nelson, PhD, Teknicare, Inc. & Mark Stevens, Naval Postgraduate School*

**Machine Learning Here, Machine Learning There… Machine Learning in Test & Evaluation? An ML Tutorial for Everyone**
*Mark Tschopp, PhD, U.S. Army Research Laboratory*

**Real World Telemetry over IP**
*Gary Thom, Delta Information Systems*

**Successful Distributed T&E with TENA, JMETC, and BDA**
*Gene Hudgins, Test Resource Management Center Joint Mission Environment Test Capability (TRMC/JMETC)*

**T&E in Support of Agile: Test and Evaluation for Information Technology Acquisition**
*Robin Poston, PhD – System Testing Excellence Program, University of Memphis, and Mark Gillenson, PhD, University of Memphis*

# 38th International Test & Evaluation Symposium

*"Evolution of T&E in an Age of Rapid Technological Change"*

Hosted by ITEA Valley of the Sun Chapter

*9/2/21 Rev L*

## 14-Sept Second Day - Plenary Sessions, Technical Sessions, & Exhibits

8:30 a.m. Welcome:
Mr. Pete Crump, ITEA President
STAR Spangled Banner
Steve Woffinden , Symposium Chair

8:45 a.m. Opening Speaker: Raymond O'Toole, PhD, Director (acting) and Principal Deputy to the Director of Operational Test & Evaluation, Office of the Secretary of Defense

9:30 a.m. Keynote Speaker: James Wells, Director, Test and Evaluation, Department of Homeland Security, Science & Technology Directorate

---

**10:00 a.m. 30- minute Break. Please take this time to visit the virtual exhibit hall**

---

10:30 a.m. **Test and Resource Perspective on Emerging Technologies**

*Moderated by James Wells, Director, Test and Evaluation, Department of Homeland Security, Science & Technology Directorate*

Panelists:

- Kyle Fox, Executive Director, Systems Integration & Test, Raytheon Missiles & Defense
- Raymond O'Toole, PhD, Director (acting) and Principal Deputy to the Director of Operational Test & Evaluation, Office of the Secretary of Defense
- George Rumford, Director (acting) and Principal Deputy, Test Resource Management
- Malcolm G. Tutty, PhD, Australian Defence Force Academy, ITEA Southern Cross Chapter - Australia

**12:00 p.m. Lunch Break – Please take this time to visit the virtual exhibit hall**

12:30 p.m. Featured Speaker: Keoki Jackson, PhD, Senior Vice President, General Manager, MITRE National Security

1:15 p.m. Featured Speaker: Kyle Fox, Executive Director, Systems Integration & Test, Raytheon Missiles & Defense

2:00 p.m. Featured Speaker: Shimon Fhima, Director Strategic Programmes, UK Ministry of Defence

2:45 p.m. Preview of day two – Steve Woffinden, Symposium Chair

***Please join us in the virtual exhibit hall from 2:30-5:00pm.***
***Attendees will be able to video chat with exhibitors during this time.***

# 38th International Test & Evaluation Symposium

*"Evolution of T&E in an Age of Rapid Technological Change"*

Hosted by ITEA Valley of the Sun Chapter

*9/2/21 Rev L*

## 15-Sept Third Day - Plenary Sessions, Technical Sessions, & Exhibits

8:00 a.m. Welcome and overview of the day's events – Steve Woffinden, Symposium Chair

8:05 a.m. Featured Speaker: Matthew C. Wroten, Lt Col, USSF, Deputy, Test Enterprise Division

8:45 a.m. Featured Speaker: Devin Cate, Director of Test & Evaluation, USAF

9:30 a.m. Featured Speaker: Carroll "Rick" Quade, Director, Navy Innovation, Technology Requirements and Test and Evaluation, Department of the Navy

10:10 a.m. Featured Speaker: James Cook, Army T&E Executive

---

**10:45 a.m. 15- minute Break. Please take this time to visit the virtual exhibit hall**

---

11:00 a.m. **T&E for AI/ML Enabled Systems**

*Moderated by Jane Pinelis, PhD, Chief of the Test, Evaluation and Assessment Branch, Department of Defense Joint Artificial Intelligence Center*

Panelists:

- Alec Banks, Defence Science and Technology Labs
- Chad Bieber, Johns Hopkins University Applied Research Lab, Department of Defense Joint Artificial Intelligence Center
- Laura Freeman, PhD, Director, Intelligent System Lab, Virginia Tech Hume Center
- John Haman, Institute for Defense Analyses
- Galen Mullins, PhD, Johns Hopkins University Applied Research Lab

12:30 p.m. Featured Speaker: Neal MacKertich, PhD, Principal Engineering Fellow, Raytheon Missiles & Defense (retired)

**1:00 p.m. Lunch Break – Please take this time to visit the virtual exhibit hall**

1:30 p.m. Technical Tracks

# 38th International Test & Evaluation Symposium

## *"Evolution of T&E in an Age of Rapid Technological Change"*

Hosted by ITEA Valley of the Sun Chapter

*9/2/21 Rev L*

| Chair | Time | Title | Presenter(s) |
|---|---|---|---|
| **Track 1: Visual Analytics** | | | |
| **Simon Su, PhD, ARL DSRC** | 1:30 | *Using Natural Language Processing to Construct a Knowledge Graph of Test Incident Reports* | Sagar Indurkhya, PhD, Hector Vazquez, Aakash Indurkhya, Ciro Donalek, Virtualitics, Inc. |
| | 2:00 | *SyncVis, a Framework for 2D and 3D Visual Analytics* | Simon Su, PhD, and Vincent Perry, ARL DSRC; Olena Tkachenko, GDIT; Michael An, Parsons |
| | 2:30 | *Integrating Machine Learning and Interactive Visualizations for Heterogeneous T&E Data Analysis* | Mariya Occorso and Vincent Perry, ARL DSRC |
| | 3:00 | *Battlespace Visualization Interaction Platform for Test Planning and Data Visualization* | Simon Su, PhD, and Vincent Perry, ARL DSRC; Jeffery Peterson, & Charles Amburn, DPG |
| **Track 2: Cyber-Physical Systems Testing** | | | |
| **Bill "Data" Bryant, PhD, MTSI** | 1:30 | *Assurance Cases as a Grand Unified Theory of Weapon Systems and Platform Cyber Risk* | Bill "Data" Bryant, PhD, MTSI |
| | 2:00 | *Cyber Test and Evaluation for Supply Chain Risk Management* | Sarah Standard, OUSD R&E, Developmental Test, Evaluation, and Assessments |
| | 2:30 | *Cyber Testing Networks Over Their Entire Life Cycle in a Digital Twin Environment* | Charles D. Burdick & Deepinder Sidhu, PhD, Cyberspace Analytics Corp. |
| **Track 3: Future of T&E and Cybersecurity – Current Gaps, Innovations, and Requirements** | | | |
| **Erwin Sabile, Booz Allen Hamilton** | 1:30 | *Cyber Survivability Test and Evaluation* | Ryan Mayer, COMOPTEVFOR, Deputy Director, Cybersecurity OT&E |
| | 2:00 | *TBD* | Scott Aughenbaugh, Acting Chief of Staff, National Security Innovation Network (NSIN) |
| | 2:30 | *TBD* | CAPT William Selk, Commanding Officer |
| | 3:00 | *Joint Artificial Intelligence (AI) Red Teaming* | Galen Mullins, PhD, Johns Hopkins University Applied Research Lab (JHUAPL) |

# 38th International Test & Evaluation Symposium

***"Evolution of T&E in an Age of Rapid Technological Change"***

Hosted by ITEA Valley of the Sun Chapter

*9/2/21 Rev L*

| | | | |
|---|---|---|---|
| **Track 4: Data Analytics & Automated Test Reporting** | | | |
| **Larisa Funk, Booz Allen Hamilton** | 1:30 | *Applying Scientific Test and Analysis Techniques in the Agile/DevSecOps T&E Environment* | William F. Rowell, PhD, retired and Kimberly L. Wolfe, Scientific Test and Analysis Techniques Center of Excellence |
| | 2:00 | *Seminal Analysis Methodology Based on Uncorrelated Data Collection in Either Benign or Contested Environments* | Nathan A. Ruprecht & Elisa N. Carrillo, 746 Test Squadron |
| | 2:30 | *Cost-effective Method to Characterize Antenna Performance* | James D. Silk & James M. Ralston, Institute for Defense Analyses |
| | 3:00 | *Data Science and Predictive Modeling* | Chelsea Owens & Larisa Funk, Booz Allen Hamilton |
| **Track 5: Agile Methods and DOE** | | | |
| **Chris Olinde, GTRI** | 1:30 | *Use of DOE tools in a Development Laboratory* | Jasveer Ramroo Beni, Master Builders Solutions Construction Chemicals, LLC |
| | 2:00 | *The Challenges of Agile Software Delivery for Military Embedded Software Solutions* | Chris Olinde, Andrew Mishoe, Stephens Summerlin & Ken Bailey, Georgia Tech Research Institute |
| | 2:30 | *Looking at Legacy I&V through an Agile Lens* | Scott A. Goodman & Doug Moshier, Raytheon Missile & Defense |
| **Track 6: International T&E** | | | |
| **Ade Britton, PhD, QinetiQ** | 1:30 | *Electro-Magnetic Spectrum Operations (EMSO) Challenges for Future Test and Evaluation Systems* | Loannis Vagias, Cranfield University; Raffaele Fiengo, National Instruments (NI); Anthony McKelvey, Defence Science and Technology Laboratory (DSTL) |
| | 2:00 | *A Software Connected Approach for Facing the Challenges Associated to New EMSO and Digital Transformation in T&E* | Raffaele Fiengo, National Instruments (NI) & Massimo Sciotti, Elettronica GmbH (ELT) |
| | 2:30 | *Globally Deployable Maritime Tracking Range Capability for Maritime Training, Mission Rehearsal, Test and Evaluation* | Steve Powell, & Rogiero Ramos Qinetiq & Darryl Newborough, Sonardyne |

***Please join us in the virtual exhibit hall from 3:30-5:00pm. Attendees will be able to video chat with exhibitors during this time.***

# 38th International Test & Evaluation Symposium

*"Evolution of T&E in an Age of Rapid Technological Change"*

Hosted by ITEA Valley of the Sun Chapter

*9/2/21 Rev L*

## 16-Sept Fourth Day - Plenary Sessions, Technical Sessions, & Exhibits

8:20 a.m. Welcome and overview of the day's events – Steve Woffinden, Symposium Chair

8:30 a.m. **T&E Workforce**

*Moderated by Terry Murphy, Deputy Director, Policy and Workforce Development, Department of Homeland Security, Office of Test and Evaluation*

Panelists:

- Darryl Ahner, PhD, Director, Scientific Test and Analysis Techniques Center of Excellence
- Pete Christensen, Consultant
- Robin Poston, PhD, Dean of the Graduate School, University of Memphis
- Laura Freeman, PhD, Director, Intelligent System Lab, Virginia Tech Hume Center
- Joseph Stasiwski, Enterprise Test Services Lead, Terminal Automation Team, Federal Aviation Administration's (FAA's)
- Kenneth Stefanek, Learning Director for Test and Evaluation, Defense Acquisition University

---

**10:00 a.m. 30- minute Break. Please take this time to visit the virtual exhibit hall**

---

10:30 a.m. **The Past is Prologue - A Panel of Past Presidents**

*Moderated by Matt Reynolds, Consultant*

Panelists:

- Patricia Sanders, PhD, Independent Aerospace Consultant
- John Foulkes
- Rusty Roberts, Georgia Tech Research Institute (GTRI)
- Stephanie Clewer, Systems Application & Technologies, Inc. (SA-TECH)
- Mark Brown, PhD, Vice President, Scientific Research Corporation
- Chas McKee, President and CEO, Taverene Analytics LLC
- Gene Hudgins, Director of Test and Training Environments, KBR
- Bill Keegan, Executive Vice President, Equator Corporation
- Pete Crump, Georgia Tech Research Institute (GTRI)

**12:00 p.m. Lunch Break – Please take this time to visit the virtual exhibit hall**

**12:30 p.m. Test & Evaluation Professional Awards Presentation**

1:00 p.m. Technical Tracks

| Chair | Time | Title | Presenter(s) |
|---|---|---|---|

# 38th International Test & Evaluation Symposium

## *"Evolution of T&E in an Age of Rapid Technological Change"*

Hosted by ITEA Valley of the Sun Chapter

*9/2/21 Rev L*

| | | | |
|---|---|---|---|
| **Track 7: STEM for T&E** | | | |
| **Virginia To, Army Research Laboratory /Parsons** | 1:00 | *Forge the Future with Computing* | Virginia To, Army Research Laboratory /Parsons |
| | 1:30 | *Topic Modeling and Visualization for Test Incident Report Data* | Sage Leone, University of Maryland; Jamie Infantolino, US Army Research Laboratory; Adam Childs & Cleon Anderson, US Army Research Laboratory/Parsons |
| | 2:00 | *Visual Analytic Application Toolset for Large-Scale Data* | Aiden Kenny, Columbia University; Mariya Occorso, Arcadia University; Vincent Perry, US Army Research Laboratory |
| | 2:30 | *Immersive Environments for Visual Analytics* | James Hughes, UMD College Park; Selena Hamilton, Penn State University; Virginia To, US Army Research Laboratory /Parsons |
| **Track 8: Digital Transformation** | | | |
| **Suzanne Beers, PhD, MITRE** | 1:00 | *T&E: Required Leadership in The Age of Transformation* | Paul McNamara, The Sente Group, Inc. |
| | 1:30 | *Accelerating Capability Delivery Through Mission-Focused MBT&E* | Joe Murphy & Bill Beuth, Analytical Graphics Inc / An Ansys Company |
| | 2:00 | *Mission Engineering, Capability Evaluation & Digital Engineering Informing DoD Technology, Prototyping and Acquisition Decisions* | Suzanne Beers, PhD, OUSD(R&E)/DTE&A (The MITRE Corporation) |
| **Track 9: Autonomous Systems** | | | |
| **Leonard Truett, PhD STAT COE** | 1:00 | *The Use and Benefits of Modeling and Simulation with Autonomous Testing* | John Whitt & Paul Bounker, Aberdeen Test Center/Ground Vehicle Systems Center |
| | 1:30 | *Evaluating Autonomous System Information Dominance* | Cierra Hall, Judith Stoer & David Scheidt, Weather Gage Technologies |
| | 2:00 | *Path Planning for Multi-Agent Semi-Autonomous Rover in an Integrated System Test Facility Without the Aid of GPS* | Lauren Blanks, Allyson Burba, Isaac Isukapati & Brian Neff, U.S. Air Force Academy and National Robotics Engineering Center |
| | 2:30 | *Advancements in Test and Evaluation for Autonomous Systems (ATEAS)* | Charlie Middleton & Lenny Truett, PhD, Scientific Test and Analysis Techniques Center of Excellence (STAT COE) |
| **Track 10: International T&E 2** | | | |

# 38th International Test & Evaluation Symposium

***"Evolution of T&E in an Age of Rapid Technological Change"***

Hosted by ITEA Valley of the Sun Chapter

*9/2/21 Rev L*

| | | | |
|---|---|---|---|
| **Ade Britton, PhD, Qinetiq** | 1:00 | *Modernizing T&E is Much Easier to Say…* | Richard Cawthorne, Defence Science and Technology Laboratory (DSTL) |
| | 1:30 | *The Queensland Flight Test Range (QFTR)* | Rob Dunn & Jim Parkes, QinetiQ Australia Pty Ltd |
| | 2:00 | *System of Systems Synthetic Evaluation* | Mark Threadgold and Ben Egan, UK Ministry of Defence, Defence Science and Technology Laboratory (DSTL) |
| **Track 11: Workforce Certification** | | | |
| **Pete Christensen, Consultant** | 1:00 | *Elevating the T&E Profession Through CTEP* | Pete Christensen, Consultant |
| | 1:30 | *DOD T&E Certification* | Paul Shaw, Professor, Cybersecurity, Defense Acquisition University |
| | 2:00 | *Micro-Credentialing for Professional Development* | Taz Daughtrey, President, Association for Testing & Software Quality Assurance |
| | 2:30 | *Value of Testing Certifications for Agile & DevSecOps* | Andrew Pollner, American Software Testing Qualifications Board |
| **Track 12: Test Technology** | | | |
| **Ben Kupferschmidt, Curtiss-Wright** | 1:00 | *Delivering Output Based T&E Capability* | Andrew Cunningham, QinetiQ |
| | 1:30 | *Making Flight Test Instrumentation Setup Easier* | Ben Kupferschmidt, Curtiss-Wright |
| | 2:00 | *Testing Mixed-Use Networks – Insights Into a Complex Test Process* | Dan Pleasant, Jarrod Tsukada & Kelly Caulk, Keysight Technologies, Inc. |
| | 2:30 | *Testing for the Future Fight: Necessary Range Capabilities to Ensure Operational Superiority of Defense Systems* | Heidi Perry, MIT Lincoln Laboratories; Hans Miller, MITRE Corporation; Derrick Hinton, SRC |

***Please join us in the virtual exhibit hall from 3:00-5:00pm. Attendees will be able to video chat with exhibitors during this time.***

# 38th International Test & Evaluation Symposium

*"Evolution of T&E in an Age of Rapid Technological Change"*

Hosted by ITEA Valley of the Sun Chapter

*9/2/21 Rev L*

## 17-Sept Fifth Day - Plenary Sessions, Technical Sessions, & Exhibits

7:30 a.m. Welcome and overview of the day's events – Steve Woffinden, Symposium Chair

7:35 a.m. State of ITEA – Pete Crump, President of ITEA

7:45 a.m. Featured Speaker: Michael Schmidt, PhD, Master Builders Solutions

8:30 a.m. **T&E Integrated Into Agile and DevSecOps**

*Moderated by Robin Poston, PhD, Director of the System Testing Excellence Program, FedEx Institute of Technology, The University of Memphis*

Panelists:

- Mike Knoll, Vice President, IT, FedEx
- Tom Meservy, PhD, Brigham Young University
- Antonia Pulley, U.S. Citizenship and Immigration Services
- John Stasulli, Parsons Corp

---

**10:00 a.m. 30- minute Break. Please take this time to visit the virtual exhibit hall**

---

10:30 a.m. **Digital Engineering**

*Moderated by Edward Kraft, PhD*

Panelists:

- Darryl Ahner, PhD, Director, Scientific Test and Analysis Techniques Center of Excellence
- COL (R) Rick Bailer, Consultant
- Tom Downs, Manager, Engineering, Weapon Systems, GD-OTS Armaments and Energetics Business Unit
- Kevin Fahey, President and CEO, Cypress International
- Camille Robbins, Director, Virtual Test and Advanced Electronics Directorate, ATC

---

**12:00 p.m. 10- minute Break. Please take this time to visit the virtual exhibit hall**

---

12:10 p.m. Closing Speaker: James Wells, Director, Test and Evaluation, Department of Homeland Security, Science & Technology Directorate

12:45 p.m. Final Remarks – Steve Woffinden, Symposium Chair

## Technical Track Abstracts

### *A Software Connected Approach for Facing the Challenges Associated to New EMSO and Digital Transformation in T&E*


Raffaele Fiengo, National Instruments (NI) & Massimo Sciotti, Elettronica GmbH (ELT)



Modern Electro-Magnetic Spectrum Operations (EMSO) urge for rapid evolution of T&E principles and architectures in order to cope with the engineering and the operational T&E challenges throughout the System Life-Cycle (SLC) of modern Radar/EW and CNS (Communication, Navigation and Surveillance) Systems.
State-of-the-art EMSO operations rely more and more on (i) SW-based capabilities, which enable agility, adaptivity, cognition and "artificial intelligence", (ii) federated and multi-domain Systems of Systems (SoS), (iii) conformal sensor-platform pairs, and (iv) cyberspace.
The EMSO scenario spans nowadays from low-observable, low probability of interception, smart, covert and camouflaged "threats" to high-power and high-density, interfering, agile emissions. Inter-dependent, time-varying, reactive and adaptive emissions are to be found in any sub domain of EMSO, which accounts for radar, communication, and navigation frequency bands, as well as the infrared, ultraviolet and visible spectrum.
In addition, the on-going digital transformation of T&E capabilities requires an integrated and cross-functional approach for improving the utilization and efficiency of the T&E resources throughout the SLC, reducing the total cost of ownership and accelerating the time to market by introducing the Software and Hardware-in-the-Loop (SiL/HiL) test in all stages of the SLC and in any test environment (office, laboratory, test rigs, anechoic chambers, integration facilities, open air ranges).

Modelling & Simulation (M&S) is thus challenged by EMSO, since the test scenario can only result from accurate representation of the actors' behavior in the gaming area. Modelling a multi-functional phased array radar is indeed a challenge in terms of compromising complexity versus realism, portability versus fidelity.

Clearly, Measures of performance (MoP) and Measures of effectiveness (MoE) in T&E are scaling up in complexity to consider the response of "intelligent" systems to "intelligent" threats. Modelling and evaluating "cognitive" functions is today's challenge, if we consider the need for "validating" the model itself. Only a validate M&S output can support a high-confidence assessment.
In parallel, the Model-based System Engineering (MBSE) principles stem out as enabler for describing both the capabilities under test and the required T&E functions. These descriptors provide the frame for M&S developments, which generate the required SW layers of the T&E capability.
The above-listed needs and trends turn into an evolutionary path for the EMSO T&E capabilities. In this speech, we address a fundamental question: are state-of-the-art digital technologies fit for supporting this path? The resulting requirements to HW modularity, programmability, scalability, reliability, computational resources, connectivity, memory storage, etc. are here discussed and analysed in order to determine the maturity of market solutions for EMSO T&E.

---

## *Accelerating Capability Delivery Through Mission-Focused MBT&E*

Joe Murphy & Bill Beuth, Analytical Graphics Inc / An Ansys Company

The recent SAF/AQ guidance around e-Program designations puts digital engineering at the forefront of future acquisition and test. Fielding of critical capabilities can be accelerated by continuously assessing Model Based Systems Engineering (MBSE) and Model Based Test & Evaluation (MBT&E) activities relative to mission performance requirements. This presentation will explore how MBSE designs and MBT&E data can be evaluated against mission performance requirements in physics-based virtual operating environments. For context, an aircraft development use case will be analyzed in three phases: pre-test planning, real-time test execution, and post-test analysis.

---

## *Advancements in Test and Evaluation for Autonomous Systems (ATEAS)*

Charlie Middleton and Lenny Truett, PhD, Scientific Test and Analysis Techniques Center of Excellence (STAT COE)

Autonomous Defense Systems (ADS) continue to garner great interest in the Department of Defense and military Services, particularly to keep U.S. advantages in a field where near-peer adversaries are quick to capitalize on capability first, while putting off or ignoring public safety. The new capabilities promised by ADS will need to operate in dynamic environments which may change the way they operate and are likely to display emergent, unexpected behavior, with or without human teaming. Requirements and measures are not well developed for many of the desired test needs of ADS, including for example trust, safety, assurance, and human-machine teaming.
The continued application of Artificial Intelligence (AI) and Machine Learning (ML) will expand the challenges associated with properly testing and evaluating these complex adaptive systems to ensure appropriate developmental and operational mission assurance are fully studied and addressed. Although T&E is well established for traditional weapons system, ADS represent a new frontier in the field of T&E as test planning and analysis for these complex adaptive systems will require innovative methods.
Much of the work that has been documented to date has been on identifying the challenges associated with the T&E of ADS. The ATEAS effort is focused identifying programs that provide practical approaches to addressing the challenges through tools, methodologies, best practices and lessons learned. The presentation will discuss the findings gathered from active programs, the framework for a guidebook to capture this

information, and future efforts to support the continued development of the guidebook as a living reference document.

## *Applying Scientific Test and Analysis Techniques in the Agile/DevSecOps T&E Environment*

William F. Rowell, PhD, Retired

The intelligent application of scientific test and analysis techniques (STAT) to Department of Defense (DoD) weapon and business systems has proven to be a key component in enabling efficient, effective, and rigorous test & evaluation (T&E) in support of decision making at all levels. With the growing implementation of Agile and Development/Security/Operations (DevSecOps) software development methodologies in the acquisition of software intensive DoD systems, a similar interest in leveraging STAT in the Agile/DevSecOps T&E environment has arisen. This paper provides guidance to the DoD T&E workforce for applying STAT to T&E of software intensive systems in the Agile/DevSecOps environment including insights into the appropriate STAT techniques to apply.

## *Assurance Cases as a Grand Unified Theory of Weapon Systems and Platform Cyber Risk*

Bill "Data" Bryant, PhD, MTSI

One of the most pressing problems with weapon systems and aviation platform cybersecurity is the lack of an agreed upon methodology for measuring mission risk for a given system within its operational environment. This problem is especially difficult, given there is so little historical data to draw from, but the lack of an agreed upon method to assess or measure risk, greatly hampers our ability to know where to position our limited resources.

The main reason why a risk measurement system has yet to gain widespread acceptance among the continuing debates over qualitative versus quantitative approaches, is that there is no single approach or tool that is a best fit in all circumstances and environments. Therefore, what is needed is a family of connected tools across the spectrum of qualitative analysis and quantitative measurement that use similar formats and outputs enabling some comparison across the tools.

This Unified Risk Assessment and Measurement System or URAMSTM provides a diverse set of integrated qualitative and quantitative tools that provides true risk management for weapon systems and aviation platforms throughout the development lifecycle and across a range of contested cyberspace environments. URAMS starts with an engineering analysis utilizing Systems-Theoretic Process Analysis for Security (STPA-Sec). This tool was developed from leveraging the safety analysis work done at MIT and has since been used with great effectiveness across a range of military weapon systems and civilian aerospace systems. 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 lifecycle, which enables true "baking in" of security.

There are three assessment and measurement tools within URAMS that can be used to score the risk scenarios developed by the STPA-Sec analysis. The first and most simple is Risk Assessment (RA) which utilizes a series of scored qualitative factors to estimate the likelihood and consequence of both financial and mission-based risks. All URAMS tools are capable of measuring risk in terms of either financial, mission loss, or both, depending on the needs of the organization. Second, Risk Assessment with Uncertainty (RAU), has the same qualitative structure as RA, but includes an ability to estimate uncertainty using a three-point estimation process. For each sub-element, analysts enter an expected score of a worst or best-case scenario. Third, and

for quantitative analysis, URAMS offers Probabilistic Risk Measurement (PRM), which utilizes direct estimation of probabilities and consequence by appropriately trained experts. With its quantitative foundation, PRM enables a more in-depth exploration of risks with less uncertainty than typically provided by RAU.
Each of these risk analysis tools have a corresponding tool that enables the combination of risks into a single risk by using Monte Carlo simulations. This enables systems, or even portfolios of systems, to be compared to each other; and perhaps more importantly, allows overall risk to be compared to a previously determined organizational risk tolerance to answer the critical question of when the design is "secure enough".

## *Battlespace Visualization Interaction Platform for Test Planning and Data Visualization*

Simon Su, PhD, Vincent Perry, Jeffery Peterson, & Charles Amburn, ARL DSRC

The Test & Evaluation (T&E) planning, data analysis and review process is so complex that it often takes months to complete due to a lack of automation, data accessibility and universal platform. Battlespace Visualization Interaction (BVI) is a research and development (R&D) testbed for experimentation in battlespace visualization and decision-making with R&D focus on human factors-related research in the areas of information visualization, multi-modal interaction, and human performance assessment. With a slight modification, BVI can also be leverage to support T&E planning, data analysis and review process. BVI functionalities in Battlespace visualization and decision-making aid can also utilized in test planning and visualization to support T&E decision-making. Furthermore, using BVI 3D terrain visualization, test planners are able to collaboratively plan and execute their T&E process with a digital archive that they reference in the future. BVI geospatial terrain information and map images allowing users to virtually plan a test events. BVI can also be a visualization platform to visualize all the test data collected during the course of the T&E events. Test data analysts and evaluators will also be able to leverage BVI to observe data as it were collected at the testing facility in the virtual environment. BVI also acts as a data server distributing the data to client applications that provide data to the users via one or more of the supported visualization modalities allowing the user to visualize the test data on Microsoft HoloLens2, web based interfaces, android table devices, Head-Mounted Display devices, and the mixed reality devices.

## *Cost-effective Method to Characterize Antenna Performance*

James D. Silk & James M. Ralston, Institute for Defense Analyses

In the test and evaluation process it is frequently required to verify that the performance of an antenna, or even a single element of an antenna array, is adequate to support its function in a larger system. This verification has generally required measurements on a large, far-field, antenna test range or else within a near-field scanning system. Although these methods are rightly considered the "gold standard" for antenna measurements, employing either of them can entail considerable cost and schedule impact to the overall program. In this presentation we will show how essential antenna characteristics can be determined in a small-scale laboratory setting.

## *Cyber Test and Evaluation for Supply Chain Risk Management*

Sarah Standard, OUSD R&E, Developmental Test, Evaluation, and Assessments

This talk presents cyber T&E strategies and activities to address cyber threats to the supply chain of U.S. Department of Defense weapon systems. These strategies may be applied to software, hardware, and firmware, expanding on the software-focused supply chain guidance in the 12 May 2021 Executive Order, "Improving the Nation's Cybersecurity". The cyber T&E strategies and activities presented will strive to address the unique

attributes of cyber threats to the supply chain. The supply chain may be vast and complex; evolve over the lifecycle of the system to be acquired from design to production, fielding, and disposal; and be subjected to attacks (which may target even trusted suppliers) that are more complex and difficult to detect than traditional cyberspace attacks. In order to address the vast scope and continuous evolution of the supply chain, and the potential for unwarned threats, supply chain T&E should be embedded within a continuous and iterative planning process comprised of threat characterization, supply chain illumination, and attack surface characterization informed by criticality analysis and results from Mission-Based Cyber Risk Assessments. Supply chain T&E will require contract language and planning ahead for specialized tools, techniques, and test personnel. This talk explores the complicated nature of preparing for and planning and performing supply chain T&E to prevent, identify, mitigate, and deter cyberspace threats to the supply chain.

---

## *Cyber Testing Networks Over Their Entire Life Cycle in a Digital Twin Environment*

Charles D. Burdick & Deepinder Sidhu, PhD, ITA International, LLC

The Test Resources Management Center is already using network digital twins for testing wireless communications devices and their connection to weapon systems. But there are other Network Digital Twin (NDT) testing systems being brought into the arena with capabilities to automatically map very large networks, create software clones from the data obtained, rapidly visualize those clones, and then reverse engineer them to be network digital twins of specific physical networks, even if they contain more than a million devices. Even at this scale, these NDTs can still meet the DoD's digital twin requirement to, "mirror and predict activities and performance over the life of its corresponding physical twin." This latter capability is a critical requirement, and the NDT can emulate a network with precise timing and bit-level fidelity from its original design stage until it is retired.
It does this by:

- In the design stage emulating the actual design using network software and virtual hardware, then as actual hardware and software comes on-line by
- Automatically reviewing every network and commercial report on the network as they come in.
- Using its data analysis tools to compare the results of each network and cyber report with the existing clone to detect changes, and
- Immediately notifying the network administrators and cyber personnel based on the alert priorities they have set

Even following Full Operational Capability, these NDTs should continue to be used with their specific network twin because we have seen how fast networks evolve from the version that was originally tested with regular software patches, hardware changes, and human modifications in rules and configurations either purposely or inadvertently.
With the NDT every software, hardware, and configuration change proposed for a network can be emulated before it is implemented ensuring that all involved have a common understanding of the effects of the change. The NDT can address complex problems arising in high-speed physical network since the NDT can be run at slower speeds that allow human observation of every interaction and the opportunity to correct the problem and immediately test the solution on the emulation.
We are now entering a long period of development and testing for the Joint All-Domain Command and Control (JADC2) network of networks, and we need NDT capabilities to:

- Search for cyber vulnerabilities, single points of failure, unnecessary redundancies, incompatibilities, and opportunities to increase the robustness of its individual parts in its design stage, well before its

pieces are built

- Test proposed interfaces with its multiple networks as well as with networks of allies and others
- Confirm the system scales.

This NDT capability has been used for several years in the Information Assurance community and is being used by several large government and commercial organizations. Their Red teams can find no differences between the NDT packets and responses to change and those of their physical twins. We believe that widespread use of this relatively low-cost technology as a continuous monitoring and test system can significantly increase our cyber readiness for specific networks over their entire life cycle.

---

## *Delivering Output Based T&E Capability*

Andrew Cunningham, QinetiQ

Since 2003 QinetiQ have been delivering core UK Military T&E Capability to the UK Ministry of Defence (MOD), in partnership with the Trials, Evaluation Services and Target (TEST) Project Team, through the Long Term Partnering Agreement (LTPA). This has historically been through a traditional availability based contract. Over the past 2 years this has transitioned from an availability based contract to an output based contract, delivering significant savings and enabling significant investment in the T&E Capabilities to ensure they remain relevant to defence in the future. This has involved redefining the delivery scope from a facilities based taxonomy to a T&E event taxonomy.

In addition to developing the new taxonomy, significant work has been done to understand the elements of each of the events, understanding the processes, people, and information required to deliver each of the events.

Throughout the transition, working in partnership with the customer has been crucial to overcome the technical and cultural challenges the transition has presented, and ensure a smooth and successful achievement of the Full Operating Capability.

The Investment process has also changed, moving from metrics focused on delivery of a new facility or infrastructure, to metrics based on the delivery of the events, and the continued relevance to defence of the outputs.

---

## *DOD T&E Certification*

Paul Shaw, Professor, Cybersecurity, Defense Acquisition University

In the DoD Test and Evaluation community, T&E certification has been a requirement for career progression and for verification of achievement of core T&E competencies. The DoD T&E community has traditionally had a range of critical test positions such as: Test Range Operator; Operational Test Lead; Test Planner; T&E Lab Manager; and Chief Development Tester. Career advancement in these test positions has required progress in technical and professional "Core Readiness Competencies." Examples of some of the technical competencies include risk identification, test readiness, test cost estimating, test control management, capabilities assessment, and validation of test results. Examples of some of the professional competencies include critical thinking, communication skills, professional ethics, technical credibility, and leadership and management. The evolving DoD framework for T&E certification includes a level of certification for Foundational and Practitioner. Per the DoD's 2020 T&E Task Force recommendations, Foundational certification signifies "a basic understanding of the T&E concepts and is developing skills on a routine set of tasks through interactions with skilled workers and on-the-job experiences. At a minimum, has demonstrated the ability to support and assist in T&E activities while interacting with multiple organizations." Per the DoD's

2020 T&E Task Force recommendations, Practitioner certification certifies "a full understanding of the concepts and basic set of skills to perform T&E activities. Has gained knowledge and experience within the T&E community by performing routine tasks with limited supervision. At a minimum, has demonstrated the ability to manage and direct T&E activities while interacting with multiple organizations." In order to succeed in the T&E community with the scope of responsibility and an increasing threat of a great power competition from near-peer nation state military capabilities, professional certification through the development and progression of core readiness competencies is critical for DoD mission assurance and achievement of national security objectives.

## *Electro-Magnetic Spectrum Operations (EMSO) Challenges for Future Test and Evaluation Systems*

Loannis Vagias, Cranfield University; Raffaele Fiengo, National Instruments (NI); Anthony McKelvey, Defence Science and Technology Laboratory (DSTL)

Radio detection and communication using the electromagnetic spectrum (EMS) enables a number of modern capabilities for military applications. The military worldwide use EM radiation to enable communications, navigation, radar, non-intrusive inspection, and other applications. The military also heavily rely on EM radiation for intelligence, surveillance, and reconnaissance (ISR) applications and self-protection, such as missile early warning, signals intelligence (SIGINT), jamming and deception of radars and communications. Governmental agencies together with the military manage access to and use of the electromagnetic spectrum (EMS). As the markets continually experience significant growth in commercial wireless services, the bandwidth demand for spectrum has increased. Governments globally have enacted policies to make additional spectrum available for commercial use, in some cases reallocating spectrum from governmental agencies and the military to commercial use, considering the following issues:

- Interoperability
- Technological superiority
- Increasing commercial interest in bands traditionally allocated to the military
- Spectrum sharing
- The interagency process for spectrum allocation
- Anticipating future spectrum needs for all users

At the same time, as bandwidth demand increases, the military are developing radio frequency systems for radiolocation and communications that enable:

a. Wide bandwidth (WB) spread spectrum (SS) waveforms.
b. Multi-beam and multimode sensors.
c. Digital processing and high sampling rates.
d. Simultaneous time and frequency domain (TD-FD) digital processing.
e. Spatial FFT processing for space-time adaptive processing (STAP).
f. Simultaneous amplitude and phase comparison for AOA determination.
g. Fused active and passive RF modes of operation for target detection and geolocation.
h. Multistatic and fused modes of operation for detecting of low observable (LO) targets.
i. On-board and off-board active or passive electronic countermeasures.

The above modes of operation have to integrated onto a combat platform with fire control sensing capabilities; the most challenging platform being the fast jet. There, these modes of operation have to be integrated with flight control aids, navigation aids, intra and inter platform data handling aids, the defensive aids and the stores management system. Beyond inter platform integration, the EMSO system has to be integrated with a multitude of other sensors such in terms of:

a. Interoperability (spectrum management)
b. Distributed, cooperative and time sensitive targeting

Each subsystem and the system overall need to operate under the Observe-Orient-Decide-Act (OODA) loop, with or without an embedded human decision incorporating elements of artificial intelligence . Thus, Development, Test and Evaluation (DT&E) phase of a system or a system of systems is very challenging as it has to be concluded within a predetermined time frame to counter obsolescence and the adversary disruptive technologies.

To understand the challenges of modern T&E equipment in the Electronic Warfare domain, we need to understand the complexity of modern sensors in terms of hardware and software. We will address that discussion in this paper.

---

## *Elevating the T&E Profession Through CTEP*

Pete Christensen, Consultant

ITEA established the Certified Test and Evaluation Program in 2014. The purpose of the program was to provide a certification venue for experienced professionals who independently demonstrated mastery of the requisite skills essential to be called a T&E professional. The CTEP program is intended to recognize individuals who demonstrate the knowledge, skills, and abilities (KSAs) identified by T&E subject-matter experts (SMEs), demonstrated a commitment to maintain currency in the field and were dedicated to advance the T&E profession. The common standards, principles, procedures, processes, and terms for the T&E profession were selected by SMEs appointed to the ITEA Board of Examiners (BOE). ITEA is committed to supporting professional development and education to enhance the KSAs of T&E professionals.

In 2020, the ITEA BOD recognized a need to revise and update the CTEP Program. A new BOE was assembled and over the last year and the CTEP Program has been significantly enhanced. Two levels of certification will be offered. The Foundational Certification will be offered for the early T&E Professional. The Practitioner Certification will continue to be offered for more senior T&E Professional. This presentation will provide an overview of the new program including the revised CTEP Domains, Body of Knowledge and discuss certification options.

---

## *Evaluating Autonomous System Information Dominance*

Cierra Hall, Judith Stoer & David Scheidt, Weather Gage Technologies

This abstract delves into the area of Autonomy as a form of command and control in which machines are allowed to make decisions. Autonomous testing and evaluation are correlated to evaluating the quality and timeliness of the decisions. One of the most significant use cases for autonomous systems is the rapid speed at which a system can acquire, process and act upon information ultimately achieving information dominance. For many use cases, autonomy utility relates directly to information dominance. A formal assessment of information gain and loss is used to measure autonomy performance. By modeling the probability that an

autonomous systems knowledge is true (1st order) and the probability that the uncertainty assessment is true (2nd order) we are able to accurately represent the knowledge products produced by autonomous perception and represent the impact of uncertain knowledge on autonomous decision processes. Using test methods that model "who knew what when" and "how that information impacted mission effectiveness" these method results can be used to expose unknown vulnerabilities in the system resulting in a more robust resilient system. Using intelligent testing optimization helps to obtain more information from each test leads to a better understanding of the limits of the autonomy performance. The tests are based on teams of UAVs and there underlining behaviors. The system is evaluated based on conditions/parameters that have a direct impact on the system under test. Metrics collected during the test can range from robustness and resiliency such as communication quality during jamming, Speed compared to mobile targets, perception in the midst of a complex engagement area. Methods used to generate simulation-based test regimes include design of experiments/coverage testing. Coverage tests guarantees that each unique stimulus is examined at least once. Randomized testing (Monte Carlo), which provides statistical evidence for control variables that are themselves random (i.e., random failures). Intelligent testing which uses stochastic optimization techniques to iteratively design test conditions on the fly. Intelligent testing is useful in identifying boundary conditions and key "knee in the curve" performance indicators. And criticality-based testing which is used to identify safety limitations of autonomy performance. Techniques used to assess test result information include linear regression, analysis of variance(ANOVA), and nonparametric tests. These techniques are utilized to analyze the data collected from the parameter and metric values obtained during the span of the test. This data is then used to assess mission effectiveness and correlate mission effectiveness and information dominance. The results show the impact parameters have on overall information metrics. The next steps with this program is to modify system to support additional use cases including air, experiment with simulation, and conduct additional analysis. In conclusion, we are able to show how autonomy utility relates to information dominance.

## *Forge the Future with Computing*

Virginia To, Army Research Laboratory /Parsons

The Department of Defense (DOD) High Performance Computing Modernization Program (HPCMP) offer workforce development initiatives to provide future scientists and engineers the computational skills and experience necessary to close the gap between the technological capability and computational skills. Much of this is accomplished through the HPCMP Internship Program (HIP) that supports undergraduate and graduate student summer internship opportunities at DOD laboratories and test centers. Students learn high performance computing skills not readily available through colleges or universities. This presentation reviews the HIP internships at the Army Research Laboratory (ARL) conducted during the summer of 2020 and how ARL rapidly adapted to the challenges of COVID-19, creating a successful virtual summer intern program that grew even larger in the summer of 2021.

## *Globally Deployable Maritime Tracking Range Capability for Maritime Training, Mission Rehearsal, Test and Evaluation*

Steve Powell & Rogiero Ramos, Qinetiq & Darryl Newborough, Sonardyne

There is an increasing requirement for navies to be forward deployed due to an increasing pressure on platform availability. Consequently training and mission rehearsal (T&MR) as well as test and evaluation (T&E) activities that are traditionally conducted in home waters face a need to be conducted overseas. To meet this emerging requirement new technical solutions are needed that allow T&MR an T&E to be delivered more flexibly. To address this challenge QinetiQ and Sonardyne have been collaborating to develop a seabed

deployable ultra-short baseline (USBL) based tracking range. This novel capability has been designed to enable maritime test, training and evaluation of naval fleets whilst overseas. The deployable tracking range exploits COTS USBL systems, used in the Oil and Gas industry, alongside robotic and autonomous systems to deliver a maritime tracking capability. This paper reports on in-water demonstration activities conducted by Sonardyne and QinetiQ. The results show how commercially available above- and under-water tracking technologies can be integrated to create fully mobile and extendable ranges in support of maritime T&MR and T&E.

---

## *Immersive Environments for Visual Analytics*

James Hughes, UMD College Park; Selena Hamilton, Penn State University; Virginia To, US Army Research Laboratory /Parsons

Collecting test data is not a problem; however, managing and analyzing the magnitude of collected data is more than challenging. The Army Research Laboratory DoD Supercomputing Resource Center in collaboration with the Aberdeen Test Center or more generally, the Army Test and Evaluation Command, are applying scientific computing and data analytics methodology to test and evaluation. This project successfully demonstrated the use of immersive environments to render the movement of the U.S. Army's Expedient Leader Follower program vehicles under test.

---

## *Integrating Machine Learning and Interactive Visualizations for Heterogeneous T&E Data Analysis*

Mariya Occorso, Arcadia University; Vincent Perry, US Army Research Laboratory

The T&E data analysis and review process is so complex that it often takes 3 to 6 months to complete due to a lack of automation and data accessibility. With test programs producing thousands of incident reports per system under test, evaluators spend ample time reviewing and scoring the incidents to determine system readiness. Our work intends to expedite the evaluation phase by integrating machine learning and visual analytics into the data analysis and scoring process for evaluators. Given the multitude of incident reports recorded for a system, machine learning algorithms can be used to augment evaluation. These include keyword extraction and text summarization to reduce the burden of reading through lines of an incident report to uncover the cause, clustering of similar incidents that provide evaluators with virtual stacks of incidents that should be scored similarly or have inherent similarities to past incidents, and classification of incidents using historical scored test incidents to support evaluators scoring the thousands of incident reports. In addition, correlating the instrumentation data from the physical system under test to these incident reports opens a world of opportunity to learn from the data to both inform and justify the evaluators' scores. Due to the heterogeneity of the data collected during test events, high-performance computing resources have been utilized to correlate these data and provide scalable software to conduct analysis. In order for the evaluators to utilize these automated tools, we created an interactive dashboard visualization suite that portrays these time-saving techniques when evaluating incidents. Using Python's Dash framework, data-driven visual analytics reduces the burden of evaluators reading through incidents line-by-line. Pulling from multiple back-end data sources, the dashboard displays visualizations of instrumentation data over time during test, when incidents occurred during those tests, the ability to filter based off parameters in the instrumentation data, all with the ability to zoom and filter interactively. Currently, the tool can display relevant incident reports based off a time range and allows similar incidents to be displayed based off a user-inputted query. As the tool progresses, the plan is to incorporate more advanced clustering techniques of incident reports using user-defined metrics into the

dashboard, as well as incorporating an online-scoring model that learns from the evaluators' final scoring evaluation. Future endeavors beyond data correlation are to provide anomaly detection, pattern recognition, and both root-cause and predictive analysis of incidents during test. As the size of the data continues to increase, the high-performance computing resources will be a necessity to produce real-time analysis as evaluators interact with the visualizations to zoom and filter their data. Ultimately, this tool will provide evaluators with data-driven algorithmic analysis to expedite their own evaluation workflows.

---

## *Looking at Legacy I&V through an Agile Lens*

Scott A. Goodman & Doug Moshier, Raytheon Missile & Defense

Legacy programs have historic 'baggage' from years (if not decades) of I&V efforts. While new features and functions quite often add to the pile of testing to be done for system verification, little is rarely done to clean out the backlog of testing and approach the verification efforts more holistically. Recently on the Cooperative Engagement Capability Program, we analyzed our current processes, and using the tenets of Agile (like 'Continuous Deployment' and 'Development Pipeline') and DevSecOps we redistributed I&V throughout the program development life cycle. This effort covers the five fundamental changes we made to our system to increase test coverage while decreasing overall lifecycle cost.

---

## *Making Flight Test Instrumentation Setup Easier*

Ben Kupferschmidt, Curtiss-Wright

Modern Flight Test Instrumentation systems are highly customizable and complex. This complexity has steadily increased as data acquisitions systems have grown larger and more data has been collected on each test flight. While it is desirable for data acquisition systems to be very flexible, that flexibility makes configuring the data acquisition system harder for the end-user. Flight Test Instrumentation systems are also monitoring a larger variety of data source during a typical test flight. Many of these new data sources are digital buses where there are many possible messages that can be sampled. This leads to the desire to reconfigure the data acquisition system more often to accommodate the unique needs of each test flight. These factors contribute to making it more difficult to configure Flight Test Instrumentation systems. As the same time that complexity has increased, the flight test industry is facing challenges as experienced engineers retire. Thus, there is a strong need to make Flight Test Instrumentation setup easier. This paper explores how Curtiss-Wright approached this problem and made setting up a Flight Test Instrumentation system easier through improvements in our TTCWare setup software.

---

## *Mission Engineering, Capability Evaluation & Digital Engineering Informing DoD Technology, Prototyping and Acquisition Decisions*

Susan Beers, PhD, OUSD(R&E)/DTE&A (The MITRE Corporation)

This presentation will describe an approach to unifying capability evaluation to inform acquisition and operational cradle-to-grave lifecycle decision-making. The approach includes combining currently separate developmental and operational test's evaluation planning into a single evaluation framework, adding in a M&S VV&A decision support evaluation framework into a Unified Evaluation Framework (UEF). The approach also unifies test and evaluation (T&E) planning with modern systems engineering (SE) methods. We establish a vision using Mission Engineering (ME) principles to enable the definition and linkage of operational and technical capabilities, while the analytical rigor of Digital Engineering practices will provide the evaluation-based information needed for acquisition, engineering, and operational decision-support in order to improve

system design and mission outcomes.

## *Modernizing T&E is Much Easier to Say…*

Richard Cawthorne, Defence Science and Technology Laboratory

The Defence Science and Technology Laboratory (DSTL), which is part of the UK Ministry of Defence (MOD), instigated a small four year project (funded by the MOD's Chief Scientific Adviser) with an aspiration of modernising some of its Test and Evaluation (T&E) capabilities. Three years into the project I'll share a journey so far that has taken us from a "wouldn't it be nice if we could…" conversation, to delivering practical tools we now use in business as usual, alongside a mental shift in our attitude towards a "this is how we will" mind set.

Using examples from particular activities performed within the project, we'll present some of the challenges faced and approaches we have taken, sharing what we found worked for us as well as what didn't. Topics will include: how do you use new tools such as augmented reality when nobody in the team has used it before, you can't recruit and security won't let the technology in the building; how do you convince a practical, hands on workshop engineer not to be afraid a robot is trying to steal their job; why machine learning will answer your questions but presents you with a whole new, longer set of questions.

We'll include how the 'importance of S&T research' narrative in the UK MOD Science and Technology (S&T) Strategy has changed during the period of research and how we have supported this, reshaping the activities and approaches being taken in MOD's wider T&E community.

## *Path Planning for Multi-Agent Semi-Autonomous Rover in an Integrated System Test Facility Without the Aid of GPS*

Lauren Blanks, Allyson Burba, Isaac Isukapati & Col Brian J Neff, PhD, U.S. Air Force Academy and National Robotics Engineering Center

Complete end-to-end testing of installed systems on an aircraft within a free-space radio frequency (RF) environment is crucial to understanding and developing its combat capabilities. The Benefield Anechoic Facility (BAF) at Edwards Air Force Base is one of the few facilities capable of measuring far-field coverage patterns of systems under test, including simultaneous measurements of multiple installed antennas for most aircraft in the Air Force inventory. However, the current test environment requires human intervention to manually reposition and reorient the emitters between discrete test points. The task of reorientation emitters within the chamber can be both cumbersome and time-consuming. Considering the current demand on integrated system test facilities such as the BAF, any lost time due to test inefficiencies directly contributes to missed opportunities for the test community. Additionally, stationary emitters only allow for snapshot characterizations of the aircraft's response since the manual set-up only allows for the test to be run in the specific and individual scenarios. Advances in distributed computing and multi-agent coordination techniques make it possible to not only automate the task of reorienting the emitters between test points, but also to simulate a continuous testing environment. As a first step in this process, this paper seeks to demonstrate the possibility of designing an autonomous process through the design of an auction-based scheduler capable of generating the optimal tour for a given set of rovers and user-defined waypoints that they ought to traverse. We propose to benchmark the performance of the system with that of current operations. Simultaneous efforts are underway to develop a semi-autonomous rover that can be used to dynamically and precisely orient emitters inside of a large integrated system test facility without the aid of GPS or RF based navigation techniques. This rover will couple vision-based navigation with odometry to precisely navigate the path specified by the multi-

agent planner above.

## *Seminal Analysis Methodology Based on Uncorrelated Data Collection in Either Benign or Contested Environments*

Nathan A. Ruprecht & Elisa N. Carrillo, 746 Test Squadron

A continual question in test is to answer how many data points are required to produce statistically significant results for decision makers and what constitutes an event. Especially in constant benign or contested environments, if each sample or a second counts as a data point. At least with flight test, profiles are determined beforehand to know limitations and parameters. With a planned profile and therefore target time, space, and position information (TSPI), this methodology can be ran prior to test to predict the number of data points to be collected. Optimized test planning will reduce test time while maintaining test rigor and data integrity. This methodology is intended for inertial guidance, GPS, or blended navigation units as universal use, it's computationally heavy which is made doable with the processing power of today's computers. This abstract is a continuation of research to both develop and characterize this methodology.

## *SyncVis, a Framework for 2D and 3D Visual Analytics*

Simon Su, PhD, Olena Tkachenko, Vincent Perry, Michael An, DEVCOM ARL

The Department of Defense Research, Development, Test and Evaluation (RDT&E) community generates and analyzes heterogeneous 2D and 3D data which includes network, application, vehicle, weapon, communication device, data link, and simulation data requiring new analysis approaches to obtain usable information from the data. Presently, there is no established framework or application program available on the DoD High Performance Computing Modernization Program (HPCMP) systems that can be leveraged to support visualization of heterogeneous 2D and 3D data processed on the HPC resources. DoD users require a visual analytics tool capable of supporting data analysis and visualization of the large complex heterogeneous data. A hybrid 2D and 3D Visual Analytics tool, SyncVis, previously prototyped has generated positive feedbacks from multiple users. However, our SyncVis prototype was at Technology Readiness Level (TRL) 1 or 2, far from a production level tool that can be deployed for general use, at TRL 6. In an attempt to harden SyncVis to a higher TRL level, we have been working on improving the stability of the prototype and develop the visualization framework to run completely on the HPC resources.

## *System of Systems Synthetic Evaluation*

Mark Threadgold, UK Ministry of Defence, Defence Science and Technology Laboratory (DSTL)

The UK Ministry of Defence (MOD) have a requirement to develop future operational capabilities and assure their effectiveness within a representative contested ElectroMagnetic Environment (EME). However, current and future platforms can no longer be considered as a collection of disparate systems, but instead form a fully integrated system of systems both within a single platform and federated across multiple platforms. Furthermore, the connectivity of systems to form a more complex integrated capability extends to the hostile environment in which our platforms are required to operate.

The role of Test & Evaluation in this develop and assure process is essential, but challenging. Current live evaluation approaches are severely hampered in their ability to bring together sufficient elements to represent the fully integrated nature of multiple platforms within a complex and congested EME. While synthetic T&E tools exist to support the approach, they typically trade detail and fidelity for scale. As such, traditional

synthetic approaches either provide physics level evaluation at the individual sub-system layer, or analytical level evaluation when increasing to mission-level interactions.

With a drive towards increasing use of complex synthetic evaluation, the MOD Chief Scientific Advisor (CSA) has commissioned research into new approaches towards mixed fidelity synthetic evaluation at the system of systems level. This System of Systems (SOS) research is fundamentally looking to reuse existing synthetic evaluation tools where feasible, but enabling multiple different tools to be used within a single evaluation framework to provide the scale demanded by the end user. Furthermore, the project is harnessing Commercial Off The Shelf (COTS) tools and technologies to provide a flexible System of Systems Architecture Design (SOSAD) that allows mixed fidelity evaluation across the breadth of the EME.

Initial demonstrations of the SOSAD approach have already shown a capability for conducting synthetic T&E that encompasses physics and analytical level models, as well as hardware in the loop facilities, to deliver complex air domain T&E for assurance of military capability within a contested environment. Future work is planned to extend this approach into multi-domain environments, while also progressing concepts for a federated capability that brings together MOD, industrial and international capabilities into a single synthetic T&E construct.

---

## *T&E: Required Leadership in The Age of Transformation*

Paul McNamara, The Sente Group, Inc.

Rapid technological change has produced an environment ripe with threats and opportunities. In response, many large and established companies have committed to transformation initiatives. These initiatives aim at making step-change improvements to speed and cost. Those who learn to exploit this rapid change first will have a huge competitive advantage. They will set themselves up to dominate for a generation or more. Those who don't will lose and lose big. It is a high-stakes game, and everyone knows it.

Complicating efforts, long-established methods for improvements are no longer up to the task. They produce incremental improvements in a world requiring step-change results. Horse and buggy manufacturers could have completely leaned out their operations and it wouldn't have mattered. People wanted and then needed the automobile. The horse and buggy went extinct. It was overthrown by a new industry that produced sustained growth for decades. Given test's location and cost in the development process and its inherent complexity, it has become an important target for improvement efforts.

The question for test professionals is are they willing to lead? Will they do what it takes to avoid extinction and to get on the happy path of business and career growth? In this presentation, we share our approach.

Software tools are not the biggest barrier to success although they are essential. Gartner and others have identified culture as the biggest challenge to transformation. Given this, there are many questions that need answering like what does this even mean? What cultural changes must take place? How must leaders structure projects for success? Who must join and where can software tools speed up culture change and results?

The presentation will answer these questions. It will share our approach and some of the practices we've used to produce the following results:

- 50% plus improvements to test speed and throughput
- 400% improvements to test equipment and facility utilization
- 99%+ asset database accuracy (by changing the culture)
- Career growth for those leaders who took on the challenge

We will present a new approach to transforming test operations with a fresh approach to asset management. To do this we will expand on historical notions of how to manage constraints to speed and throughput. We will distinguish "culture" and the role it plays as a constraint in complex operations. We will include

approaches and practices appropriated from leaders like Bezos, Musk, Jobs, and others. These approaches are essential for dealing with human constraints. Next, we will show how to leverage technology to produce results at scale. Finally, we will address Agile's role in these efforts.

---

## *Testing for the Future Fight: Necessary Range Capabilities to Ensure Operational Superiority of Defense Systems*

Heidi Perry, MIT Lincoln Laboratories; Hans Miller, MITRE Corporation; Derrick Hinton, SRC

The National Academies of Sciences, Engineering, and Medicine will release a report in early September 2021 assessing the physical and technical suitability of DoD test and evaluation ranges and infrastructure across all domains (land, sea, air, space, and cyberspace). This study was commissioned by the DoD Director for Operational Test and Evaluation and the report will include a comprehensive set of key findings and recommendations to address range deficiencies for testing technologies anticipated to arrive between now and 2035. The report findings are based on testimony from senior military officers and officials from operational, acquisition, and test backgrounds, as well as test and training experts, leading technologists, leaders from relevant commercial enterprises, and individuals with experience in DoD and Congressional budget processes. The report and presentation focus on three fundamental themes: 1) future combat will demand connected kill chains in a joint all-domain operations (JADO) environment; 2) digital technologies are dramatically reshaping the nature, practice, and infrastructure of test; and 3) speed-to-field is today's measure of operational relevance. The presentation will include the report recommendations, which span five broad categories: 1) develop the "range of the future" to test complete kill chains in JADO environments; 2) restructure the range capability requirements process for continuous modernization and sustainment; 3) bootstrap a new range operating system for ubiquitous M&S throughout the weapon system development and test life-cycle; 4) create the "TestDevOps" digital infrastructure for future operational test and seamless range enterprise interoperability; and 5) reinvent the range enterprise funding model for responsiveness, effectiveness, and flexibility.

---

## *Testing Mixed-Use Networks – Insights Into a Complex Test Process*

Dan Pleasant, Jarrod Tsukada & Kelly Caulk, Keysight Technologies, Inc.

Near-future battlefield environments will contain dynamic ad-hoc networks that include not just communications but also radar/EW and cyber offensive and defensive capabilities. RF signals will be subjected to jamming and spoofing by adversaries, while users must be able to reliably consolidate all of the available information into a single view of the current situation. These systems must be robust against all enemy attacks and survive – and reconfigure themselves – even in the face of lost network nodes, cyber attacks, or even endpoints that have been captured by the enemy. It's critical that such networks be test thoroughly against hundreds or thousands of scenarios before they can be fielded. This paper gives an overview of the kinds of techniques that will be used to test them, and gives some initial insights into the requirements that must be met by the network to allow it to be tested at all. Thinking further ahead, some backend services for warfighter applications will move from the edge to the cloud. Validating network resilience and security against assets that cannot be physically accessed presents a whole new set of testing challenges.

---

## *The Challenges of Agile Software Delivery for Military Embedded Software Solutions*

Chris Olinde, Andrew Mishoe, Stephens Summerlin & Ken Bailey, Georgia Tech Research Institute

Significant progress has been made in the maturation of tools and processes to shorten the development life

cycle of software with the goal of providing continuous integration and delivery of high-quality software, commonly referred to as DevOps. There remain several challenges to deploying these practices easily on packaged embedded software solutions for military applications. Two specific challenges will be addressed in this technical session – navigating security requirements and software build reproducibility.
The first challenge is navigating security requirements, while minimizing vulnerability of the system while simultaneously enabling unclassified development and continuous delivery.
A second challenge is meeting the sponsor requirement of delivery of source code and reproducing executable binaries. The challenge is primarily due to the increased complexity of the build and test pipeline, which has to be replicated at a sponsor site in order to satisfy all external dependencies to reproduce software. The effort to standardize around common pipelines and environments frequently makes assumptions about hardware, storage access, and network connectivity that do not hold for secure embedded systems.
Finally, the unique nature of embedded system hardware and interfaces creates tight coupling between components and increases test complexity. Often, availability of hardware or representative interfaces lead to bottlenecks which can limit a team's ability to test and verify functionality.
This paper will present the roadblocks encountered during a recent DoD embedded system development project, the approach that was taken to overcome each issue, and lessons learned that should help to enable continuous delivery to embedded systems of the future.

---

## *The Queensland Flight Test Range (QFTR)*

Rob Dunn & Jim Parkes, QinetiQ Australia Pty Ltd

The Queensland State Government has a vision to be a world leader in Unmanned Aerial System (UAS) technology and their applications. A key enabler in achieving this vision is the development of the Queensland Flight Test Range (QFTR). The QFTR is a Queensland Government sponsored facility operated by QinetiQ Australia. The QFTR is located at Cloncurry Aerodrome in north-west Queensland.
The QFTR is a commercially operated facility available for both civilian and military use and provides the ideal location to test, trial and evaluate UAS and payloads in a safe, managed environment. The QFTR has relatively low levels of conventional aircraft traffic, an excellent 2000-metre primary runway and a graded 1150-metre secondary runway.
The surrounding uncontrolled Class G airspace has low traffic levels and large areas of unpopulated land below. There are significant airspace volumes of varying dimensions available to meet specific UAS trial requirements. Whilst similar capabilities exist in Europe and North America, this commercially operated Range is unique to Australia and the Asia Pacific (APAC) region. It is designed to cater for large, high speed platforms, platforms with long endurance and/or high-altitude attributes, and teaming/multi-platform operations in a common user environment.
QinetiQ used first-principles risk management approach to develop the operational procedures for the QFTR in a regulatory environment silent on Flight Test Range operations.
The QFTR currently offers:

- Approved and licensed operators to fly and test UAS of all weight classes in a safe and controlled environment;
- BVLOS approvals through the provision of ground based surveillance systems;
- Airspace design and approvals to meet the needs of the user trials and manage ground and air risk;
- Safety Case development (SORA based) and Regulatory approvals;
- Local aviation stakeholder/land owner engagement on behalf of users;
- Range surveillance safety services through tactical identification of air risks;
- Operations under a QinetiQ Remotely Piloted Aircraft Operator's Certificate (ReOC);

- Emergency response planning and coordination; and
- Provision of aeronautical information and weather briefings to facilitate FTR operations.

The option for QFTR staff to secure complex regulatory approvals on behalf of users within existing regulatory frameworks, including detailed SORA applications and extensive stakeholder engagement, reduces the time and resource burden on users of the Range.
Additionally, QinetiQ holds a ReOC and can offer an induction service for users that may not hold an Operator's Certificate. This is particularly beneficial for smaller operators or international clients.
With planned future capability enhancements, the Range will expand into a broader Aerospace Test Facility that will facilitate:

- Complex flight trials and payload testing supported by a Live Virtual Constructive (LVC) capability;
- Integration of T&E instrumentation with more complex mission control systems and infrastructure;
- Counter UAS research;
- Manned/Unmanned Teaming testing;
- UAS air weapon testing; and
- Long range and guided munitions testing.

"The QFTR is a critical component of the overall drone ecosystem, as both an enabler and connector of high-technology UAS programs and initiatives." Greg Barsby - Managing Director QinetiQ Australia.

---

## *The Use and Benefits of Modeling and Simulation with Autonomous Testing*

John Whitt & Paul Bounker, Aberdeen Test Center/Ground Vehicle Systems Center

Autonomous systems are becoming more and more pertinent in today's world. The question that comes into play is how do you accurately and effectively test these systems to ensure safety and functionality? As these systems evolve so have modeling and simulation tools. Current gaming engines have high fidelity physics based simulation capabilities making it easier to simulate autonomous vehicle testing. Using these tools, testers can now integrate terrain models, mobility models, sensor models, and autonomy software together creating a virtual testing ground. Simulations will now be able to take digital twins of autonomous vehicles and have them interact with weather, obstacles, and even degraded sensors. With the use of these tools, testing can now be broken down into three stages. The first stage would require the autonomy software to be provided and integrated into the simulation. Once it is integrated, as vehicles undergo development before prototypes are built, thousands of simulations can be run using the autonomy kits software. This saves money as no hardware has to be built. Issues noticed during the simulations can be fixed in the software before prototypes are ever built. With the use of a high-performance computer, thousands of scenarios can be run quickly and returned producing edge case scenarios. Even still, pass and fail criteria can be assigned, and the results coming from the high-performance computer can not only give you edge case scenarios but also how many times it passed and failed. After the software is developed enough, the software can now progress to stage 2. Stage 2 testing involves the simulation mentioned in stage 1 but adding a Hardware in the Loop (HITL) element. HITL allows the use of simulators that already exist such as Aberdeen Test Center's (ATC) Roadway Simulator (RWS). The RWS is a treadmill for military vehicles. The RWS can integrate a live autonomous vehicle with virtual autonomous vehicles as well as virtual weather and obstacles. Stage 2 ensures that the software is not telling the vehicle to perform certain functions that the actual hardware is not capable of. One advantage of this approach is you have taken your 1000s of simulation and have reduced it down to 100s while still only using one live prototype. After the goals of stage 2 are met, the autonomous vehicle under test can now proceed to stage 3. Stage 3 involves live testing of the vehicles in a controlled environment. This environment involves

using a hardware/software applique kit that is installed over top of the autonomy software. This kit would shut the vehicle down if it went out of locational boundaries, came too close to obstacles, noticed a vehicle failure, or noticed a change in status of the emergency stop that part of the vehicle. How the vehicle is stopped and when it is stopped is completely configurable. Stages 1, 2 and 3 are already being implemented as part of a project called the Autonomous Systems Test Capability (ASTC).

---

## *Topic Modeling and Visualization for Test Incident Report Data*

Sage Leone, University of Maryland; Jamie Infantolino, US Army Research Laboratory; Adam Childs & Cleon Anderson, US Army Research Laboratory/Parsons

Testing at the U.S. Army Aberdeen Testing Center (ATC) is producing data at an increasing volume, velocity, and variety. In order to handle this growing amount of data, less resource-heavy approaches to analysis need to be implemented. This research leverages the big data resources available at the Army Research Laboratory High Performance Computing (HPC) center to create an automated classification system for data records. Using Apache Spark's MLlib and SQL libraries, a topic model using Latent Dirichlet Allocation was developed to quickly and correctly categorize Test Incident Reports from vehicle testing at ATC. These results were compiled into clear visualizations, allowing trends to be quickly observed and used to directly improve the testing process. Building on previous research on the use of HPC resources, this topic model further reduces the time and personnel spent on data analysis. In the future, further integration with Hadoop, Hive, and new Spark Streaming technologies, as well as generalizing this product for other datasets, could lead to faster and more widely accessible results.

---

## *Use of DOE Tools in a Development Laboratory*

Jasveer Ramroo Beni, Master Builders Solutions Construction Chemicals, LLC

Our technical team throughout the ORA (Orient-Russia-Africa region) underwent training for Design for Six Sigma with the aim to improve the efficiency of our laboratory and development activities. As a result, one of the concepts from Design for Six Sigma is the Design of Experiment (DoeE) which we nowadays use in several of our research and development projects. DoEs help us to use our technical expertise in a streamlined and scientific approach for several benefits such as:

- Screening large numbers of raw materials for a particular project
- Identifying interesting interactions which impact product performance
- Building predictive modelling or formulation guidelines
- Optimization of existing product formulations with respect to cost or performance benefits
- Sharing of information across countries using statistical tools, and data derived from DOE is like having one common language

I will put up a short presentation, summarizing the above points with some examples of completed Design of Experiments and data analysis and how the above-mentioned benefits contribute to the scientific testing and evaluation in a R&D lab as ours.

---

## *Using Natural Language Processing to Construct a Knowledge Graph of Test Incident Reports*

Sagar Indurkhya, PhD, Hector Vazquez, Aakash Indurkhya, and Ciro Donalek, Virtualitics, Inc.

It is often the case that a Test Incident Report (TIR) makes note of an observation that has been reported in many other TIRs, so that the noted observation is an example of a pattern. The Test and Evaluation (T&E) operator desires to identify such patterns so that related observations can be consolidated, a comprehensive solution can be developed, and duplication of work can be avoided. However, the identification of such a pattern is slow and difficult because of the tedious manual effort required to identify members of the large corpus of TIRs that are relevant (with respect to the observation of interest) and then analyze the large volumes of data associated with each identified TIR. To this end, we have developed a computational procedure for automatically identifying common patterns observed across many TIRs and producing relevant visualizations of the datasets associated with the identified TIRs.

The procedure takes as input a corpus of TIRs and associated datasets (of sensor data). The procedure outputs a Knowledge Graph, which is a type of network graph in which each node corresponds to a TIR and two nodes are connected by an edge if the two corresponding TIRs are semantically related; two TIRs are semantically related if they make the same observation (e.g. "the engine is leaking coolant") even if each TIR expresses the observation differently (e.g. "the engine is leaking coolant" vs. "coolant has leaked out of the engine"). The procedure takes the form of a Natural Language Processing (NLP) pipeline, consisting of a sequence of stages involving: the preprocessing and tokenization of each TIR in the input corpus, recognition of named entities, alias resolution, semantic parsing (using a Universal Dependency parser), and extraction of event-representations (i.e. a structural representation of the semantic roles encoding knowledge of "who did what to whom") via the application of a (small) set of structural transformations.

A T&E operator can query the Knowledge Graph output by the procedure in order to automatically identify subsets of the corpus of TIRs, with the TIRs in each identified subset being semantically related to one another (and thus likely to be sharing related observations that collectively form a pattern). Additionally, the Knowledge Graph can be queried in more specific ways by requiring that the identified subsets either be semantically related to a specific TIR of interest or pertain to a specified named-entity or event - in this way, a T&E operator can rapidly search the corpus of TIRs to identify related TIRs that are relevant to an observation of interest. Finally, each identified subset of TIRs is automatically associated with a set of 2D and 3D visualizations of relevant sensor data (i.e. sensor data that corresponds to the named-entities and events extracted by the NLP pipeline that are common to the identified subset of TIRs) – these visualizations can help a T&E operator quickly perform a quantitative analysis of the identified pattern.

## *Value of Testing Certifications for Agile & DevSecOps*

Andrew Pollner, American Software Testing Qualifications Board

Testing is in a state of evolution as Agile, DevOps, and now DevSecOps are the new methods and practices by which software-intensive systems are developed and maintained. Great efficiencies have been gained with automation playing an ever larger role in development, testing, and deployment. With a "shift-left" mindset, T&E professionals are now more involved, and earlier in the process. The challenge for T&E professionals is to learn and adapt skills to better align with this new approach to the software lifecycle. Test techniques need to be honed, and automation needs to be better understood and applied. Automation is not only used in testing but has been integrated to quickly and efficiently build and move code from unit to system testing, then to production, and finally with monitoring, feedback, and continuous improvement steps that lead to increased quality. This iterative process adds overall complexity which needs to be understood by T&E professionals so that it can be managed and mitigated when problems surface. ASTQB facilitates and enables upskilling through a rich body of knowledge in testing that addresses this evolving software lifecycle shift while ultimately helping advance the T&E profession.

*9/2/21 Rev L*

## *Visual Analytic Application Toolset for Large-Scale Data*

Aiden Kenny, Columbia University; Mariya Occorso, Arcadia University; Vincent Perry, US Army Research Laboratory

The goal of our project was to build the framework for an interactive user dashboard to help analyze large-scale data sets, with specific emphasis on creating interactive visuals. We focused our dashboard on data collected from the Army's Expedient Leader Follower program, since the project's success is of great importance to the Army and much data has been collected. We wrote code in R that automatically accesses the database and then cleans the data and returns useful summary statistics. We also wrote code in Python that build the actual dashboard itself and accepts the cleaned data and produces interactive visualizations. The result is an interface that allows a user to easily create visualizations of the data that allows them to derive important insights.