Edge-Enabled Tactical Systems

Grace A. Lewis

Jeff Boleng

Software Engineering Institute
Carnegie Mellon University
Pittsburgh, PA 15213

This material is based upon work funded and supported by the Department of Defense under Contract No. FA8721-05-C-0003 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center.

Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the United States Department of Defense.

DM-0002794

Background: Focus on Small, Tactical Units

Tactical units are losing decisive advantage as adversaries adapt to and adopt new technologies at a pace faster than our refresh cycles.

EETS investigates and adapts cutting-edge technologies and builds prototypes of assured, efficient and rapidly-fieldable systems for enhanced SA and decision support for tactical users

For FY15/FY16 we are addressing assurance

- Trusted Nodes
  - Trusted identities in disconnected tactical environments
- Confidence in Information
  - Confidence in information derived from social media streams
  - Confidence in information through fused social and physical sensor data

Prior Results

Tactical Cloudlets

Forward-deployed, discoverable, virtual machine (VM) based cloudlets that can be hosted on vehicles or other platforms to provide infrastructure to (1) offload computation, (2) forward data-staging for a mission, (3) filter data to remove unnecessary data from streams intended for dismounted warfighters, (4) serve as collection points for data heading for enterprise repositories

Edge Analytics (EA)

Infrastructure for end-to-end near-real-time analysis (seconds to minutes) of social media and other sensor streams to provide actionable intelligence, trends, and summaries in resource-constrained edge environments

Information Superiority to the Edge (ISE)

Group-context-aware reference architecture, middleware, data model, and prototype implementation to reduce cognitive load and conserve resources by using sensor, role/task, and event information to deliver the right information, at the right time, to the right soldier

FY15 Results and Accomplishments

Prototypes in Action

- EA is being used by AF/A2I to analyze OSINT (Evaluation)
- EA used by PA 3rd WMD CST at Boston Marathon and in Pope Francis visit (Deployment)
- EA integration with MIT LL NICS and DCGS for first responder inter-agency board (Demonstration)
- EA/GCF integration deployed at CreationFest for public safety awareness (Experiment)
- ISE, DTN, and reliable UDP integrated with PRC-117(g) radios for USMC TID (Field Demonstration)
- EETS architectures and research experience are being applied to TALOS (Rapid Prototyping)
- KD-Cloudlets: KVM-based Discoverable Cloudlets (Open Source Project)

Publications (FY15 Only)

- Tactical Cloudlets: Moving Cloud Computing to the Edge, MilCom 2014, October 2014
- Fusing Open Source Intelligence and Handheld Situational Awareness: Benghazi Case Study, MilCom 2014, October 2014
- On-Demand VM Provisioning for Cloudlet-Based Cyber-Foraging in Resource-Constrained Environments, MobiCASE 2014, November 2014
- Using Multiple Contexts to Detect and Form Opportunistic Groups, CSCW 2015, March 2015
- The Group Context Framework: An Extensible Toolkit for Opportunistic Grouping and Collaboration, CSCW 2015, March 2015
- A Catalog of Architectural Tactics for Cyber-Foraging, WICSA/CompArch 2015, May 2015
- Cloudlet-Based Cyber-Foraging in Resource-Limited Environments, Book chapter for IGI Global publication on Emerging Research in Cloud Distributed Computing Systems, June 2015
- Characterization of Cyber-Foraging Usage Contexts, ECSA 2015, September 2015
- Cyber-Foraging for Improving Survivability of Mobile Systems, MilCom 2015, October 2015

Task 1: Trusted Identities in Disconnected Environments — Summary

Problem: Systems operating in disconnected environments cannot rely on the availability of an online trusted authority to validate the credentials of a requester

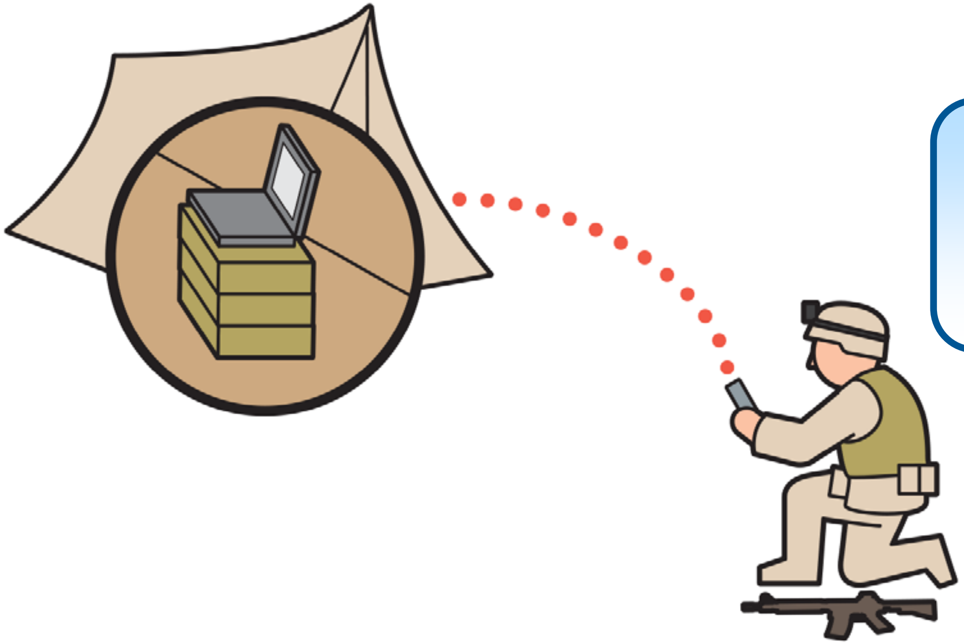
Solution: Create a trusted identity solution that provides greater or equal identity strength as traditional PKI and addresses threats of tactical environments

Tasks

- Threat model for tactical edge environments
- Review of research, open-source and industry solutions for trusted identities: *hardware, software, hybrid, human-centric*
- Validation of identity solutions against threat model
- Implementation of a selected solution in the context of our tactical cloudlet infrastructure
- Evaluation of implementation in tactical cloudlets prototype using threat model, vulnerability analyses, and identity strength metrics

Task 1 Context: Mobile Client/Server in Fully-Disconnected Environments

Is the service request coming from a valid mobile device?



Is the discovered node a real cloudlet?

No reach back to validate credentials

Task 1: Implementation

Step 1: Bootstrapping

- Generation of *Server Credentials* using IBE (Identity-Based Encryption)
- Setup of *RADIUS Server* with *Server Credentials*

Step 2: Pairing

- Generation of *Device Credentials* using IBE
- Transfer to device using Bluetooth or USB, plus visual confirmation
- Transfer to *RADIUS Server*

Step 3: WiFi Authentication

RADIUS Server implements Wi-Fi WPA2-Enterprise 802.1X EAP-TTLS with PAP

- Device receives server credentials and validates
- Devices sends its credentials for validation

Step 4: API Requests

- Device exchanges encrypted messages with the server
- Each exchange is validated against authorized device list

Termination

- Automatic due to timeout: Bootstrapping requires setting up mission length
- Manual due to known loss or compromise: Server Management component has revocation option

Task 1: Evaluation Against Threat Model

Fully Addressed by Implementation
1: Impersonating a Device
2: Finding an Active Client
3: Finding a Device
7: Sniffing Wireless
14: Server Impostor

Partially Addressed by Implementation
6: Lost Credentials (usability tradeoff)

Not Addressed by Implementation
4: Altered Software
5: Daisy Chaining

Addressed Outside the Implementation
8: Site Intrusion
9: On the Net
10: On the Box
11: Super-User Compromise
12: Application Compromise
13: Seeing Everything

Task 2: Assigning Credibility Scores to Social Media Streams in Real-time — Summary

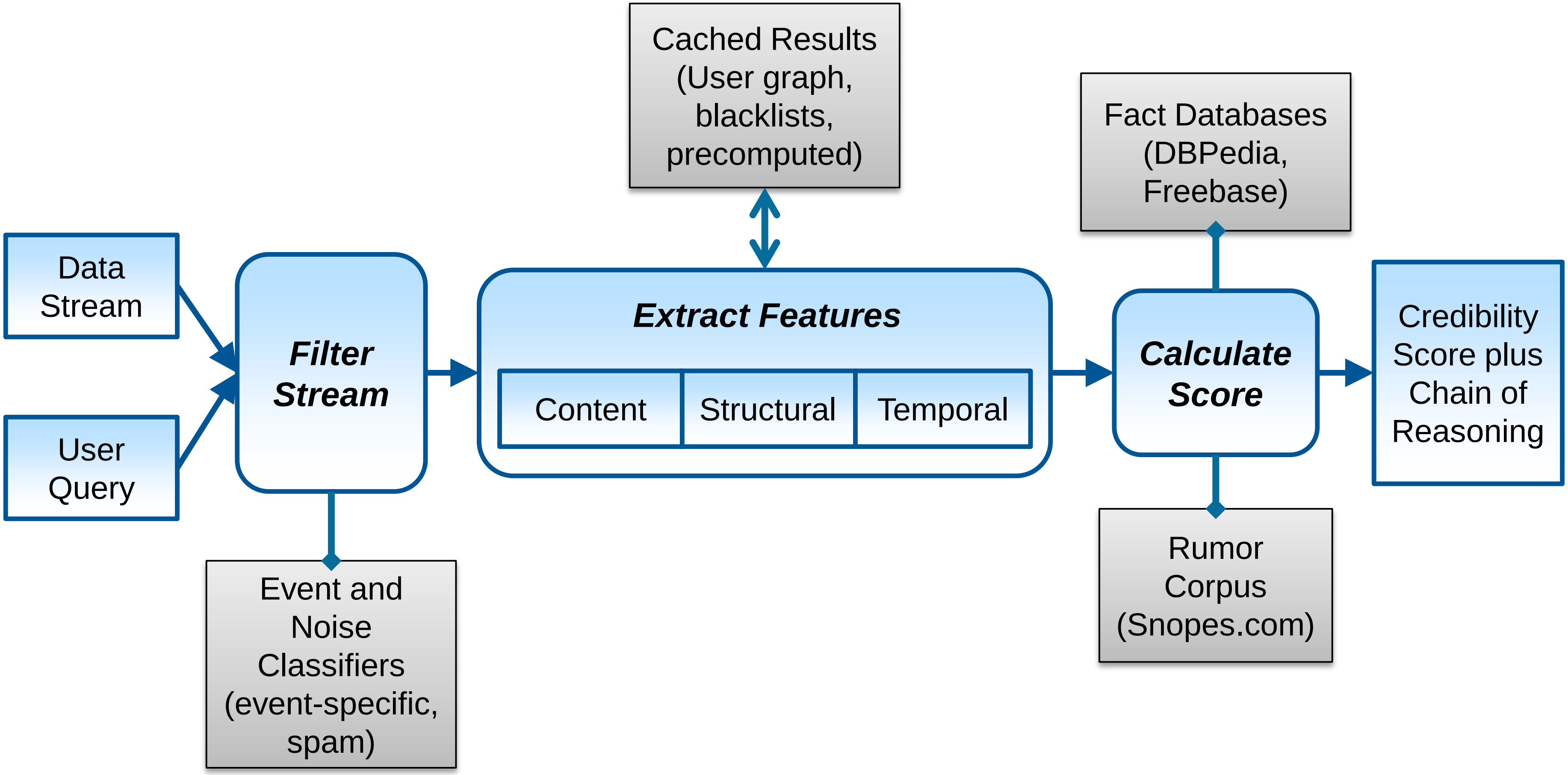
Problem: Social media data generated by adversaries can be mined for intelligence, but establishing trust in this content is difficult because the reliability of information is often poor due to spam, rumors, or compromised accounts.

Solution: Combine, enhance, and extend existing work to create a prototype that assigns credibility scores *quickly* and provides a *human-understandable chain of reasoning* to speed data-to-decision.

Tasks

- Literature review
- Prototypes of existing algorithms to completely understand their strengths and limitations
- Development of an algorithm to assign credibility scores and provide chains of reasoning to streaming social media data
- Architecture that implements developed algorithm based on an ensemble of existing algorithms
- Extension of existing Edge Analytics prototype to implement the architecture and validate research

Task 2: Credibility Calculation Pipeline


Cached Results (User graph, blacklists, precomputed)
Fact Databases (DBPedia, Freebase)
Data Stream
User Query
Filter Stream
Extract Features
Content
Structural
Temporal
Calculate Score
Credibility Score plus Chain of Reasoning
Event and Noise Classifiers (event-specific, spam)
Rumor Corpus (Snopes.com)

Task 2: Rule-Based Event Extraction

Process

For each tweet in dataset:

If no applicable rule already exists:

1. Write rule to extract veracity and rumor description, either by creating a new rule or modifying a similar rule
2. Test accuracy of rule
3. Test generalizability of rule

Challenges

- Balancing specificity and generality
- Involved and time consuming process
- Characteristics of Twitter data

Example

Their (sic) is a rumor going around John Doe will join #ISIS.

```
  name: "rumor-1"
  label: [Rumor]
  priority: 3
  type: token
  pattern: |
   (?<trigger>
   [lemma=/rumor/ & tag=/NN/]) (going
around|flying around|doing the rounds)
[tag="DT"]? [tag="WDT"]? [tag="NN"]?
[tag="IN"]? []*?
   @about: Word+ []*?
```

Conclusion

⇒ **Approach does not scale**

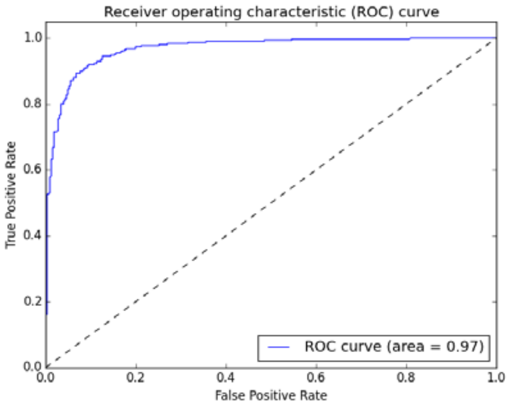
Task 2: Filter Stream

Binary classification of Twitter data; e.g., “shooting” and “non-shooting” categories.

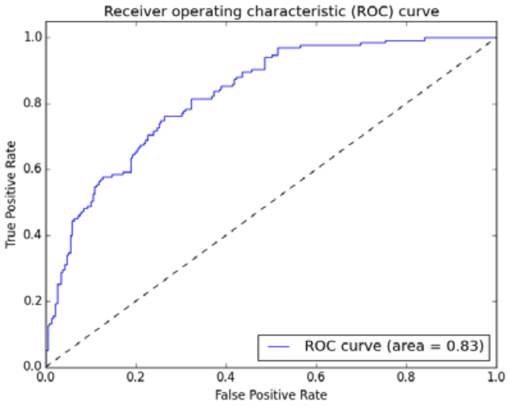
Classifiers Investigated

- Linear SVC (SVM) ⇒ **best**
- Random Forest
- Multinomial Naïve Bayes
- Logistic Regression

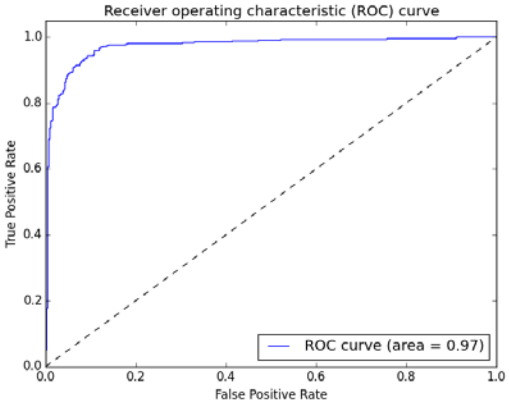
Classification Results (SVM)



Train, test on original dataset



Test on new data



Retrain, test on original dataset + new data

Task 3: Fusion of Social and Physical Sensor Data

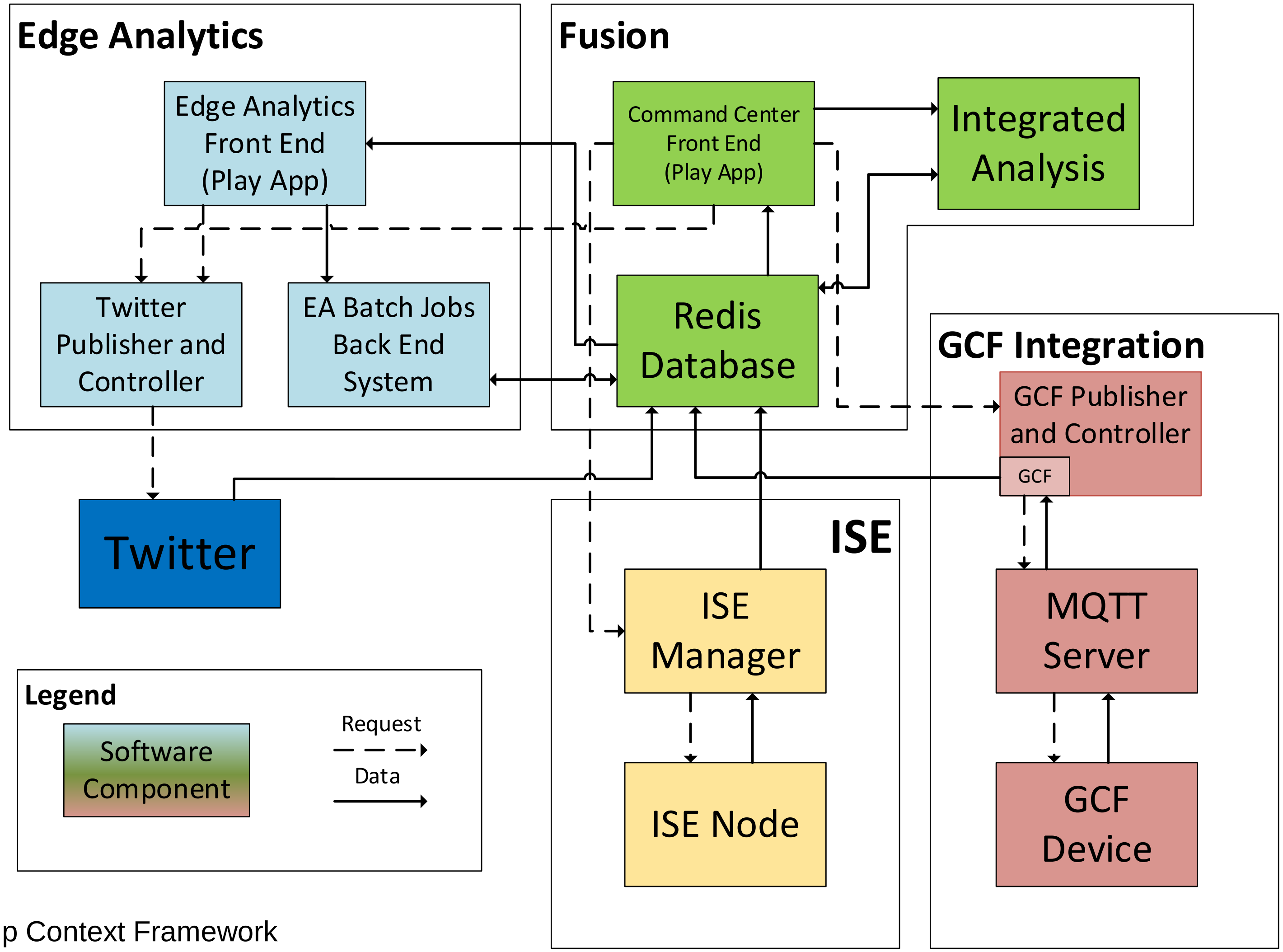
Problem: Fusion of electronic signals (SIGINT), images (IMINT), and other intel data with open source (OSINT) data is currently an analyst-intensive activity and is not available in tactical time frames.

Solution: Create a fusion strategy that relates open source intelligence from Twitter streams with images, audio, and potentially other data sources derived from opportunistic access to handheld devices.

Tasks

- Literature review
- Generation of common format for incoming sensor feeds that can be analyzed by Edge Analytics
- Build experimentation platform
- Validation of effectiveness of Edge Analytics with data streams incorporating social media and physically-sensed data
- Analysis of resulting capability for Performance, Usability, and Trust

Task 3: High-Level Architecture


Edge Analytics
Edge Analytics Front End (Play App)
Twitter Publisher and Controller
EA Batch Jobs Back End System
Twitter
Fusion
Command Center Front End (Play App)
Integrated Analysis
Redis Database
GCF Integration
GCF Publisher and Controller
GCF
MQTT Server
GCF Device
ISE
ISE Manager
ISE Node
Legend
Software Component
Request
Data

GCF = Group Context Framework
MQTT = MQ Telemetry Transport
ISE = Information Superiority to the Edge Research Prototype

Task 3: Relationships and Scenarios

Cooperative Context = ISE research prototype

Opportunistic Context = GCF (Group Context Framework)

Possible Relationships

- Tweet + location (actual or inferred) cues GCF sensors
- Trending topic + similar mission keywords cues ISE sensor (events)
- ISE sensor/event + location cues GCF sensors

Scenarios (objective)

- Geo-tagged tweet triggers GCF sensors for collection
- Trending keyword matches with ISE event description
  - Use ISE to task GCF for additional sensor data
- ISE event triggers GCF sensor collection

Experiment results

- Validated sensor feed architecture from GCF
- GCF sensors accurately locate and track via Bluetooth and location
- GCF sensors may be able to geolocate via audio, analyzing noisy data

Plans for FY16 — Divided Into Two Projects

Tactical Computing and Communications (TCC): Efficient and secure computing and communications for teams operating in tactical environments

- Extension of trusted identities to server clusters and data at rest
- Integration of Delay-Tolerant Networking (DTN) for mobile device to cloudlet to cloud communications
- Refinement of application layer reliable UDP transport
- Architectures for data staging at the tactical edge

Tactical Analytics (TA): Innovative capabilities for data-to-decision in tactical environments

- Use-case refinement and enhanced cueing, tipping, and fusion of sensor data
- Leverage analytics for context (geo-inferencing, name entity recognition, and topic modeling)
- Real-time credibility scoring
- Proactive and transfer learning for recognizing rapidly emerging events
- Identification of storylines in streaming media data

EETS: Assured, Efficient, *Fieldable* Solutions

Impact – a good problem

- Transition: AF/A2I Dewey, TALOS, WMD CST, MIT LL NICS, USMC TID
- Research: Moving forward, multiple refereed publications, multiple collaborations

Reliance 21 alignment: C4I (synthesis/analysis/decision tools; HCII)

Team Members

- W. Anderson, MS, SSD
- J. Boleng, PhD, SSD
- B. Bradshaw, BS, MBA, SSD
- S. Echeverría, MS, SSD
- A. Henderson, MS, SSD
- D. Klinedinst, MS, CERT
- G. Lewis, MSE, PhD(C), SSD
- E. Morris, MS, SSD
- M. Novakouski, MSE, SSD
- K. Pitstick, BS, SSD
- J. Root, BS, SSD
- S. Simanta, MSE, SSD
- K. Williams, BS, SSD

Key Engaged Stakeholders

- J. Carbonell, PhD, LTI, CMU
- A. Dey, PhD, HCII, CMU
- M. Satyanarayanan, PhD, SCS, CMU
- E. Xing, PhD, LTI, CMU

Key Collaborators

- Col F. Deutch, PhD AF/A2I
- Maj. R. Flick, PA Nat. Guard
- A. Miller, IAB S&T Co-chair