

CMU SEI Collaboration

Jeff Boleng

Chief Technology Officer (acting) and Deputy CTO

Software Engineering Institute
Carnegie Mellon University
Pittsburgh, PA 15213

The SEI is a DoD R&D Federally Funded Research and Development Center



Established in 1984 at Carnegie Mellon University

~615 employees (ft + pt), of which about 70% are engaged in technical work

Initiated CERT cybersecurity program in 1988

Offices in Pittsburgh and DC, with several operating locations near customer facilities

About \$150M in funding (~\$20M DoD Line)

DoD Technology Trends

Offset Strategy

“Asymmetrically compensating for a disadvantage”

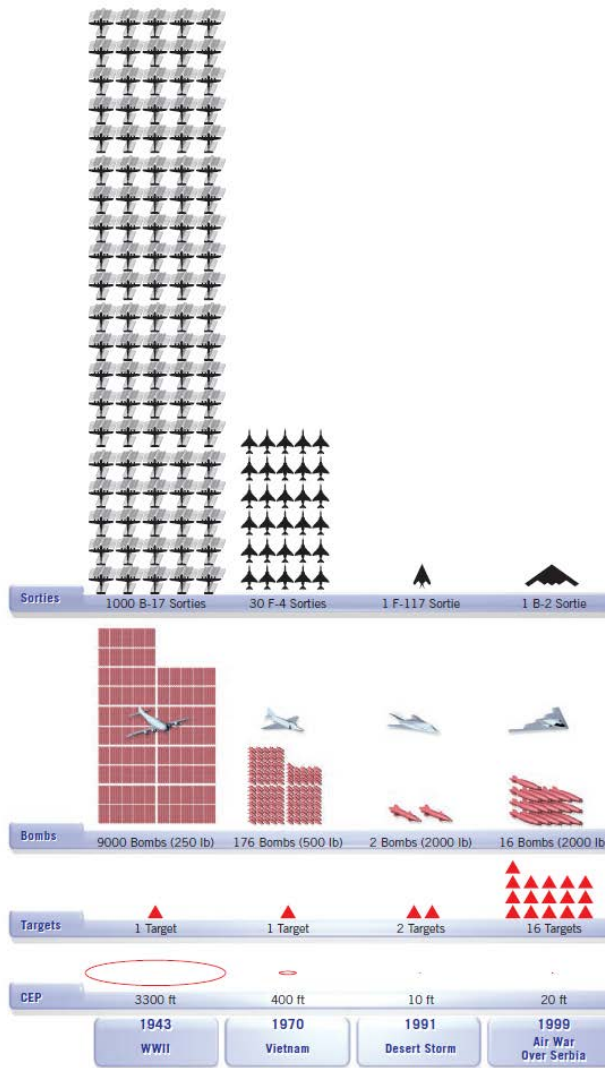
- 1st Offset
 - 1950's – nuclear deterrence to deter Warsaw Pact
 - Avoid large expenditures needed for conventional deterrence
- 2nd Offset
 - 1975-1989 – technology superiority to offset quantitative gap
 - \$100B decline in defense spending, forces outnumbered 3-1
 - Intelligence, Surveillance and Reconnaissance (ISR) platforms
 - Low Observability (LO, aka stealth technology)
 - Precision Guided Munitions (PGM)
 - Night Vision

ISR, LO, & PGM

Technology combined with new strategy

- Effects based operations
- Revolutionary command and control
- Rapid decisive operations

Operating inside the adversary's decision cycle

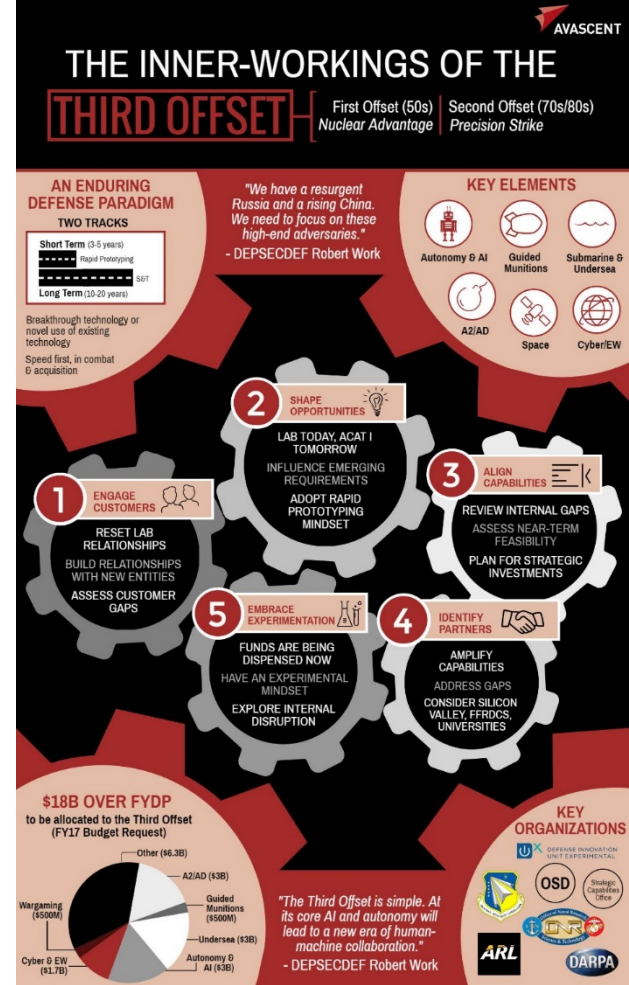


3rd Offset - 2014

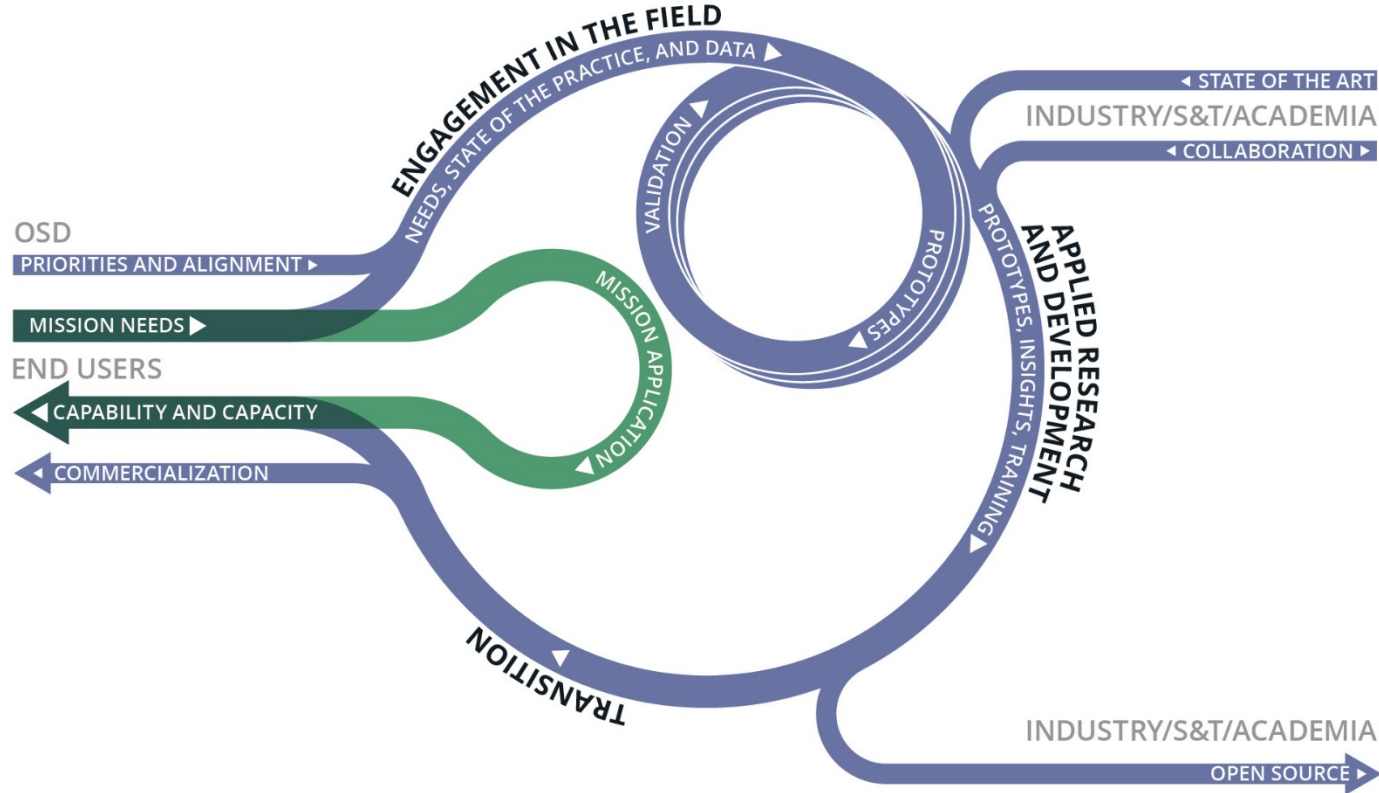
Announced by SecDef Chuck Hagel

- Key elements
 - Hypersonics
 - Autonomy and AI
 - Guided Munitions
 - Undersea Warfare
 - Resilient Space
 - Anti-access/Area Denial (A2/AD)
 - Cyber and Electronic Warfare (EW)
- \$18B over five year defense plan

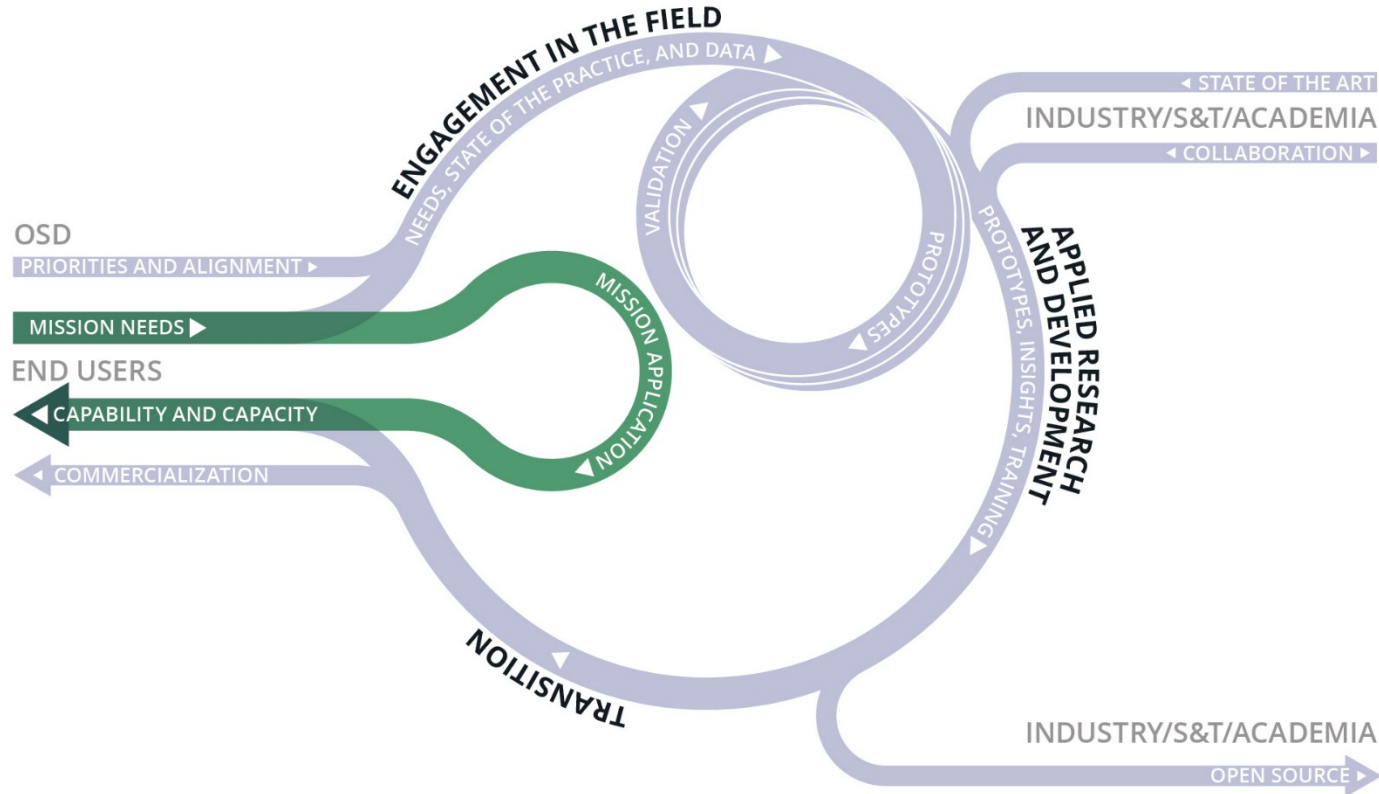
Source: Avascent Infographics, <http://www.avascent.com/2016/06/war-rocks-publishes-josh-pavluk-august-cole-avascent-infographic-third-offset-strategy/>



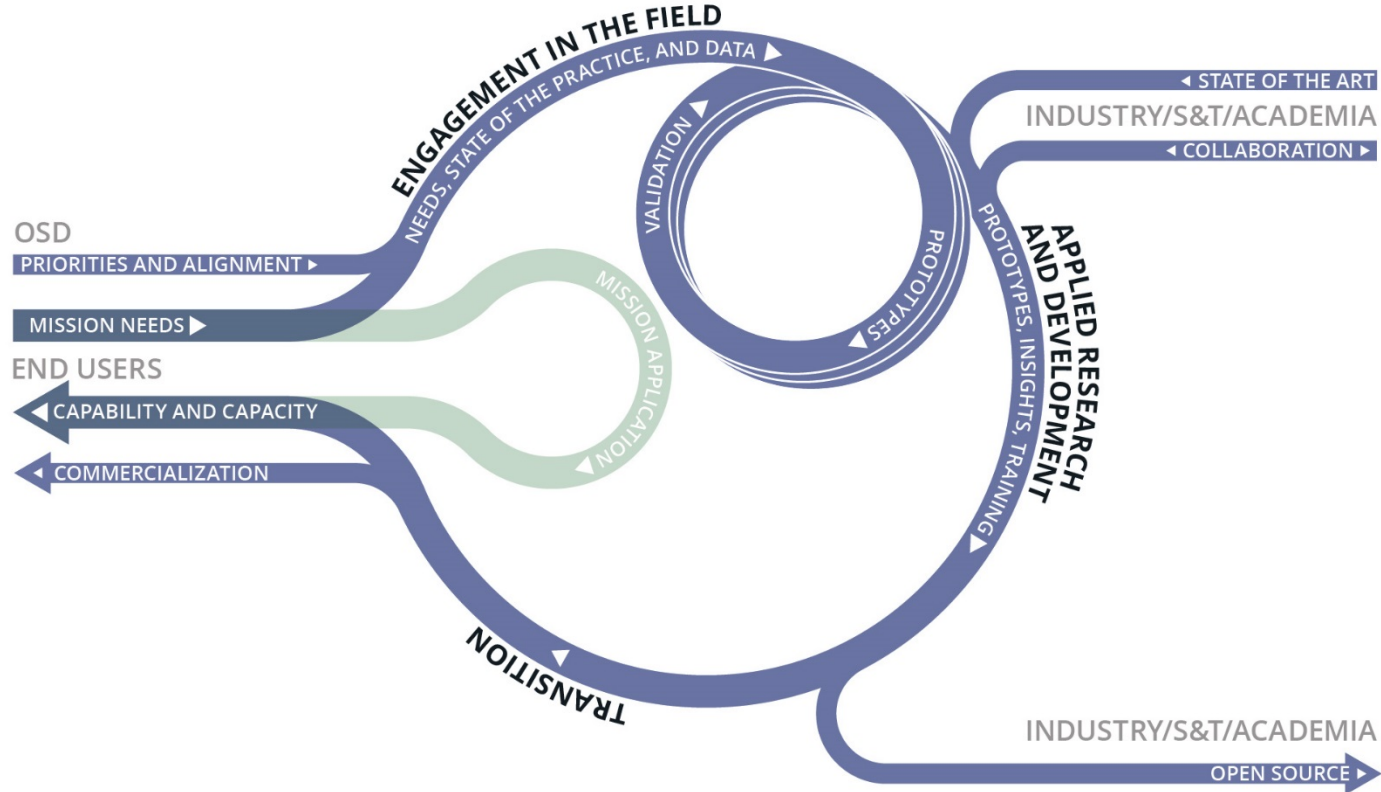
SEI Execution Model



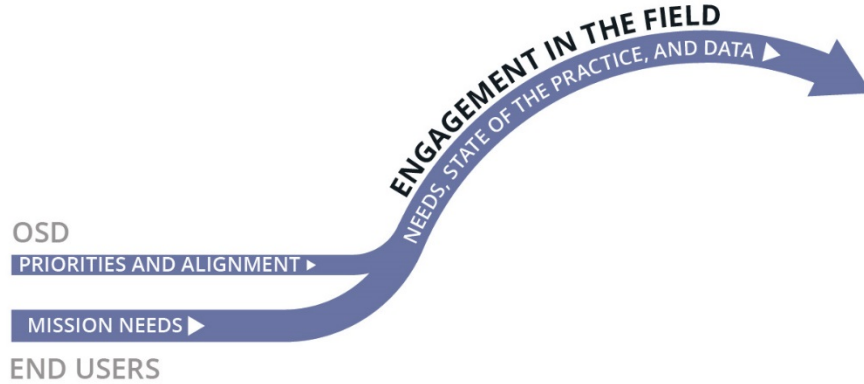
SEI Execution Model



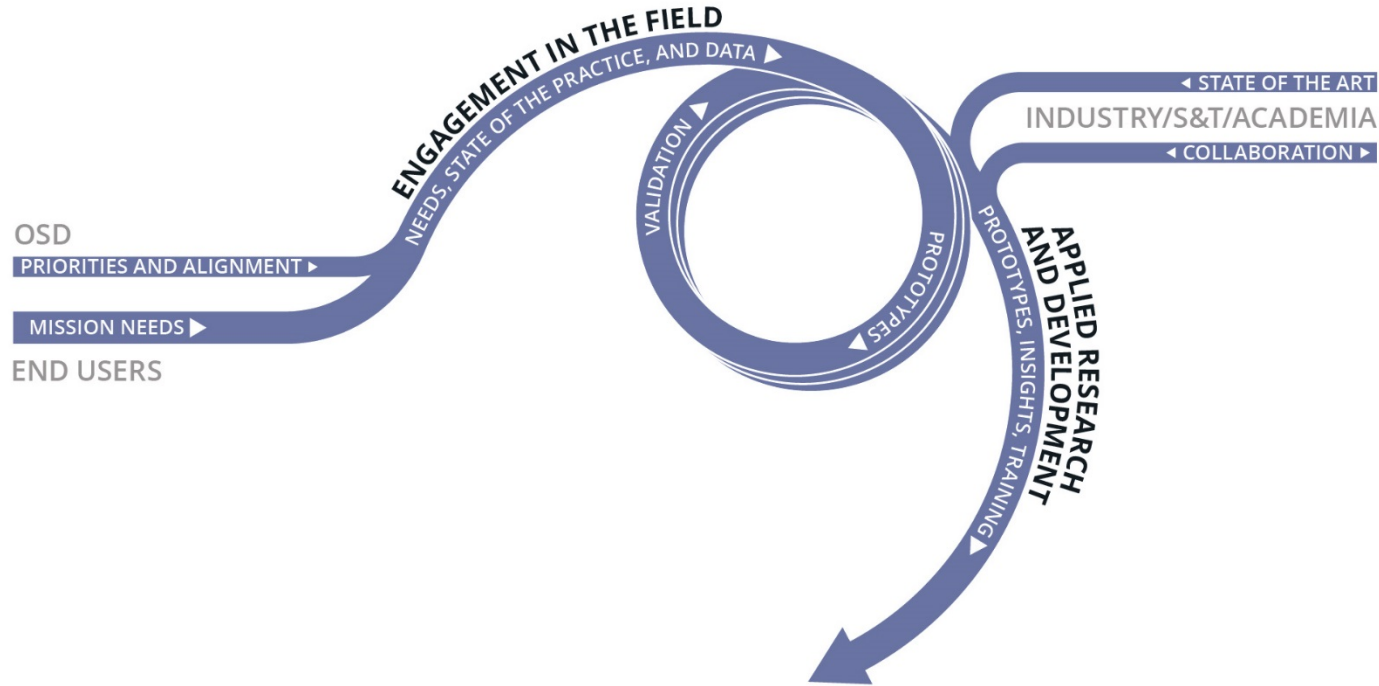
SEI Execution Model



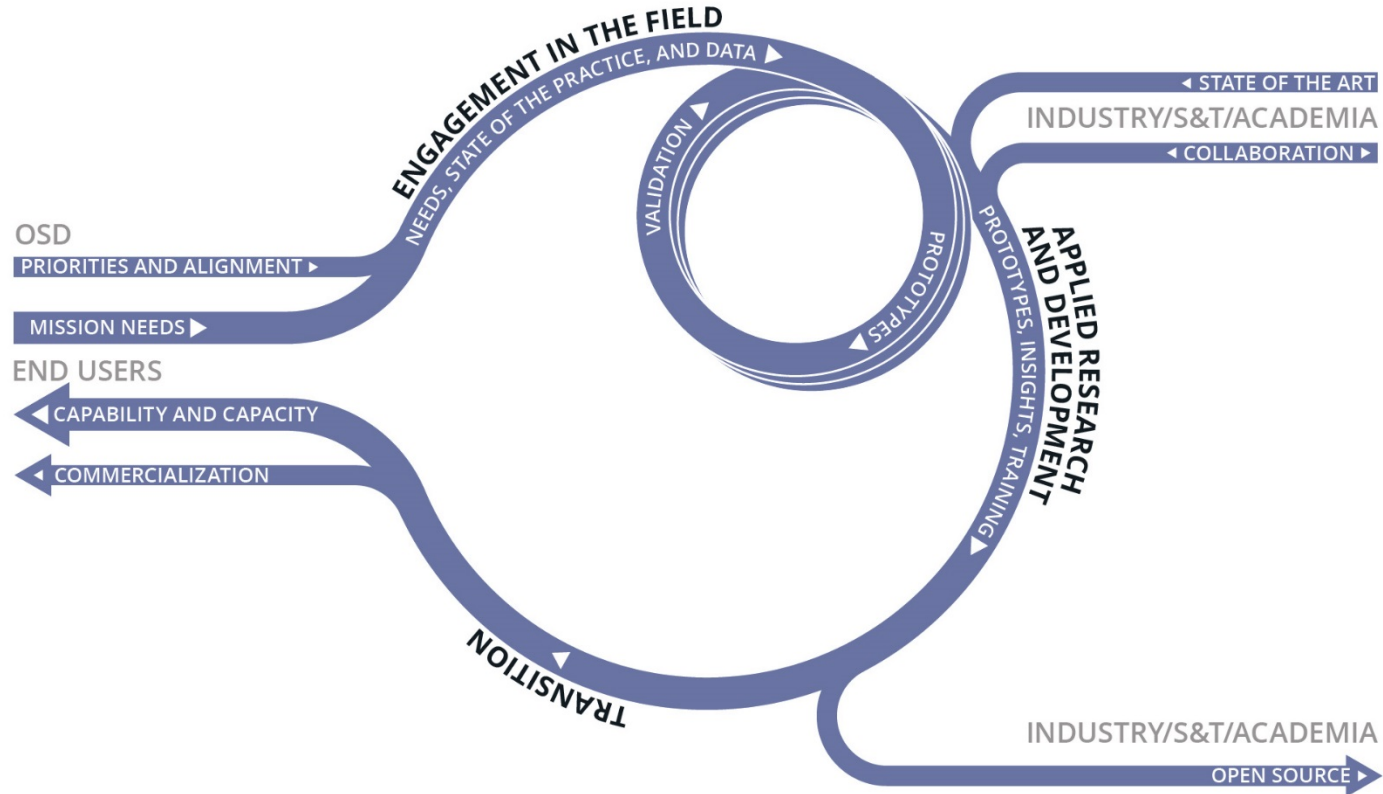
SEI Execution Model



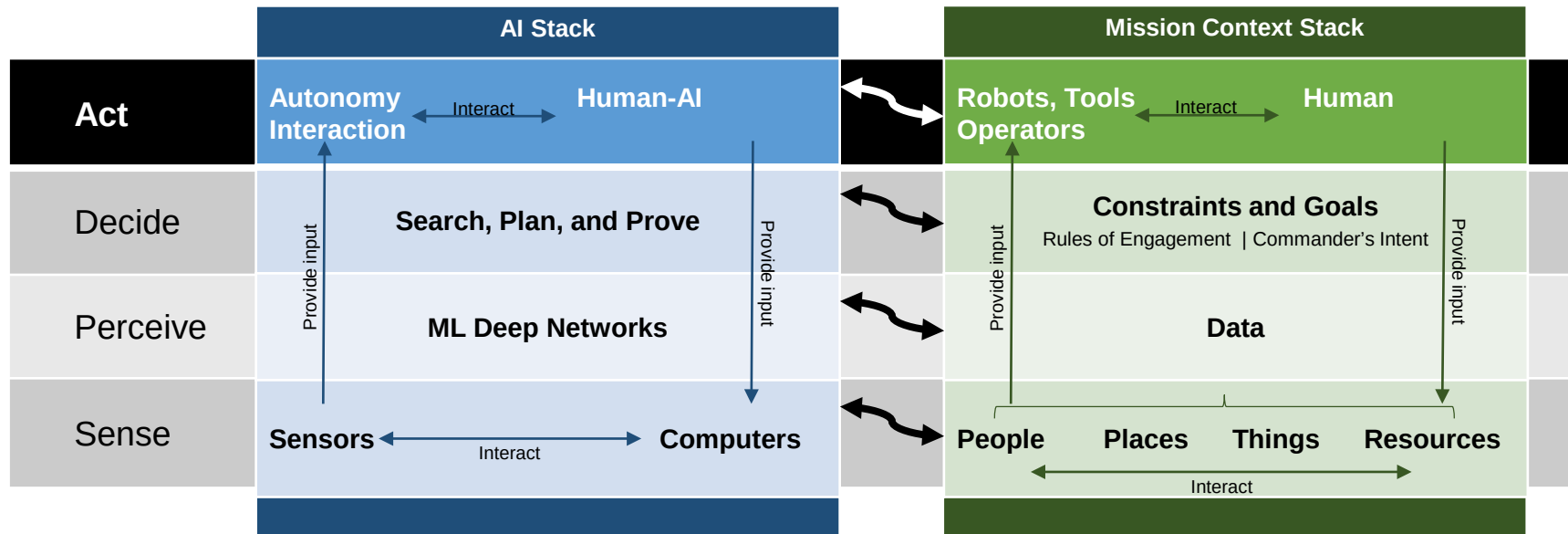
SEI Execution Model



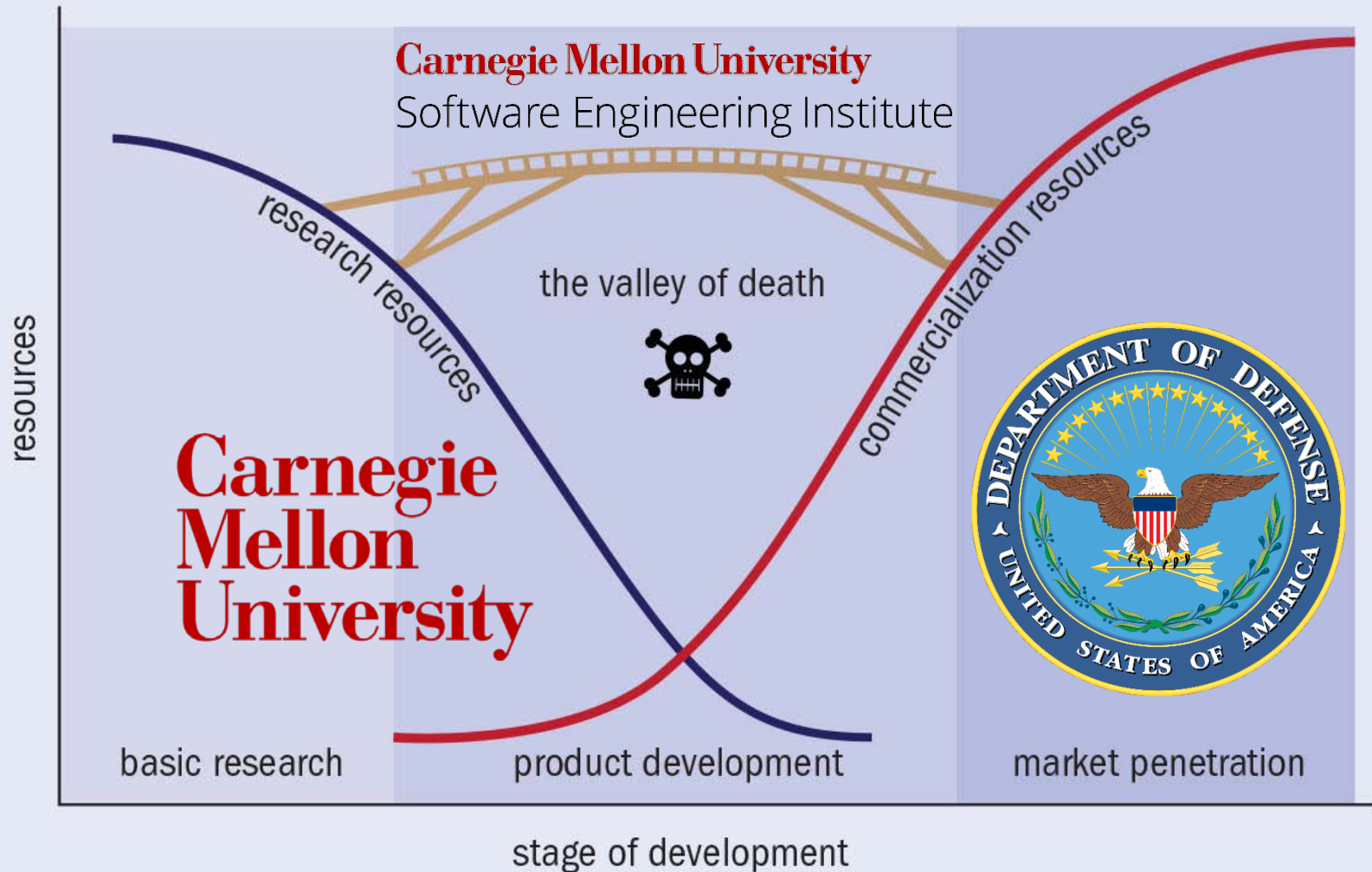
SEI Execution Model



Carnegie Mellon University



Carnegie Mellon University
Software Engineering Institute



SEI Funding

- Line Funding

- Approximately \$20M per year, congressionally allocated
- Line Funded Strategic Initiatives (75%)
 - 1-3 years in length
 - >\$1M per year
 - Intended to develop capability that can be transitioned to DoD
 - Primary output is proof of concept or prototype
- Line Funded Exploratory New Starts (25%)
 - 1 year in length
 - \$350-\$500k
 - Intended to demonstrate that an emerging technology may (or may not) have DoD applicability
 - Can become LSI

- Project Work Plans

- Approximately \$130M per year, customer (DoD, DHS, Gov't Agency, commercial)
- Must be approved by DoD
- Very applied, must be US citizen, set deliverables and timeline

Fundamental Research (FR) Designation

- By default, “Distribution authorized to the Department of Defense and U.S. DoD contractors only. (i.e. Distribution D)”
- Fundamental Research Designation allows “Approved for public release. Distribution is unlimited. (i.e. Distribution A)”
- Most LENS work, and some portion of LSI can be designated as FR
- PWPs can be FR if required by customer and approved by DoD
- FR allows unrestricted collaborators and work location
- Non-FR requires US citizens and work only in SEI space on SEI network
- Approval process by DoD has been onerous, but is getting better

We work to address gaps with R&D in 7 technical areas

Enduring



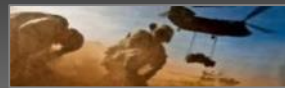
Software Engineering & Information Assurance

Enable high quality, secure software-based systems in a predictable, affordable manner



Cyber Security

Develop improved systems, repeatable practices, and capable personnel to enable cyber missions



System Verification & Validation

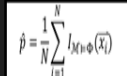
Enhance confidence in the systems engineering lifecycle with evidence-based methods and tools

Make software less costly and more resilient and mission capable by ruthlessly automating all aspects of design, development, integration, testing, deployment, operations, defense, and sustainment of software systems

Emerging



Data Modeling & Analytics: Develop and apply mathematically rigorous data collection, analysis, and visualization techniques



C4ISR Mission Assurance: Enable reliable and predictable mission support by software and systems, which are resilient to adversary actions



Autonomy & Counter-Autonomy: Develop evidence that indicates the trustworthiness, dependencies, & vulnerabilities of autonomous systems



Human-Machine Interactions: Invent, assess, improve comprehensible, safe, and trustworthy technologies for humans to use and team with machines

FY18 Line-funding investment overview by technical area

Enduring technical areas

Primary Technical Area	Project Title	Period	Tech Area \$	Tech Area %
Software Engineering & Information Assurance	Automated Code Repair to Ensure Memory Safety: Source and Binary	FY18-20	\$4.53M	23.2%
	Automated Repair of Incorrect Usage of Security APIs	FY17-18		
	Predicting Security Flaws through Architectural Flaws	FY18		
	Rapid Construction of Accurate Automatic Alert Handling System: Model & Prototype	FY18-19		
	Rapid Software Composition by Assessing Untrusted Components	FY18		
	What to Fix? Automating Technical Debt Analysis through Software Analytics	FY17-18		
Cyber Security	Automated Executable Program Transformation	FY18-20	\$3.30M	16.9%
	Cyber Integration with Kinetic Training	FY18		
	Device Fingerprinting for Commercial Maritime Control Systems	FY18		
	Extracting Key Features of Expert Performance in Cyber Security Tasks	FY18		
	Identifying Vulnerabilities Through Binary Software Composition Analysis	FY18		
	Modeling the Operations of the Vulnerability Ecosystem	FY18		
	Machine Learning Evaluation of Semantic Malware Feature Spaces	FY18		
System Verification & Validation	Certifiable Distributed Runtime Assurance	FY17-18	\$4.10M	21.0%
	Infrastructure as Code	FY18		
	Innovating Air Force Jet Engine System Reliability Test using Machine Learning Integrated with Causal Modeling	FY17-18		
	Integrated Safety and Security Engineering for Mission-Critical Systems	FY18-20		
	Timing Verification of Undocumented Multicore	FY18		

FY18 Line-funding investment overview by technical area

Emerging technical areas

Primary Technical Area	Project Title	Period	Tech Area \$	Tech Area %
Data Modeling & Analytics	An Integrated Causal Model for Software Cost Prediction & Control	FY18-20	\$3.99M	20.4%
	Automated Code Generation for Future-Compatible High-Performance Graph Libraries	FY17-18		
	Building a COTS Benchmark Baseline for Graph Analytics	FY18		
	Summarizing and Searching Video	FY18-20		
C4ISR Mission Assurance	High Assurance Software-Defined IoT Security	FY18-20	\$0.97M	5.0%
Autonomy & Counter-Autonomy	A Series of Unlikely Events	FY18-20	\$1.25M	6.4%
	What will the robot do next?	FY17-18		
Human-Machine Interaction	A Novel Approach to Emotion Recognition from Voice	FY18-19	\$1.40M	7.2%
	Can Deep Learning Predict Security Defects in Synthetic Code?	FY18		

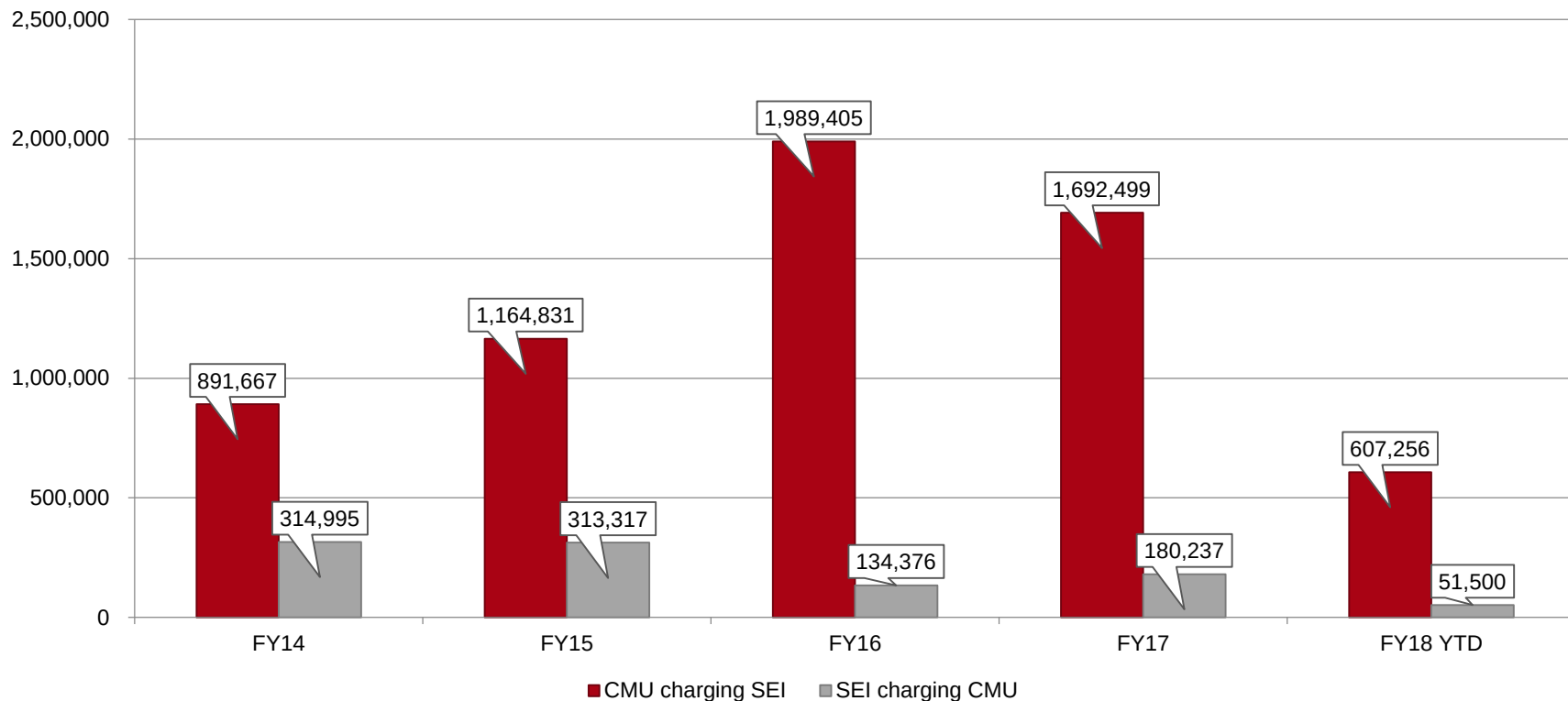
Line-Funding Investment by Technical Area, FY14-18

~60% of funding has been in enduring areas

Investments in new areas have grown in past 2-3 years

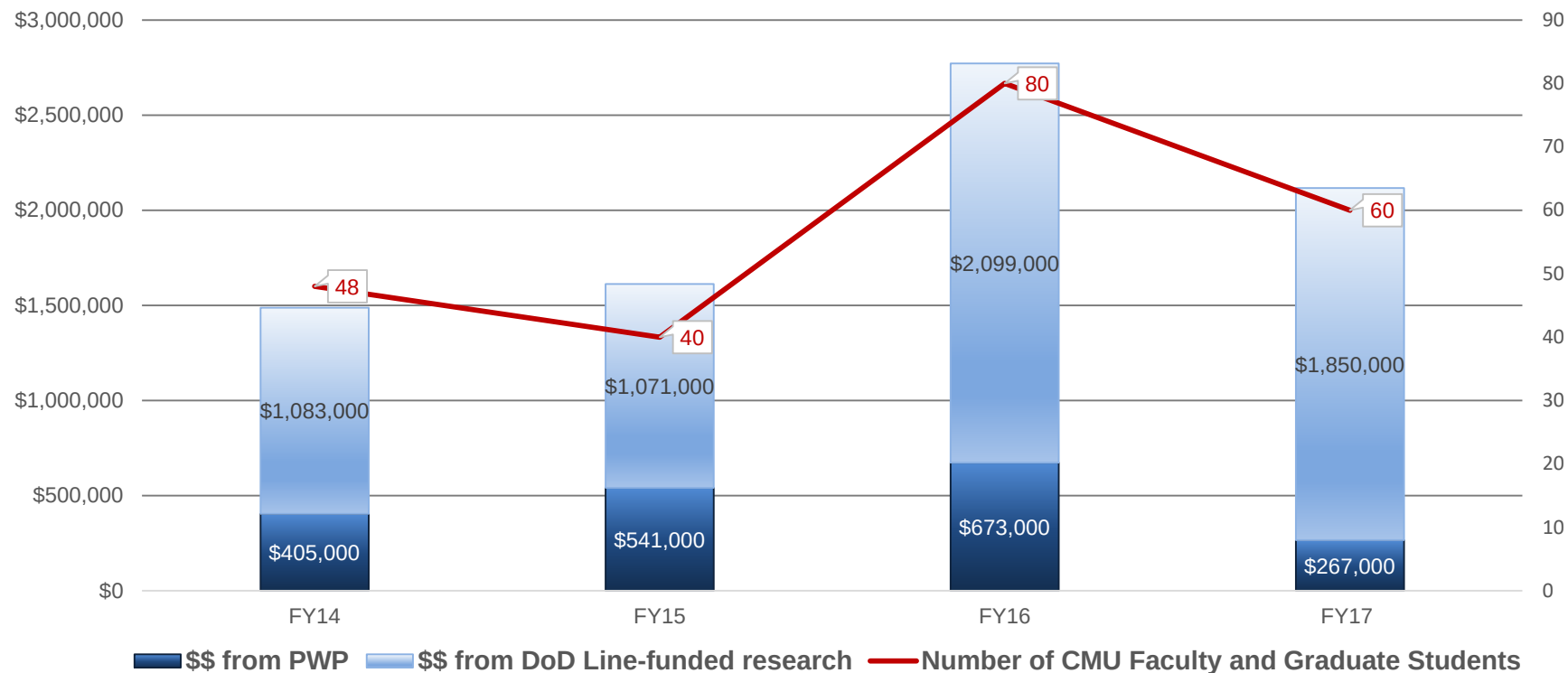


CMU Charging SEI & SEI Charging CMU (CMU FY)



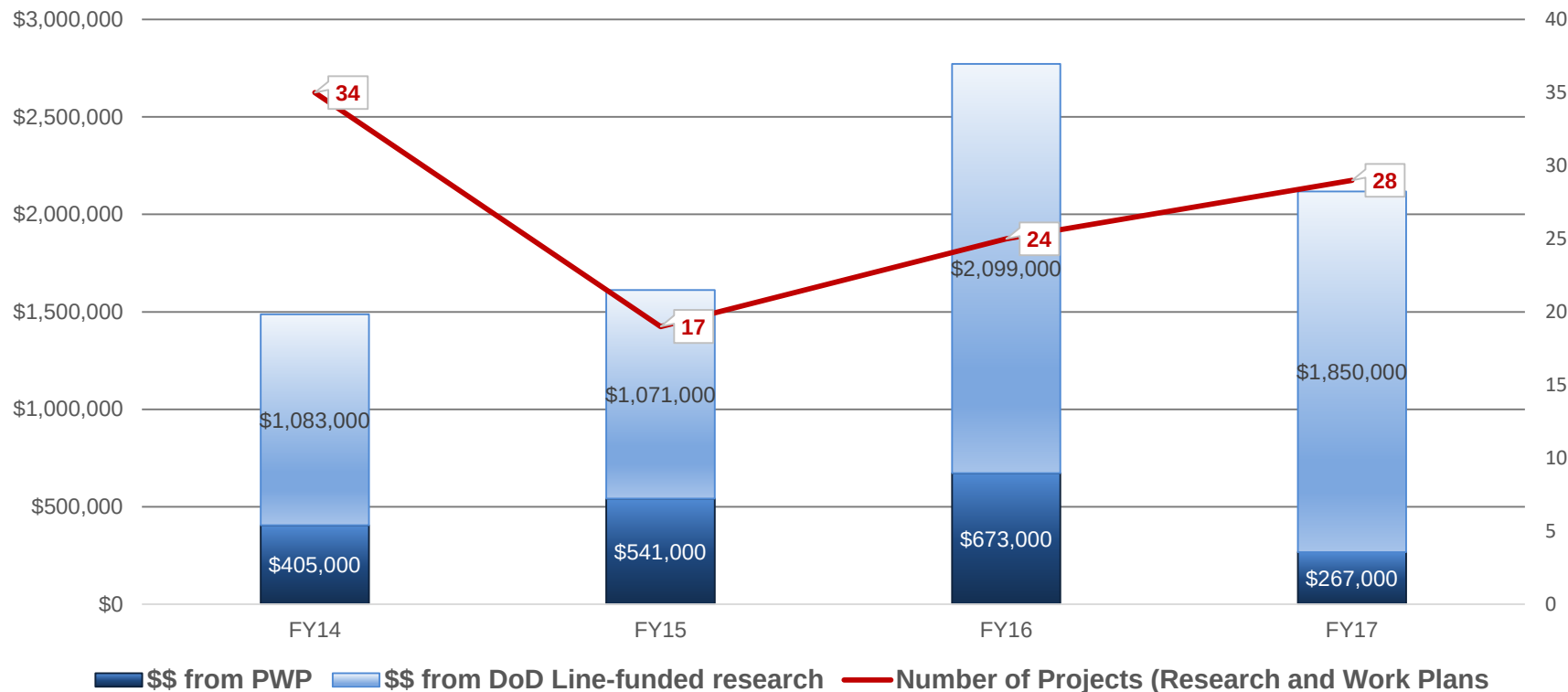
Contribution from SEI Technical Work to CMU (SEI FY)

showing total funding and number of CMU faculty and graduate students



Contribution from SEI Technical Work to CMU (SEI FY)

showing total funding and number of projects



FY19 LSI

Project Name: Untangling the Knot: Automated Component Refactoring Assistant

PI: Ipek Ozkaya (SSD/AP; 70%), James Ivers (SSD/AP; 60%)

Team Members: Robert Nord (SSD/AP; 55%), Jay Marchetti (SSD/CTSD; 40%), Reed Little (SSD/AP; 25%), Rick Kazman (SSD/AP; 25%)

Collaborators: C. Le Goues (CMU ISR), C. Izurieta (Montana State), C. Guthrie (Army)

Project Name: Using all processor cores while being confident about timing

PI: Bjorn Andersson (SSD/CPSULS; 60%), Team Members: Dionisio de Niz (SSD/CPSULS; 25%), Mark Klein (SSD/CPSULS; 20%), Jeffery Hansen (SSD/CPSULS; 15%)

Collaborator(s): CMU/Stat: John Lehoczky; UCR: Hyoseung Kim; CMU/ECE: Ragunathan Rajkumar; AMRDEC: Alex Boydston

Project Name: Using all processor cores while being confident about timing

PI: Bjorn Andersson (SSD/CPSULS; 60%), Team Members: Dionisio de Niz (SSD/CPSULS; 25%), Mark Klein (SSD/CPSULS; 20%), Jeffery Hansen (SSD/CPSULS; 15%)

Collaborator(s): CMU/Stat: John Lehoczky; UCR: Hyoseung Kim; CMU/ECE: Ragunathan Rajkumar; AMRDEC: Alex Boydston

FY19 LENS Proposals

- Proposals in June/July, selection in August, work starts in October
- Encouraging campus collaboration
- Focus areas (Better, Faster, Cheaper, more Secure Software)
 - Development tool chain automation
 - Increased formal rigor and expressiveness
 - Automated generation of lifecycle phase artifacts
 - Automated verification of lifecycle phase artifacts
 - Automated integration of artifacts across lifecycle phases (up and down stream)
 - Autonomous cyber operations
 - Align cyber defense around mission
 - Prepare the cyber force
 - Hardware software co-design
 - Human machine teaming
 - Advanced uses of AI and ML

FY14 Projects with CMU Collaborators

SEI Project Name	Faculty
Acquisition Dynamics	Gonzalez
Automated Cyber Readiness Evaluator	Sheikh, Gordon
Autonomous Vehicle Health Monitoring (FY13-14)	Scerri
Avatar	Rudnický
Concurrent Crowdsourcing of Requirements and Architectures for Socio-Technical Eco-system (STE) Infrastructure Improvement	
Context-Aware Network Analysis in Resource-Constrained Environments	Camara Moreno, Schmerl
Contract-Based Virtual Integration of CPS Analyses	Garlan, Schmerl
Cyber Security Expert Performance and Measurement	Young
Deep Focus: Increasing User Depth of Field to Improve Threat Detection	Maxion
Developing Coordinated Multi-UGV Reliability Analysis Techniques	
Edge-Enabled Tactical Systems (FY13-14)	Satyanarayanan
Efficacy of Agile Methods (FY13-14)	
Enhanced Vulnerability Discovery and Exploitation (FY13-14)	Brumley
ETC NSA R2	Camara Moreno, Garlan, Schmerl
Extracting Insider Threat Indicators from Large Data Sets	
High-Confidence Cyber-Physical Systems (FY13-14)	

SEI Project Name	Faculty
Improving Synthetic Data for Cyber Security (FY13-14)	Rudnický
Insider Threat Mitigation	Carley
Malware Analysis	
Prioritizing Malware Analysis	
Real-Time Mobile Applications in Intermittently Connected Networks	Seshan
Secure and Assured Mobile Computing Components	
Secure Coding	Bauer, Jia
Self-governing Mobile Adhocs with Sensors and Handhelds	
Software Assurance Engineering	
Software Model Checking for Verifying Distributed Algorithms	
Verifying Evolving Software	
5-308J DHS S&T CMAS	
5-356H3 DARPA	
5-444C1 and C2 DIA	Frederking, Gershman
5-444E DIA	
5-555B2 CSEP CICPA	
5-555E3 CyberSec Eval PM	Maxion
5-555E4 SRAM ProgSup	Maxion

FY14: Highlights

Edge-Enabled Tactical Systems

- **Total: \$102K**
- Algorithms for group autonomy of mobile systems
- Prototypes for information superiority at the edge technologies, analyzing data streams (e.g., social media), and tactical cloudlets

Automated Cyber-Readiness Evaluator

- **Total: \$172K**
- System to convert computer screen activities into human readable format
- Intelligent “tutor” to parse output from system into objective measures of skills for a defined skill set

Contract-Based Virtual Integration and Cyber-Physical System Analysis

- **Total: \$96K**
- Algorithm to discover conflicts between analyses based on their contracts
- Implementation in Open Source AADL Tool Environment

Insider Threat

- **3 projects, Total:\$140K**
- Develop insider threat mitigations that form an architectural foundation for next-generation DoD enterprise systems



FY15 Projects with CMU Collaborators

SEI Project Name	Faculty
Acquisition Dynamics (FY14-15)	Gonzalez
API Usability and Security	Myers, Sunshine
Automated Cyber Readiness Evaluator/Generalized Automated Cyber Readiness Evaluation (FY14-15)	Gencaga, Sheikh, Gordon
Edge-Enabled Tactical Systems (FY13-15)	
Extending AADL for Security Design Assurance of the Internet of Things (IoT)	Rosso-Llopart
Human-Computer Decision Systems for Cyber Security	
Insider Threat Mitigation (FY14-15)	
Line Plus up CSC AFSOR	Bass
Machine Learning to Support Big Data System Acquisition	Yang

SEI Project Name	Faculty
Malware Analysis (FY14-15)	
Secure Coding (FY14-15)	
Verifying Evolving Software	
5-356H3 DARPA	
5-444C1 and C2 DIA	Frederking, Gershman
5-555E4 and 6-555E1 SRAM ProgSup	Maxion
5-577A1 SOCOM TALOS	Atkeson
5-579A1 NTIA	Breaux, Gralan, Le Goues, Scherlis, Scmerl, Sicker

FY15: Highlights

Generalized Automated Cyber-Readiness Evaluation

- **Total: \$410K (and \$422K in FY15)**
- Extends Automated Cyber-Readiness Evaluator for additional job roles
- Builds automated evaluations platform
- Identifies new criteria for cyber operational mission readiness

Technical Assessment: Frequency Records Management System for the Department of Commerce (PWP 5-579A1)

- **Total: \$206K**
- Verify SOA implementation
- Analyze source code quality
- Analyze software architecture for desired quality attributes

Machine Learning to Support Big Data Acquisition

- **Total: \$165K**
- Algorithm for building acquisition decision knowledge bases
- Automatically extensible knowledge base for acquiring Big Data systems

Critical Infrastructure Resilience for DHS (PWP 5-555E4/6-555E1)

- **Total: \$208K**
- Provide stakeholder risk assessment and mitigation
- Perform data analytics
- Update and maintain the supply chain risk management toolset
- Provide threat and resilience analysis

FY16 Projects with CMU Collaborators

SEI Project Name	Faculty
Automated Code Repair	Kastner, Le Goues
Big Learning Benchmark	Gibson, Xing
Critical Infrastructure Assurance Through Electronic Components Attribution	
Data Validation for Large-Scale Analytics	
Edge-Enabled Tactical Systems (FY14-16)	
Effecting Large-scale Adaptive Swarms Through Intelligence Collaboration (ELASTIC Autonomy)	
Generalized Automated Cyber-Readiness Evaluation (FY15-16)	Gencaga, Sheikh, Gordon
Generalizing Supervised Latent Dirichlet Allocation (LDA) for Analyzing Open Source Data	Xing
Human-Computer Decision Systems for Cyber Security (FY15-16)	
Insider Threat Mitigation (FY14-16)	
Machine Learning to Support Big Data System Acquisition (FY15-16)	

SEI Project Name	Faculty
Multi-Agent Decentralized Planning for Adversarial Robotic Team	
Quantifying Uncertainty for Early Lifecycle Cost Estimation (QUELCE)	Nyberg
Real-time Extraction of Biometric Traits from Video	Kitani, Savvides
Structural Multi-Task Transfer Learning for Improved Situational Awareness	Carbonell
Supporting Software Engineering Best Practices in Additive Manufacturing	Hudson, Mankoff
Tactical Analysis	Carbonell
Verifying Distributed Adaptive Real-Time (DART) Systems	Muelliling
Why did the robot do that?	
5-308N1 and 6-308N1 DHS S&T NON DOD	Garlan, Schmerl
5-577A1 SOCOM TALOS	Atkeson, Choset
5-579A1 NTIA	
5-579A2 SSD NTIA	Le Goues
6-555E1 7.1 SRAM	Maxion

FY16: Highlights

Human-Computer Decision Systems for Cyber Security

- **Total: \$124K**
- Investigated methods using both supervised (active) and unsupervised learning to augment the abilities of analysts

Why did the Robot do That?

- **Total: \$207K**

Project aims to build trust of robots by human teammates



- Explainable AI: algorithms to explain robot behavior automatically and incorporate explanations into robot user interfaces

Data Analytics for Situational Awareness

- **2 projects, Total: \$232,000**
- Built prototypes to demonstrate new capabilities for script learning
- Delivered ML techniques that enable analysts to understand emerging situations quickly

Reference Implementation for Unified Moving Target Defense (MTD)

(PWP 5-308N1/6-308N1)

- **Total: \$323K**
- Developed a fully IT-managed MTD platform
- Performed research to further software architecture concepts for self-adaptive systems

FY17 Projects with CMU Collaborators

SEI Project Name	Faculty
Automated Code Generation for Future-Compatible High-Performance Graph Libraries	Franchetti, Low
Automated Code Repair (FY16-17)	
Big Learning Benchmark (FY16-17)	Gibson
Certifiable Distributed Runtime Assurance (CDRA)	Datta
Critical Infrastructure Assurance Through Electronic Components Attribution	
Data Validation for Large-Scale Analytics (FY16-17)	
Developing static analysis techniques for blockchain application development	Aldrich, Sunshine
Enhanced Vulnerability Discovery and Exploitation	
Events, Relationships, and Script Learning for Situational Awareness	Carbonell
Generalized Automated Cyber-Readiness Evaluation (FY15-17)	Sheikh
Generalizing Supervised Latent Dirichlet Allocation (LDA) for Analyzing Open Source Data	
Human-Computer Decision Systems for Cyber Security (FY15-17)	
Innovating Air Force Jet Engine System Reliability Test using Machine Learning integrated with Causal Modeling	Danks
Micro-expressions - More than meets the eye	Savvides

SEI Project Name	Faculty
Modeling the decision making of expert incident coordinators/responders for use by non-experts	
Multi-Agent Decentralized Planning for Adversarial Robotic Team	
Ongoing, Dynamic Design Analysis of Configurable Systems	Garlan, Schmerl
Real-time Extraction of Biometric Traits from Video (FY16-17)	Savvides
Structural Multi-Task Transfer Learning for Improved Situational Awareness (FY16-17)	Carbonell
Summarizing and Learning Latent Structure in Video	Liang
Supporting Software Engineering Best Practices in Additive Manufacturing (FY16-17)	
Verifying DART Systems (FY15-17)	Oh
What will the robot do next?	
Why did the robot do that? (FY16-17)	
Why does software cost so much? Towards a Causal Model	Danks, Zhang
5-308N1 and 6-308N1 DHS S&T NON DOD	
6-555E1 7.1 SRAM	Maxion
6-599B2 JIDA	Savvides

FY17: Highlights

Micro-Expressions (ME): More than Meets the Eye

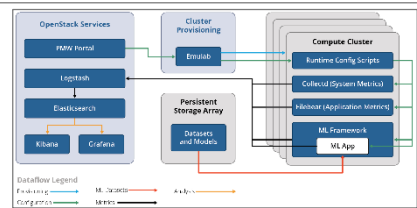
- **Total: \$197K**
- ME expose genuine emotions
- Building a prototype that outperforms humans in spotting and recognizing facial micro-expressions in near real time

Automated Code Generation for Future-Compatible High-Performance Graph Libraries

- **Total: \$323K**
- **Also, funded \$250K supercomputer, still in use by team**
- advances state-of-the-art in code
- generation by introducing concepts to capture graph algorithm primitives

Big Learning Benchmark

- **Total: \$397K**
- Building a workbench to measure and report performance of large-scale ML platforms designed to operate on Big Data workloads



Video Search Web Application for the Joint Improvised Threat-Defeat Agency (PWP 6-599B2)

- **Total: \$175K**
- Engage CyLab Biometrics Center in achieving best classification accuracy on types of videos of interest
- Develop and test a video search web application

FY18 Projects with CMU Collaborators

SEI Project Name	Faculty
Automated Code Generation for Future-Compatible High-Performance Graph Libraries (FY17-18)	Franchetti, Low
Automated Code Repair to Ensure Memory Safety for Source and Binary	Le Goues, Claire Jia, Limin
Building a COTS Benchmark Baseline for Graph Analytics	Tze Meng Low Franz Franchetti James Hoe
Certifiable Distributed Runtime Assurance (CDRA) (FY17-18)	Datta
High-Assurance Software-Defined IoT Security	Sekar, Vyas Vasudevan, Amit
Innovating Air Force Jet Engine System Reliability Test using Machine Learning integrated with Causal Modeling (FY17-18)	Danks

SEI Project Name	Faculty
Integrated Causal Model for Software Cost Prediction & Control	Danks
A Novel Approach to Emotion Recognition from Voice	Singh ; Stern
Rapid Construction of Accurate Automatic Alert Handling System: Model & Prototype	Le Goues, Claire
A Series of Unlikely Events (using Inverse Reinforcement Learning)	Dey, Anind
Summarizing and Searching Video	Xing; Liang
Timing Verification of Undocumented Multicore	Lehoczky

Copyright 2018 Carnegie Mellon University. All Rights Reserved.

This material is based upon work funded and supported by the Department of Defense under Contract No. FA8702-15-D-0002 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center.

The view, opinions, and/or findings contained in this material are those of the author(s) and should not be construed as an official Government position, policy, or decision, unless designated by other documentation.

NO WARRANTY. THIS CARNEGIE MELLON UNIVERSITY AND SOFTWARE ENGINEERING INSTITUTE MATERIAL IS FURNISHED ON AN "AS-IS" BASIS. CARNEGIE MELLON UNIVERSITY MAKES NO WARRANTIES OF ANY KIND, EITHER EXPRESSED OR IMPLIED, AS TO ANY MATTER INCLUDING, BUT NOT LIMITED TO, WARRANTY OF FITNESS FOR PURPOSE OR MERCHANTABILITY, EXCLUSIVITY, OR RESULTS OBTAINED FROM USE OF THE MATERIAL. CARNEGIE MELLON UNIVERSITY DOES NOT MAKE ANY WARRANTY OF ANY KIND WITH RESPECT TO FREEDOM FROM PATENT, TRADEMARK, OR COPYRIGHT INFRINGEMENT.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution. Please see Copyright notice for non-US Government use and distribution.

This material may be reproduced in its entirety, without modification, and freely distributed in written or electronic form without requesting formal permission. Permission is required for any other use. Requests for permission should be directed to the Software Engineering Institute at permission@sei.cmu.edu.

DM18-0457