

Headquarters U.S. Air Force

Autonomous Horizons

System Autonomy in the Air Force



Dr. Greg Zacharias Air Force Chief Scientist (AF/ST)

Integrity - Service - Excellence







- Background and context
- Challenges to overcome
- Approaches to solutions
- Next steps

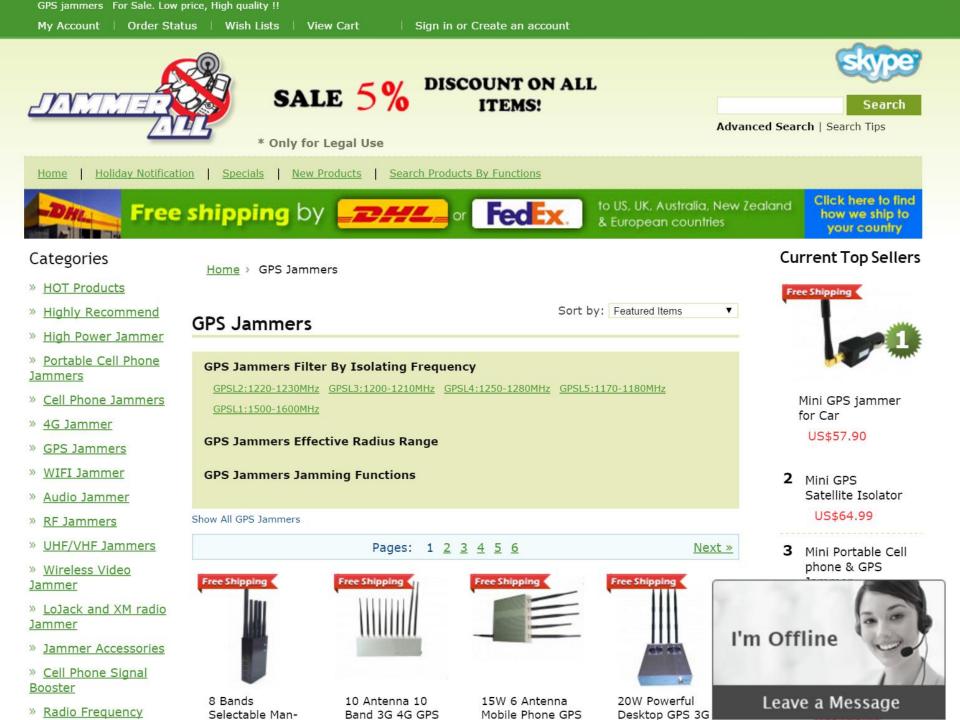






Background and context

- Challenges to overcome
- Approaches to solutions
- Next steps





Autonomy Could Transform Many Air Force Missions





Remotely Piloted Vehicles







Cyber Operations



C2&ISR



Air Traffic Control



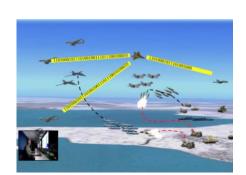
DSB 2012 Autonomy Study: Recommendations



- The Assistant Secretary of Defense for Research and Engineering (ASD(R&E)) should work with the Military Services to establish a coordinated S&T program with emphasis on:
 - Natural user interfaces and trusted human-system collaboration
 - Perception and situational awareness to operate in a complex battle space
 - Large-scale teaming of manned and unmanned systems
 - Test and evaluation of autonomous systems
- These emphasis areas have driven DoD's Autonomy Community of Interest Tier I Technology Areas*:



Human/Autonomous System Interaction and Collaboration (HASIC)



Machine Perception, Reasoning and Intelligence (MPRI)



Scalable Teaming of Autonomous Systems (STAS)



Test, Evaluation, Validation, and Verification (TEVV)

*Dr. Jon Bornstein, "DoD Autonomy Roadmap: Autonomy Community of Interest", NDIA 16th Annual Science & Engineering Technology Conference, Mar 2015.



DSB 2015 Autonomy Study: Terms of Reference



- The study will ask questions such as:
 - What activities cannot today be performed autonomously? When is human intervention required?
 - What limits the use of autonomy? How might we overcome those limits and expand the use of autonomy in the near-term as well as over the next two decades?
- The study will also consider:
 - Applications to include:
 - Decision aids, planning systems, logistics, surveillance, and war-fighting capabilities
 - The international landscape, identifying key players (both commercial and government), relevant applications, and investment trends
 - Opportunities such as:
 - Use of large numbers of simple, low cost (ie, "disposable") objects
 - Use of "downloadable' functionality (e.g., apps) to repurpose basic platforms
 - Varying levels of autonomy for specific missions rather than developing missionspecific platforms

The study will deliver a plan that identifies barriers to operationalizing autonomy and ways to reduce or eliminate those barriers



DSB 2015 Autonomy Study: Status



- Still awaiting release of the Report
- But we can infer some conclusions from DepSecDef (Mr. Work) from his comments last December's CNAS Inaugural National Security Forum





- Autonomous deep learning systems
 - Coherence out of chaos: Analyzes overhead constellation data to queue human analysts (National Geospatial Agency)
- Human-machine collaboration
 - F-35 helmet portrayal of 360 degrees on heads up display
- Assisted human operations
 - Wearable electronics, heads-up displays, exoskeletons
- Human-machine combat teaming
 - Army's Apache and Gray Eagle UAV, and Navy's P-8 aircraft and Triton UAV
- Network-enabled semi-autonomous weapons
 - Air Force's Small Diameter Bomb (SDB)



A Spectrum of Autonomous Solutions*

Autono

M



Assisted/enhanced human pe

- Wearable electronics, heads-
- 711th HPW enhanced sensory, architecture

Human-machine collaboration

- Humans teaming with autono
- Cyborg Chess; Pilot's Associa

Human-machine collaboration

- Humans teaming with autonc
- AFSOC Tactical Off-board Se Technology Demonstration (*i*

Autonomous "deep learning"

- Autonomous systems that lea emergent behavior, ...
- AFRL's Autonomous Defensiv
- Cyber-secure and EW-harde
 - AF's Small Diameter Bomb (Sk)

s, exoskeletons motor

n-aiding)

ems elmet teaming) forms anced



711th Human Performance Wing BATMAN project



Altius UAV Demo

he and "big data"; tactical learning,

berations (ADCO) onomous weapons 'S-denied operation



Need Effective Synergy of the Human/Autonomy Team



- Main benefits of autonomous capabilities are to extend and complement human performance, not necessarily provide a direct replacement of humans
 - Extend human reach (e.g., operate in more risky areas)
 - Operate more quickly (e.g., react to cyber attacks)
 - Permit delegation of functions and manpower reduction (e.g., information fusion, intelligent information flow, assistance in planning/replanning)
 - Provide operations with denied or degraded comms links
 - Expand into new types of operations (e.g., swarms)
 - Synchronize activities of platforms, software, and operators over wider scopes and ranges (e.g., manned-unmanned aircraft teaming)

Synergistic human/autonomy teaming is

critical to success

- Coordination and collaboration on functions
- Overseeing what each is doing and intervening when needed
- Reacting to truly novel situations









Background and context

Challenges to overcome

Approaches to solutions

Next steps

AUTONOMOUS HORIZONS

System Autonomy in the Air Force - A Path to the Future

Volume I: Human-Autonomy Teaming



United States Air Force Office of the Chief Scientist

> AF/ST TR 15-01 June 2015

Distribution A. Approved for public release; distribution is unlimited. Public Release Case No 2015-0267





Traditional approaches to automation lead to "out-of-theloop" errors (low mission SA)

- Loss of situation awareness
 - Vigilance and complacency, changes in information feedback, active vs. passive processing
- Slow to detect problems and slow to diagnose
- Previous systems have led to poor understanding of the system's behavior and actions (low system SA)
 - System complexity, interface design, training
 - Raft of "mode awareness" incidents in commercial aviation after flight management systems (FMS) introduced
- Can actually increase operator workload and/or time required for decision-making
- Trust and its impact on over- and under-usage



Does Automation Reduce Workload?



- Automation of least use when workload highest (Bainbridge, 1983)
- Pilots report workload same or higher in critical phases of flight (Wiener, 1985)
- Initiation of automation when workload is high increases workload (Harris, et al, 1994; Parasuraman, et al, 1994)
- Elective use of automation not related to workload level of task (Riley, 1994)
- Subjective workload high under monitoring conditions (Warm, et al, 1994)





- Autonomous decisions can lead to high-regret actions, especially in uncertain environments → Trust is critical if these systems are to be used
 - Current commercial applications tend to be in mostly benign environments, accomplishing well understood, safe, and repetitive tasks. Risk is low.
 - Some DoD activity, such as force application, will occur in complex, unpredictable, and contested environments. Risk is high.
- Barriers to trust in autonomy include those normally associated with human-human trust, such as low levels of:
 - Competence, dependability, integrity, predictability, timeliness, and uncertainty reduction
- But there are additional barriers associated with human-machine trust:
 - Lack of analogical "thinking" by the machine (e.g., neural networks)
 - Low transparency and traceability; system can't explain its own decisions
 - Lack of self-awareness by the system (system health), or environmental awareness
 - Low mutual understanding of common goals, working as teammates
 - Non-natural language interfaces (verbal, facial expressions, body language, ...)







Background and context

Challenges to overcome

Approaches to solutions

Next steps

AUTONOMOUS HORIZONS

System Autonomy in the Air Force - A Path to the Future

Volume I: Human-Autonomy Teaming



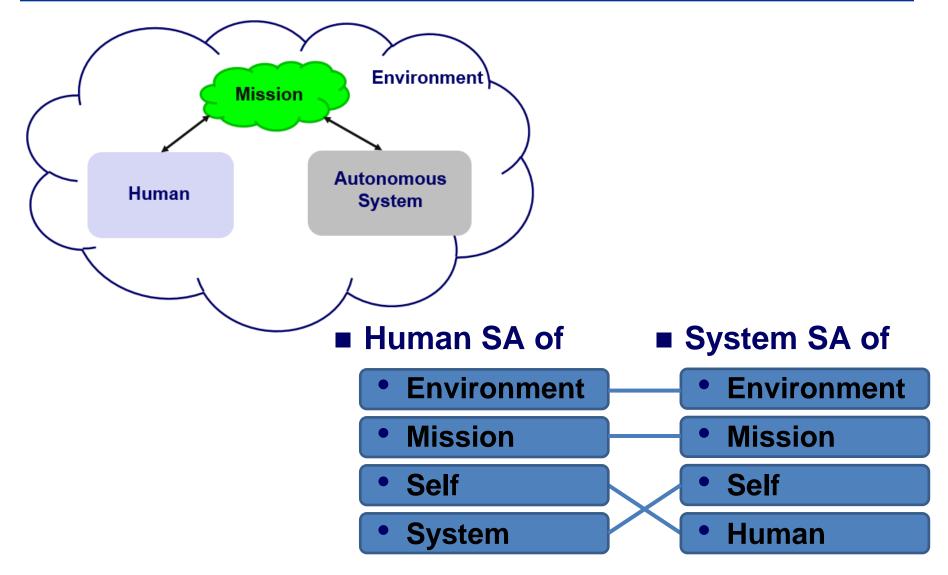
United States Air Force Office of the Chief Scientist

> AF/ST TR 15-01 June 2015

Distribution A. Approved for public release; distribution is unlimited. Public Release Case No 2015-0267



SA is Critical to Autonomy Oversight and Interaction

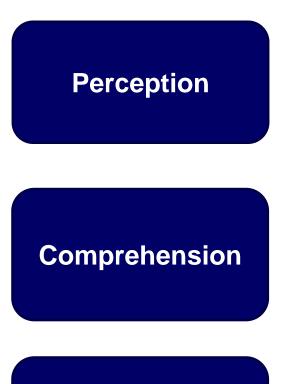






Human

- Data validity
- Automation Status
- Task Assignments
- Task Status
- Current Goals
- Impact of Tasks on Autonomy Tasks
- Impact of Tasks on System/Environment
- Impact of Tasks on Goals
- Ability to Perform Assigned Tasks
- Strategies/Plans
- Projected actions



Projection

Autonomy

- Data validity
- Human Status
- Task Assignments
- Task Status
- Current Goals
- Impact of Tasks on Human Tasks
- Impact of Tasks on System/Environment
- Impact of Tasks on Goals
- Ability to Perform Assigned Tasks
- Strategies/Plans
- Projected actions

Reducing Workload and Reaction Time, and Improving Performance

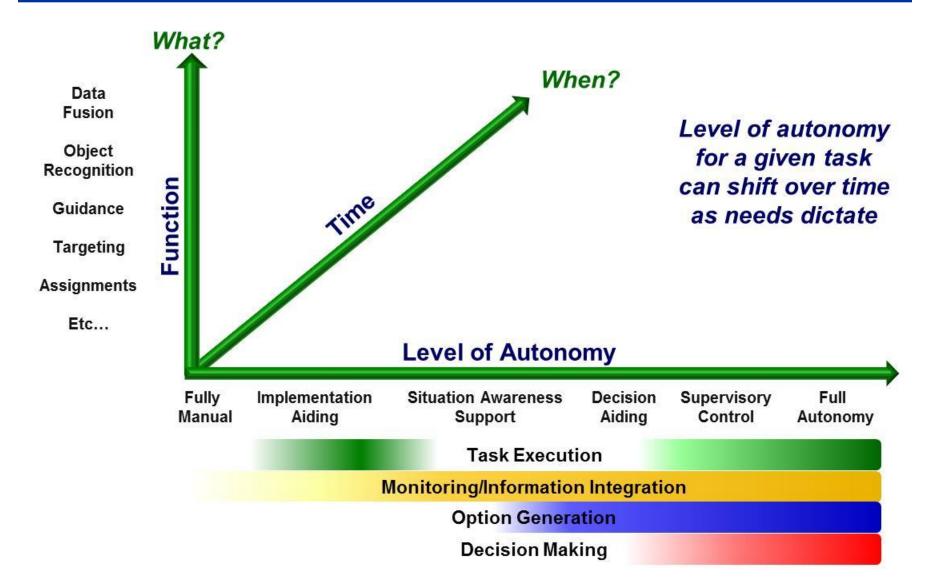


- Supervised, flexible autonomy
 - Human in ultimate control: Can oversee, modify behavior as needed
 - Autonomy levels available that can shift over time as needed
- Benefits of autonomy depend on where applied
 - Significant benefits from autonomy that transfers, integrates, and transforms information to that needed (Level 1 and Level 2 SA)
 - But filtering can bias attention, deprive projection (Level 3 SA)
 - Significant benefit from autonomy that carries out tasks
 - Performance can be degraded by autonomy that simply generates options/strategies
- Flexible autonomy: Ability to switch tasking from human to automation and back over time and changes in mission tasks
 - Provides maximum aiding with advantages of human
 - Must be supported through the interface
 - Keep humans in the loop



Flexible Autonomy

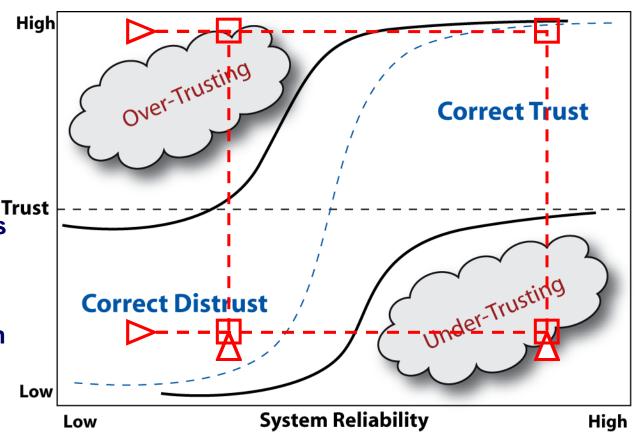








- Simple model showing partitioned trust/reliability space*
- Can use to explore transitions in trust and reliability over time
- But trust depends on many other factors
- And trust, in turn, drives other system-related behaviors, particularly usage by the operator
- But there's more we can do in the way of design and training...





Ways to Improve Human Trust of Autonomous Systems (1 of 2)



Cognitive congruence or analogical thinking

- Architect the system at the high level to be congruent with the way humans parse the problem
- If possible, develop aiding/automation knowledge management processes along lines of the way humans solve problem
- Example is convergence of Endsley's SA model with the JDL fusion model

Transparency and traceability

- Explanation or chaining engines
- If the system can't explain its reasoning, then the human teammate should be able to drill down and trace it
- Context overviews and visualizations at different levels of resolution
- Reducing transparency by making systems too "human-like" has the added problem of over-attribution of capability by the human user/teammate
 - Visually, via life-like avatars, facial expressions, hand gestures, ...
 - Glib conversational interface (e.g., Eliza)



Ways to Improve Human Trust of Autonomous Systems (2 of 2)



Self-consciousness of system health/integrity

- Metainformation on the system data/information/knowledge
- Health management subsystems should monitor the comms channels, knowledge bases, and applications (business rules, algorithms, ...)*
- Need to go far beyond simple database integrity checking and think in terms of consistency checkers at more abstract levels, analogs to flight management health monitoring systems, ...

Mixed initiative training

- Extensive human-system team training, for nominal and compromised behavior
- To understand common team objectives, separate roles and how they co-depend
- To develop mutual mental models of each other, based on expectations for competence, dependability, predictability, timeliness, uncertainty reduction, ...







Background and context

- Challenges to overcome
- Approaches to solutions
- Next steps



Four Tracks Towards Autonomy (1 of 2)



Cybernetics

- 1940's: The scientific study of control and communications in the animal and the machine (Norbert Weiner)
- 50's 70's: Manual control (e.g., flight simulators)
- 70's 90's: Supervisory control (e.g., FMS)
- 90's present: Cognitive models with a systems bent (e.g., COGNET, SAMPLE)

Symbolic Logic ("hard" AI)

- 50's: Turing Test, "Artificial Intelligence" Dartmouth Symposium, General Problem Solver (Newell and Simon)
- 60's 80's: Symbolic/linguistic focus, expert systems, logic programming, planning and scheduling
- 80's present: Cognitive models with a logic bent (e.g., Soar)



Four Tracks Towards Autonomy (2 of 2)



- Computational Intelligence ("soft" AI)
 - 40's: Artificial Neural Networks (ANNs)
 - 50's: ANNs with Learning (Turing again, Hinton, LeCun)
 - 60's present: Genetic/Evolutionary Algorithms (Holland, Fogel)
 - 60's 90's: Fuzzy Logic (Zadeh)
 - 80's present: Deep Learning
 - We've ceased to be the lunatic fringe. We're now the lunatic core. (Hinton)
 - Merging architectures for Big Data and Deep Learning, to influence cognitive architectures

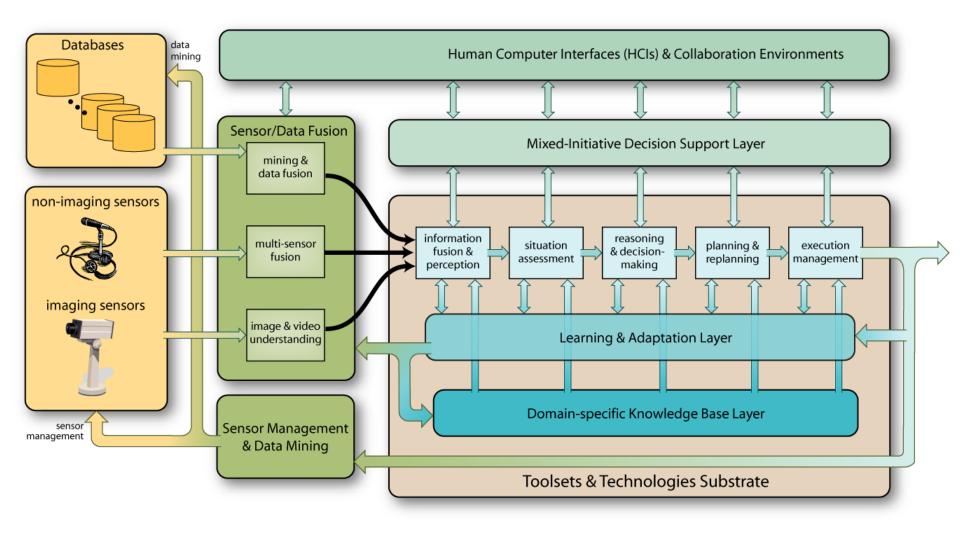
Robotics

- ~1900's: Remote control of torpedoes, airplanes
- 30's present: "Open loop" in-place industrial robots
- 40's 70's: Early locomoting robots
- 70's present: "Thinking" locomoting robotics
 - Actionist approach (e.g., Brooks' iRobot, Google Cars, ...)
 - Sensor-driven mental models of "outside" world; drive to "cognition"



Potential Framework for Autonomous Systems R&D









Autonomous Horizons Volume II

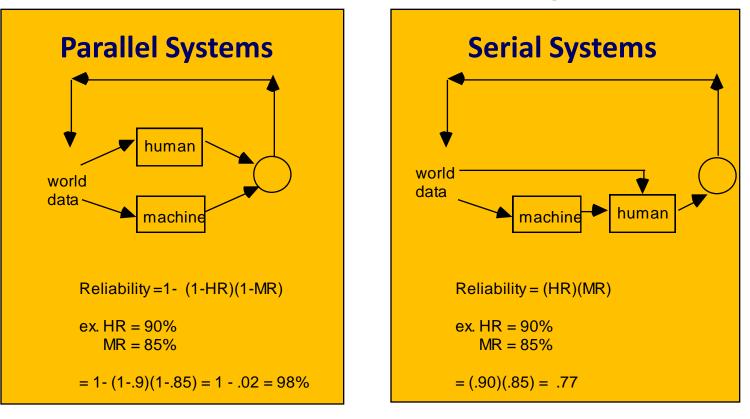
- Focus on developing a framework that will reach across communities working autonomy issues
 - Identify high payoff AF autonomous systems applications
 - Identify technical interest groups working these problems, via Autonomy COI, others
- Specify key "under the hood" functions included in that framework (e.g., planning)
- Evaluate key technologies that can support implementation of these functions (e.g., optimization)
- Lay out a research strategy and demonstration program
- Autonomous Horizons Volume III
 - Focus on critical implementation issues, including: cyber security, communications vulnerability, V&V



Independent, Objective, and Timely Science & Technology Advice



People take the recommendation as another information source to combine with their own decision processes





Human-Autonomy Interaction



Robustness

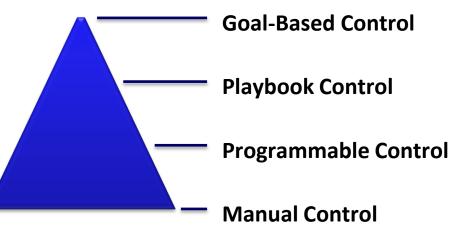
 The degree to which the autonomy can sense, understand, and appropriately handle a wide range of conditions

Span of Control

 From only very specific tasks for specific functions, up to autonomy that controls a wide range of functions on a system.

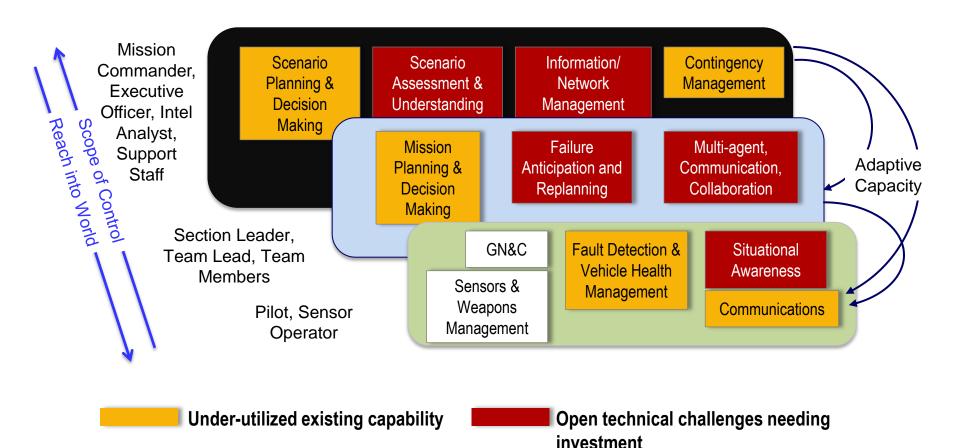
Control Granularity

 Level of detail in the breakdown of tasks for control



Missed Opportunities and Needed Technology Developments





*Defense Science Board , Task Force on the Role of Autonomy in the DoD Systems, 2012



STATES 42P

Overtrust

- A DC-10 landed at Kennedy Airport, touching down about halfway down the runway and about 50 knots over target speed. A faulty auto-throttle was probably responsible. The flight crew, who apparently were not monitoring the airspeed, never detected the over-speed condition.
- In 1981 a DC-10 crashed into Mt. Erebus in Antarctica. The accident was primarily due to incorrect navigation data that was inserted into a ground-based computer, and then loaded into the on board aircraft navigation system by the flight crew. The inertial navigation system (INS), erroneously programmed, flew dutifully into the mountain.

Misuse

 While climbing to altitude, the crew of a DC-10 flying from Paris to Miami programmed the flight guidance system to climb at a constant vertical speed. As altitude increased, the autopilot dutifully attempted to comply by constantly increasing the pitch angle, resulting in a high-altitude stall, and loss of over 10,000 feet of altitude before recovery.

(Bad) Human-System Teaming in the Commercial Cockpit (2 of 2)



Differing intentions across teammembers

- In a China Airlines Airbus A300 accident at Nagaya Japan, the autopilot continued to fly a programmed go-around, while the crew tried to stay on glide slope. The autopilot applied full nose-up trim and [the] aircraft pitched up at a high angle, stalled, and crashed.*
- Confusion over flight mode was the cause of a fatal A320 crash during a non-precision approach into Strasburg-Entzheim Airport in France. The crew inadvertently placed the aircraft into 3300 feet per minute descent when a flight crewmember inserted 3.3 into the flight management computer while the aircraft was in vertical descent mode instead of the proper flight path control mode. Pilots intended to fly a 3.3 glide slope.*
- The DHL B757 and Tu154M mid-air over Germany in 2002 might have been avoided if both crews had followed their onboard TCAS advisories: the B757 was told to dive, the Tu154M to climb. ATC, unaware of the advisories, told the Tu154M to dive. The B757 crew, trusting TCAS in a close conflict situation, dove. The Tu154 crew, trusting ATC, did also.**



Building Trust in Autonomous Systems

- Understanding autonomous system capability and limitations
 - Develop models, tools, and datasets to understand system performance
 - Experimentation with systems that change over time with the environment, and because of learning
- Understanding the boundaries within which the system is designed to operate, and the systems "experience"
 - Boundaries are situational, may evolve, and may violate the original system design assumptions
 - Systems will change over time because of learning, changing operator expectations
- Supporting effective man-machine teaming
 - Provide mutual understanding of common goals
 - Support ease of communication between humans and systems
 - Train together to develop CONOPS and skilled team performance, across wide range of mission, threat, environment, and users
- Assuring the operator of the system's integrity
 - Provide for transparency, traceability, and "explainability",
 - Support machine self-awareness, including boundary operation violations
 - Performance within boundaries must be reliable and secure
 - Awareness of operating outside the boundaries
- Identifying and addressing potential vulnerabilities
 - Red teaming early and often



Hierarchy for Supporting Collaboration



- Goal Alignment
 - Desired goal state actions need to support
 - Requires active goal switching based on prioritization
- Function Allocation/Re-allocation
 - Assignment of functions and tasks across team
 - Dynamic reassignment based on capabilities, status
- Decision Communication
 - Selection of strategies, plans and actions
 - needed to bring world into alignment with goals
- Task Alignment
 - Coordination of inter-related tasks for
 - effective overall operations

Shared Situation Awareness



Autonomy Functions



Machine Perception

- Vision
 - Image Processing and Computer Vision
 - Image Understanding
- Tactile Sensing
- Specialized Sensor Processing
 - EO, IR, Radar, Sonar,...
- Event Detection
- Situation Assessment
 - External Environment
 - Internal Environment
 - Health Awareness
 - Confidence specification (of assessments)

Reasoning



Autonomy Functions



Planning and Scheduling

Motor Control

- Locomotion
- Motor Control (manipulation)
- Sensor control
- Learning
 - Knowledge Acquisition
 - Adaptation/Learning
- Performance Monitoring/assessment
 - Performance awareness
 - Capability awareness (operating envelope)
- Reconfiguration/repair (of self)



Autonomy Functions



Human Computer Interface

- Auditory Channel
 - Alarms
 - Natural Language Processing
 - Signal Processing
 - Speech Recognition
 - Signal Processing
 - Computational Linguistics
 - Speech Synthesis

- Haptic Channel
- Visual Channel
 - Image Processing
 - Face recognition
 - Gesture Recognition
 - Object Recognition
 - Display/Visualization