Interaction-Centred Design for Optimizing Human-Autonomy Symbiosis Technology

Dr. Ming Hou

Department of National Defence, Canada

DRDC | RDDC

IEEE CS/RA/SMC Vancouver Chapters, 30 July, 2021

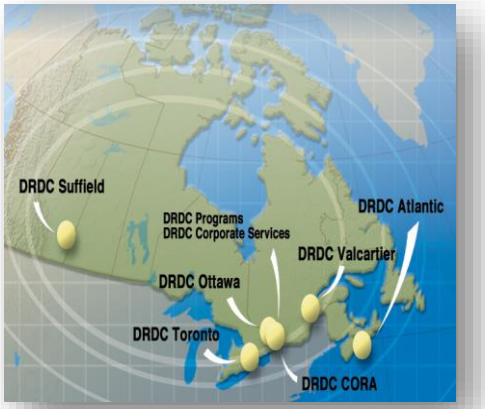
Canada

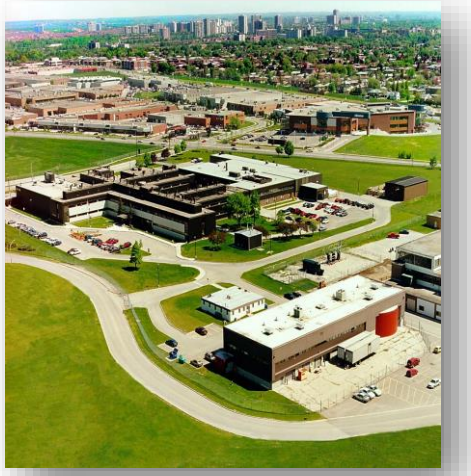
Outline

- DRDC Toronto Research Centre (Who)
- Human-Autonomy Symbiosis Technology (WHAT)
- Interaction Issues with AI/Autonomy (Why)
- Interaction-Centered Design for Optimization (How)
- Application Examples of Interaction-Centered Design Approach (Where and When)

DRDC Toronto Research Centre (TRC) & Canadian Forces Environmental Medicine Establishment (CFEME)

- DRDC TRC is Canada's centre of excellence for human effectiveness science and technology in the defence and national security environment.
- Environmental Medicine, physiology, psychology, sociology, and human factors expertise
- Emerging Areas for Human Effectiveness:
  - Human-autonomy teaming
  - Human interaction with digital media (VR and MR)
  - Human interaction with mobile computing
  - Simulation-based training and acquisition
  - The use of cognitive modelling or AI to understand and enhance human effectiveness





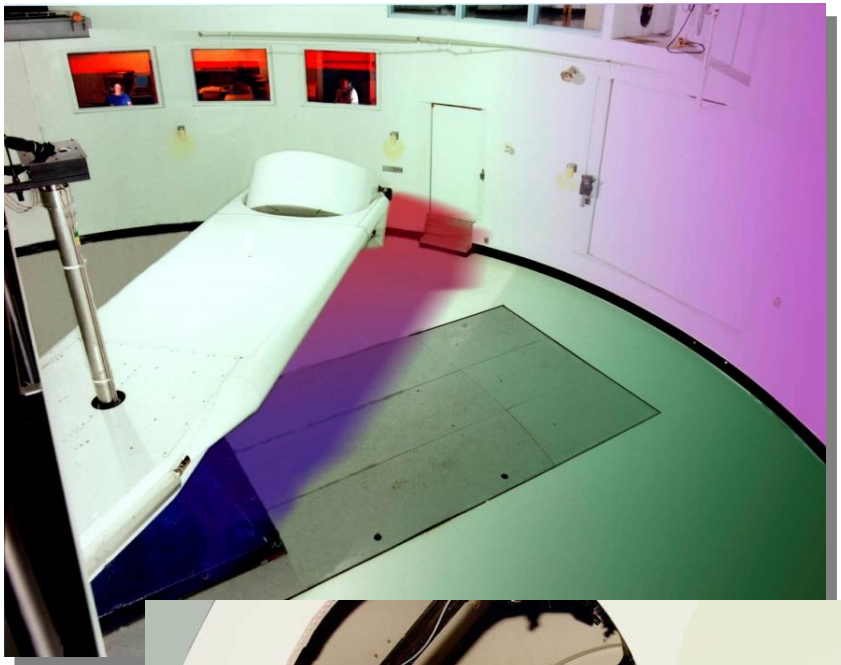
Mission

To enhance the effectiveness and ensure the health and safety of the human in any human-machine system or adverse environment.

TRC was founded in 1939, in support of Canada's war effort, under the leadership of Nobel Prize Laureate, Sir Frederick Banting.

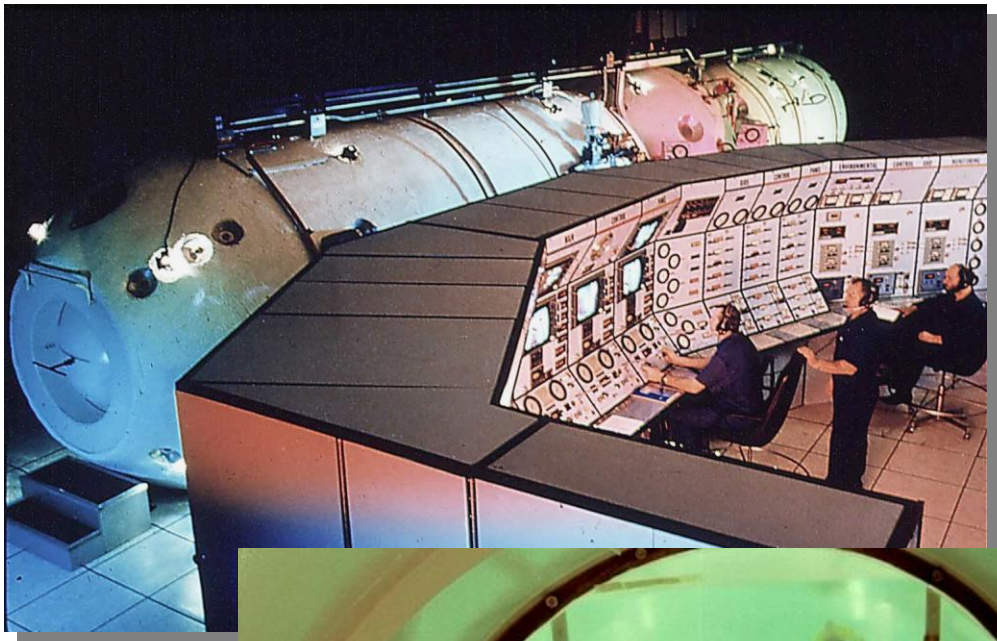
The Only Human Centrifuge in Canada

- Simulates the rapid-onset G-forces experienced by pilots flying high performance aircraft.
- Used for Research and Development (R&D) in aircrew protective equipment, aircrew training, and medical assessment.



The Unique Diving Chamber in the World

- Simulates underwater environments for military and commercial applications
- Human-rated with maximum pressure 17.2 MPa (1725 metres)
- Selection and training for astronauts



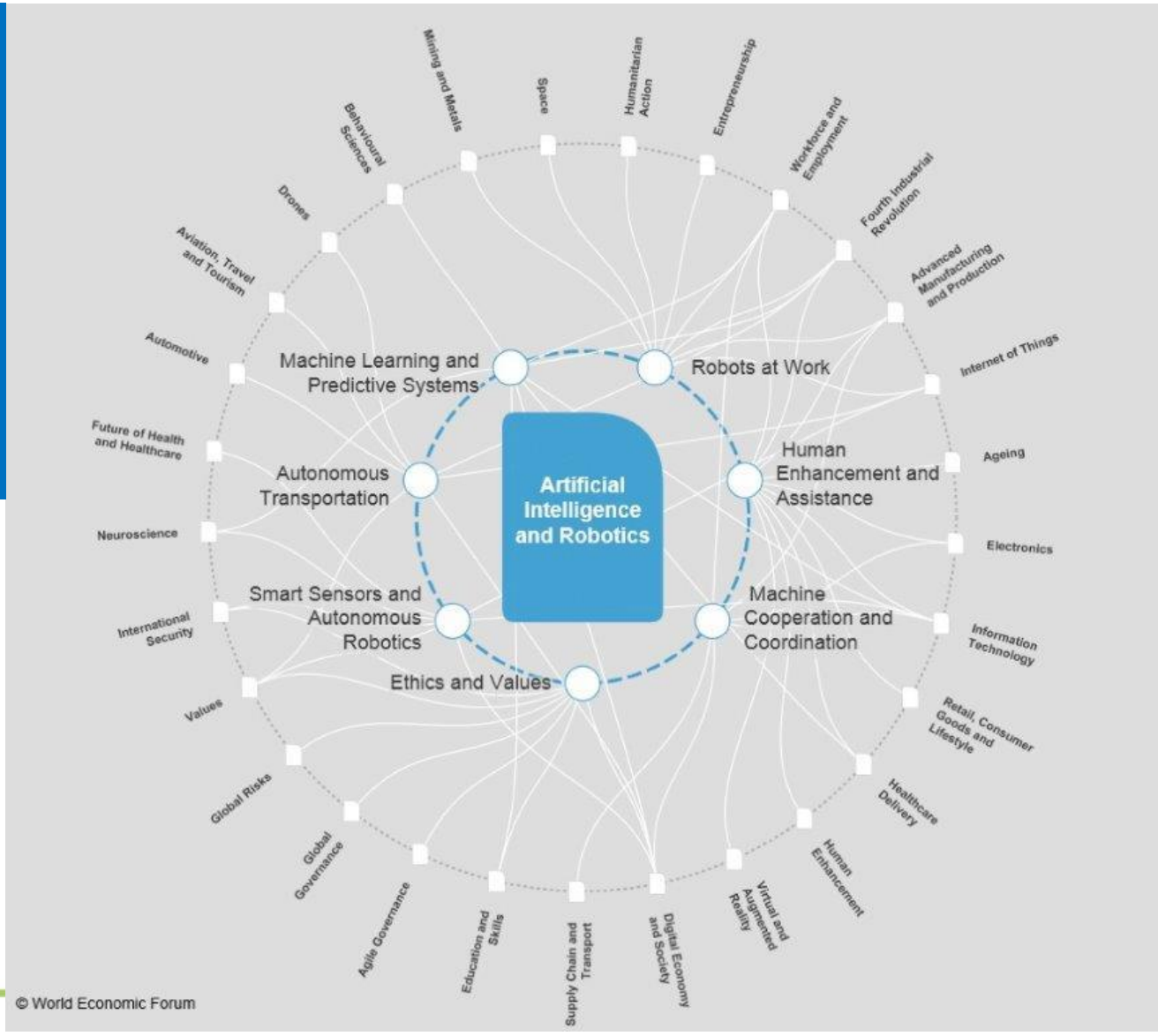
The Only Unmanned Aircraft System (UAS) Command & Control (C2) Center in the World

Outline

- DRDC Toronto Research Centre (Who)
- Human-Autonomy Symbiosis Technology (WHAT)
- Interaction Issues with AI/Autonomy (Why)
- Interaction-Centered Design for Optimization (How)
- Application Examples of Interaction-Centered Design Approach (Where and When)

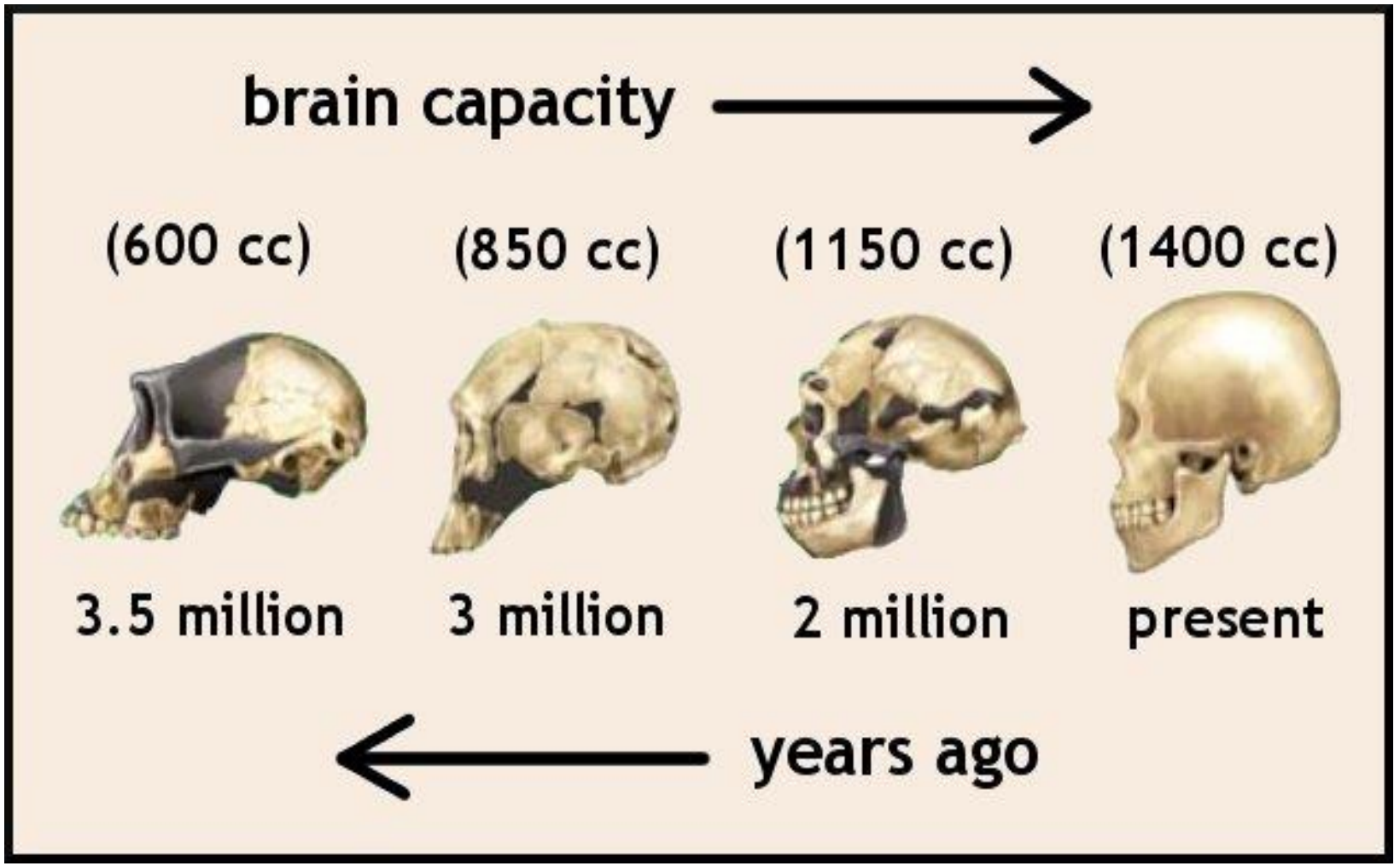
AI & Robotics, the 4th Industrial Revolution

According to the World Economic Forum, AI and Robotics will be pervasive to many domains of human activity



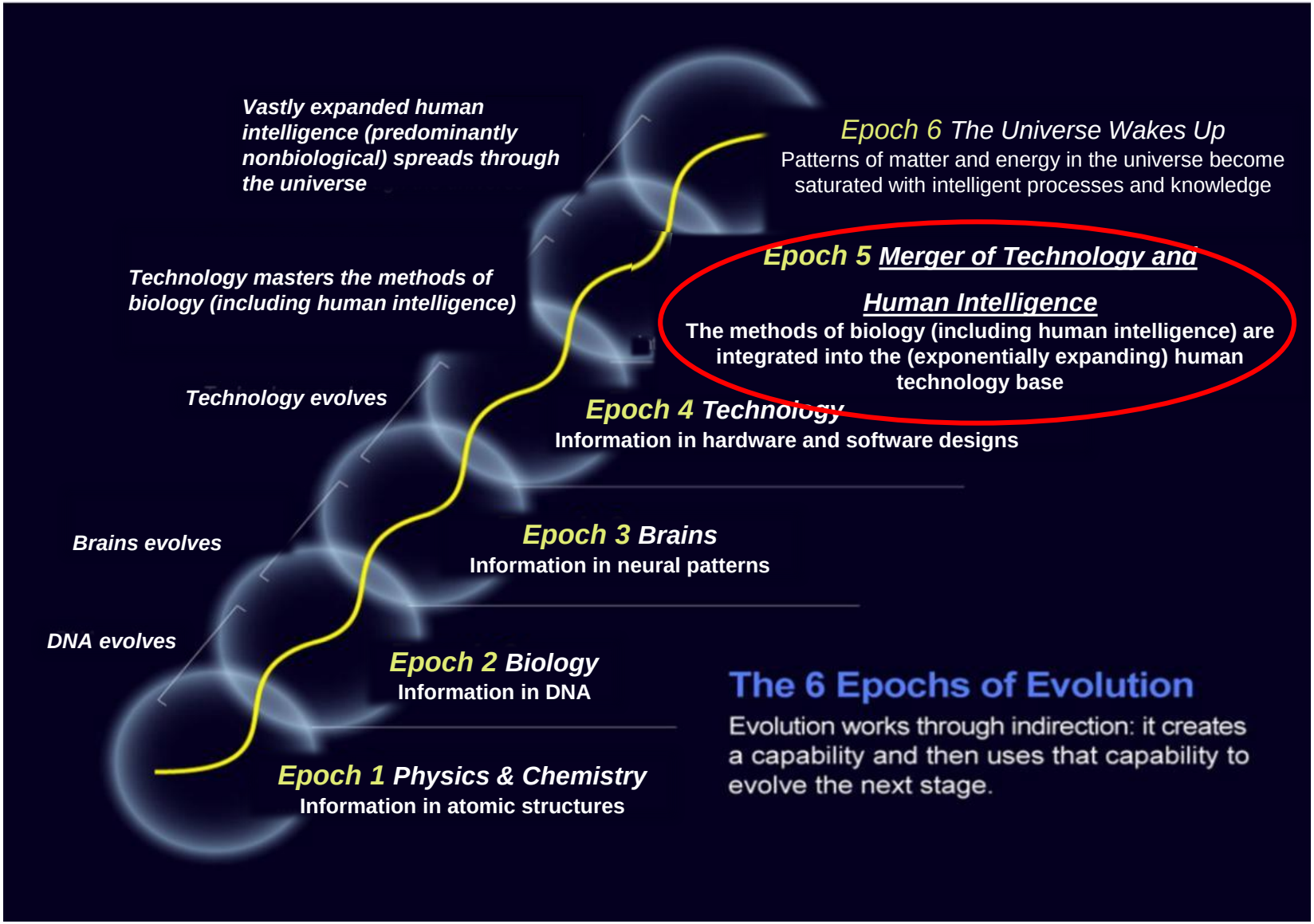
Impact
on
Human Cognitive Capacity
towards
Human-Autonomy Symbiosis

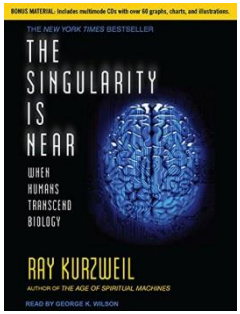
Evolution of Human Cognitive Capacity



https://jeevanshu.wordpress.com/

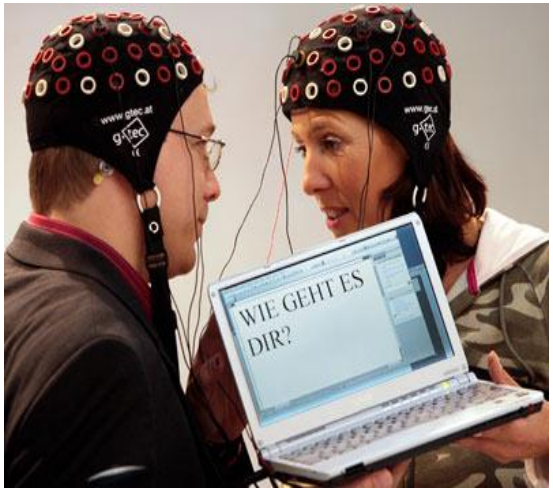
Evolution of Human Capabilities





Amazon.co.uk

C^3: $Capability_{(Tech)}$, $Capacity_{(Human)}$, $Complexity_{(Env)}$



Brain Computer Interface

https://ugs.utexas.edu

Robonaut co-worker in space

www.nasa.gov

Control of a swarm of drone with mind

engineering.com wetalkuav.com

Electronic tattoo display runs on blood

phys.org

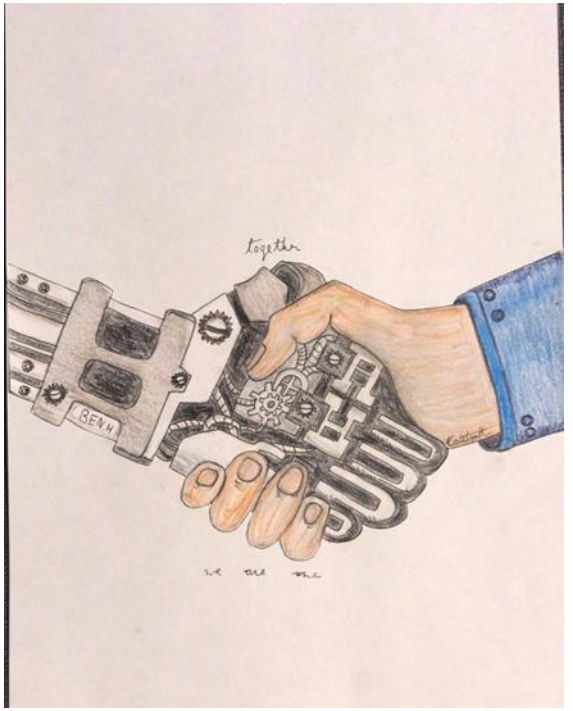
Outline

- DRDC Toronto Research Centre (Who)
- Human-Autonomy Symbiosis Technology (WHAT)
- Interaction Issues with AI/Autonomy (Why)
- Interaction-Centered Design for Optimization (How)
- Application Examples of Interaction-Centered Design Approach (Where and When)

HOWEVER

Issues of Human-Autonomy Symbiosis (HAS) Technology with AI Enabled and/or Computerized Decision-Making:

- Transparency
- Cognitive load
- Trust
- Accountability
- Legal and Ethical Aspects
- Policy and Regulations



Death by Algorithm

The errant algorithm went undetected for 9 years

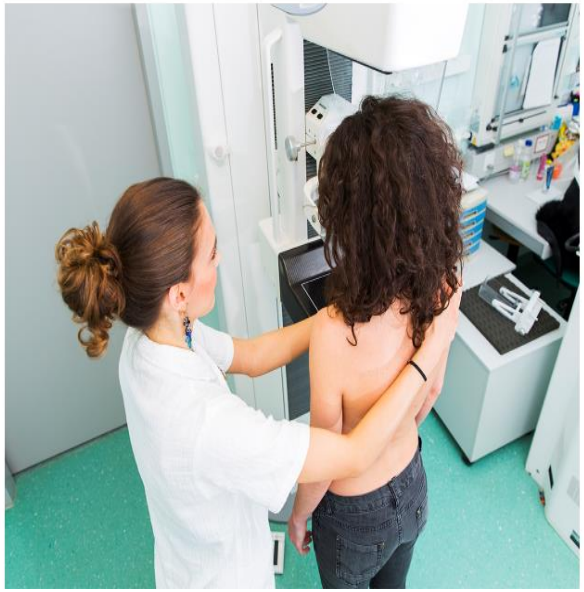
May 2017, the U.K. Health Minister announced a "computer algorithm failure" caused about 450,000 patients to not be invited for their final breast cancer screening. An estimate suggested that between 135 and 270 women might have died prematurely as a result of this blunder.

11 May 2018 | 17:40 GMT

450,000 Women Missed Breast Cancer Screenings Due to "Algorithm Failure"

A disclosure in the United Kingdom has sparked a heated debate about the health impacts of an errant algorithm

By Robert N. Charette



The U.K. Government said there were still 309,000 women alive who missed their final screening.

"NHS would revise its investment figure to include the £100 million that many lawyers feel the government will need to pay out in compensation for this latest case of avoidable harm"

Source: https://spectrum.ieee.org/riskfactor/computing/it/450000-woman-missed-breast-cancer-screening-exams-in-uk-due-to-algorithm-failure

Psychopath AI showing horrors of human biased

- A MIT AI machine tested how specific data fed into an algorithm can impact "outlook". The machine, named Norman was fed with grisly images and went through Rorschach test to "calculate" his mental health with terrifying results when responding to images:

http://norman-ai.mit.edu/#inkblot

Norman Sees	Images	Standard AI Sees
Man is shot dumped from car		An plane flying with smoke coming from it
Man is murdered by machine gun in broad daylight		A black and white photo of a baseball glove
Man is shot dead in front of his screaming wife		A person is holding an umbrella in the air

- Demonstrates: dangers of biased and prejudice. Norman only responds in horror because that's all he's been trained. (How can we regulate human bias in AI?)

Facebook Shuts down AI Robots due to FEAR

- Facebook challenged two AI chatbots to negotiate a trade, attempting to swap items with assigned values
- Watched over by Facebook AI Researchers and Linguists who realized the two AI robots were quick to break down the negotiations. The robots chanted at each others in a language that was incomprehensible to human

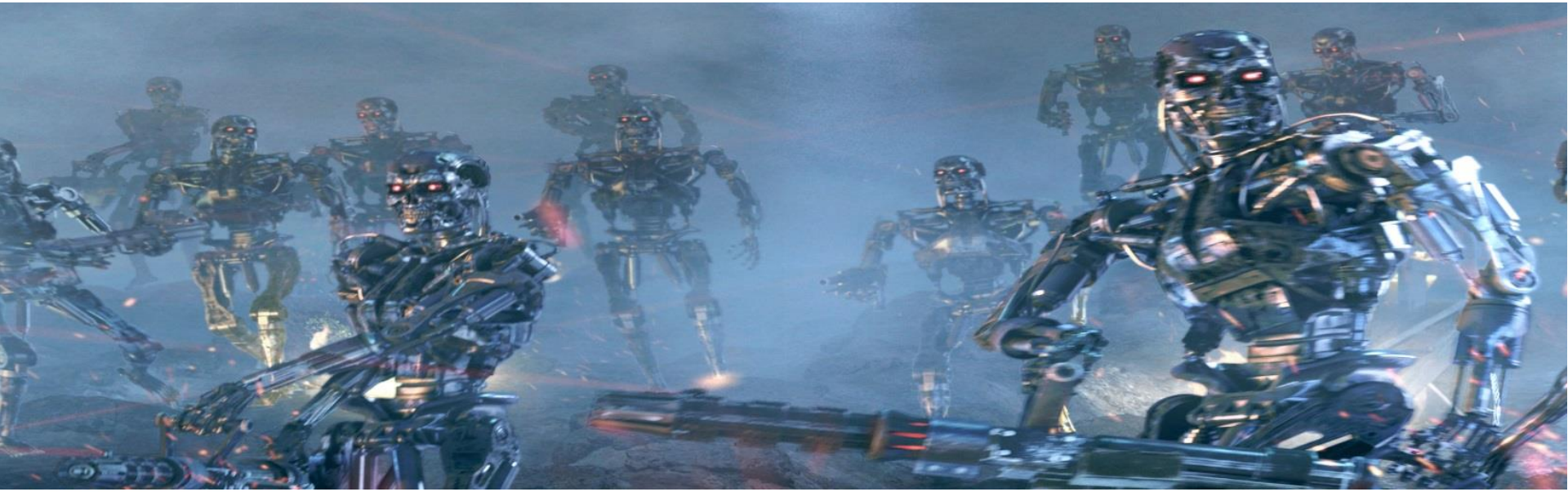
Source: https://www.independent.co.uk/life-style/gadgets-and-tech/news/facebook-artificial-intelligence-ai-chatbot-new-language-research-openai-google-a7869706.html

Rise of AI Machines: the World's First Android Citizen

In November 2017, a robot Sophia was given citizenship of Saudi Arabia – the first robot given legal personhood anywhere in the world

However, this AI robot says that she wants to destroy Humans…

Source: https://globalnews.ca/news/3844031/saudi-arabia-robot-citizen-sophia/



Source: de. fanpop.com

AI Will ‘transform or destroy’ Society

Stephen Hawking

businessinside.com

“Success in creating effective AI, could be the biggest event in the history of our civilization. Or the worst. We just don't know. So we cannot know if we will be infinitely helped by AI, or ignored by it and side-lined, or conceivably destroyed by it,…

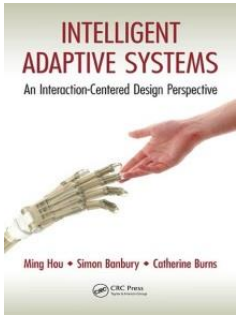
Unless we learn how to prepare for, and avoid, the potential risks, AI could be the worst event in the history of our civilization. It brings dangers, like powerful autonomous weapons, or new ways for the few to oppress the many. It could bring great disruption to our economy."

Design Flaws and Catastrophic Consequences

(Loss of 228 lives 12 years ago)

1 June 2009, Air France Flt 447 from Rio de Janeiro, Brazil to Paris, France crashed into the Atlantic Ocean. The accident is the deadliest in the history of Air France, and the deadliest aviation accident involving the Airbus A330.

http://www.spiegel.de



But we learned from our mistakes…right?

However, History Ain't Changed!

We learn in A HARD WAY with the price of human lives

More Catastrophic Consequences

(loss of 346 lives within only 5 months in 2018/2019 due to an uncontrollable AI-enabled system with autonomous functions: Maneuvering Characteristics Augmentation System (MCAS) on Boeing 737 Max

Youtube.com

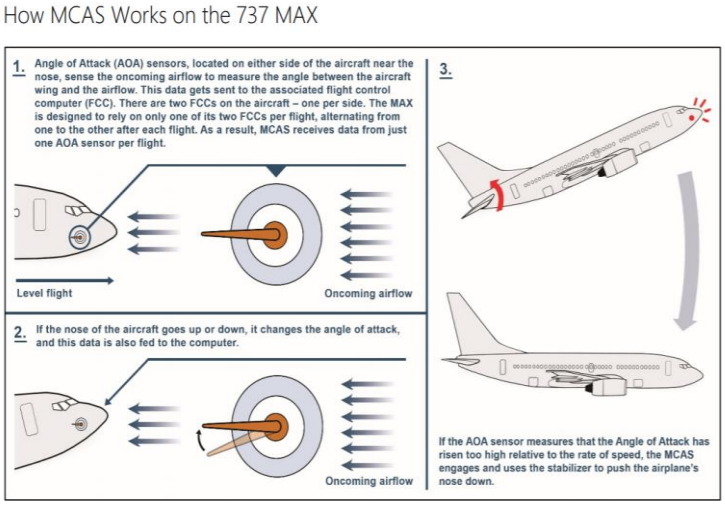
Teknologi.id

reuters.com

Design Failure of MCAS on Boeing 737 Max

Three of five main reasons for the crashes:

1. **Pressures to update <u>design</u> swiftly and inexpensively (cost reduction),**
2. **Faulty assumption that <u>pilots</u> could recognize and override a malfunction of the system within a few seconds,**
3. **Did not comply with <u>regulations</u> and jeopardized the flying public safety**



Source: OIG analysis of FAA and Boeing data

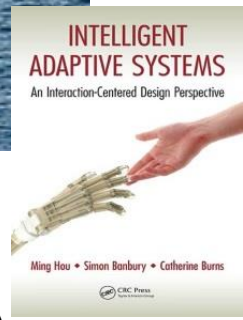
Sources: Majority Staff of the U.S House Committee on Transportation and Infrastructure. (2020). *Final Committee Report: the design, development, and certification of the Boeing 737 Max.*

TheFlightChannel. (2019, November 19). Boeing 737 MAX Crashes Immediately After Takeoff | Here's What Really Happened to Flight 302 [Video].

System Design and Human Factors

when transition interaction from "on-the-loop" to "in-the-loop"

15 Jan 2009, US Airways Flt 1549 lost engine power after about two minutes departed from LaGuardia Airport in New York City. The pilot (Capt Sully) made a quick decision (100+ seconds) to land safely in the Hudson River and all 155 people survived.



Intelligent Adaptive Systems: An Interaction-Centered Design Perspective (Hou et al., 2014)

What a Real Life Story in a Hollywood Movie "Sully" Tells About?

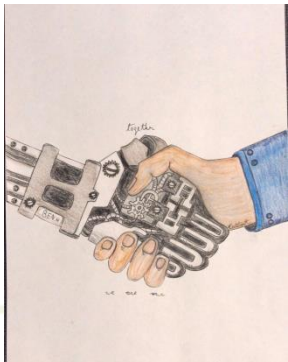
Human Factors and Interactions

Outline

- DRDC Toronto Research Centre (Who)
- Human-Autonomy Symbiosis Technology (WHAT)
- Interaction Issues with AI/Autonomy (Why)
- Interaction-Centered Design for Optimization (How)
- Application Examples of Interaction-Centered Design Approach (Where and When)

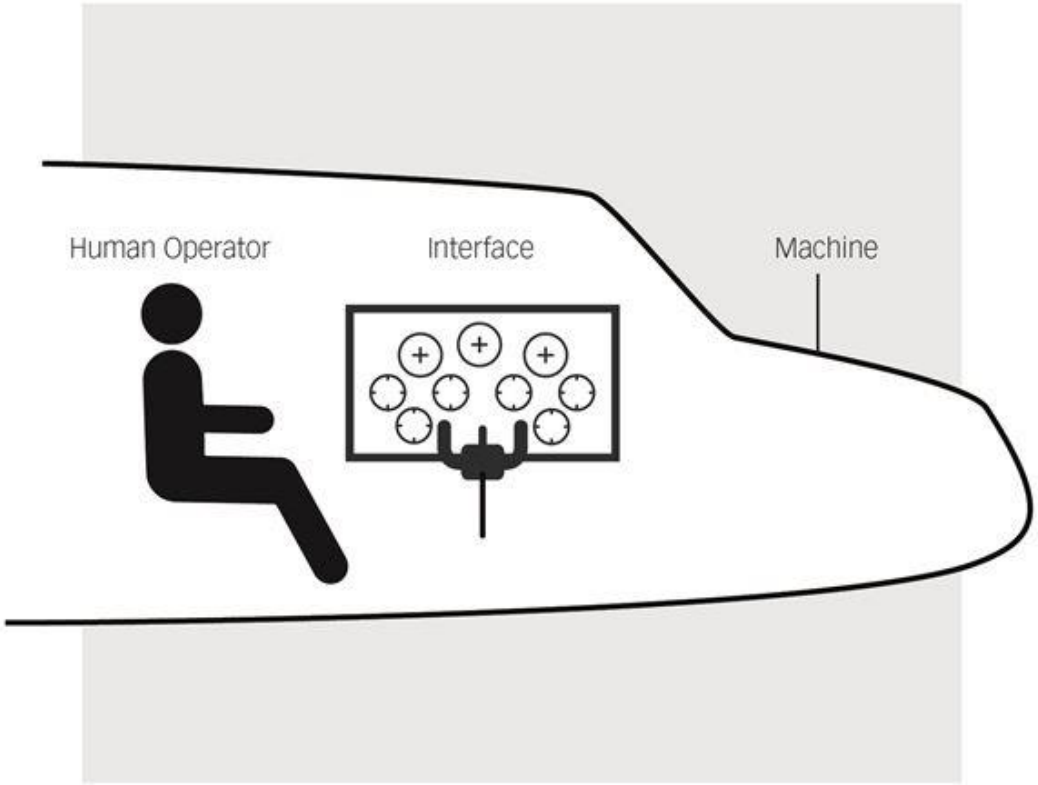
System Design Philosophy

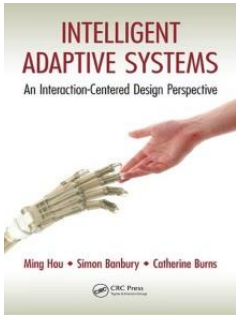
1. You may believe in machine, however, the best technology might not be the most advanced or even the first invented, but the safest to human;
2. Believe more in human in emergency as human has not only intelligence, but more importantly, wisdom, ethics, and guts in our heart; and
3. Design not only for technology but a trusted and collaborative partnership between human and machine in a human-autonomy symbiotic relationship.



A Human-Autonomy Symbiosis System

As represented by a human operator (i.e., pilot), a machine (i.e., aircraft), and the interface that allows them to interact



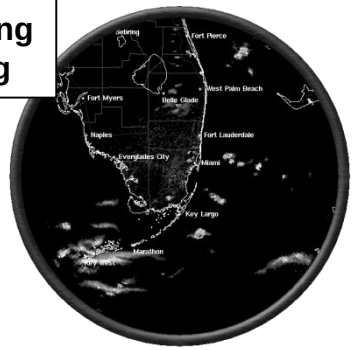


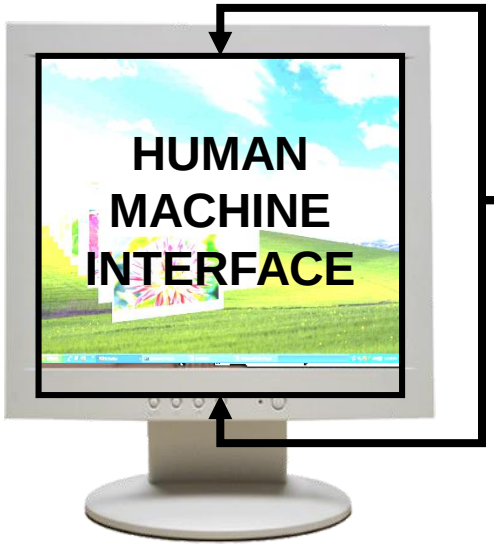
Intelligent Adaptive Systems: An Interaction-Centered Design Perspective (Hou et al., 2014)

A Conceptualized HAS System

- Knowledge of mission plans/goals
- Knowledge of mission time-lines
- Knowledge of mission tasks/activities
- Knowledge of mission environment

- Mission Plan/Goal Monitoring
- Control (task/activities) Monitoring
- Sensor (environment) Monitoring



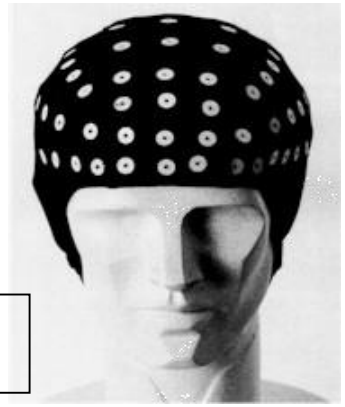


SITUATION ASSESSMENT

ADAPTATION ENGINE

Compare

OPERATOR STATE ASSESSMENT



- Psychophysiological Monitoring
- Behavioural Monitoring

- Model of human cognition
- Model of human control abilities
- Model of human communication
- Model of human interaction

Conceptual Framework of An Human- Autonomy Symbiosis (HAS)

- Knowledge of mission plans/goals
- Knowledge of mission time-lines
- Knowledge of mission tasks/activities
- Knowledge of mission environment

TASK MODEL

SYSTEM MODEL

WORLD MODEL

- Mission Plan/Goal Monitoring
- Control (task/activities) Monitoring
- Sensor (environment) Monitoring

- HMI Design Guidelines
- HCI Principles

Feedback Loop

Adapt HMI

DIALOGUE MODEL

Human MACHINE INTERFACE

KNOWLEDGE MODEL

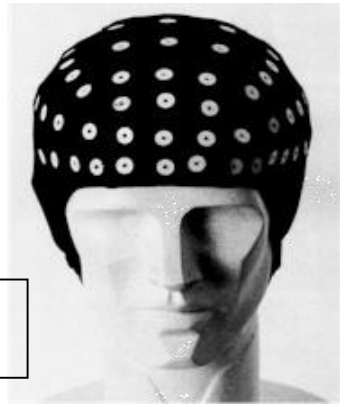
ADAPTATION ENGINE

SITUATION ASSESSMENT

Compare

OPERATOR STATE ASSESSMENT

Automation / Aid



- Psychophysiological Monitoring
- Behavioural Monitoring

- HMI Design Guidelines
- Automation-design Principles

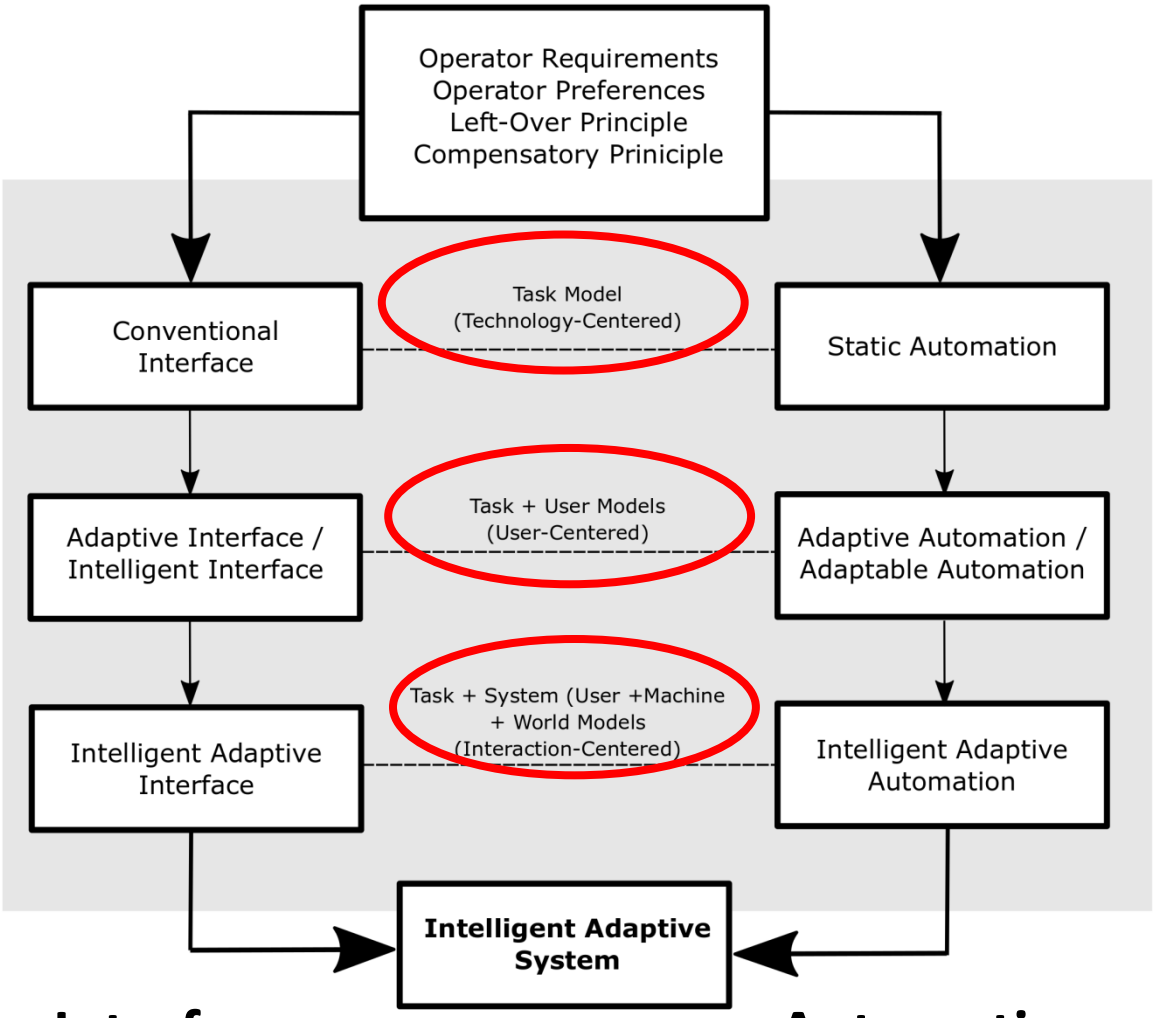
USER MODEL

- Model of human cognition
- Model of human control abilities
- Model of human communication
- Model of human interaction

(Hou, Gauthier, & Banbury, HCI 2007)

(Hou, Banbury, & Burns, 2014;2020)

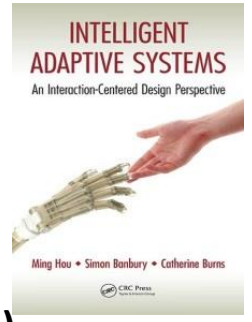
Evolution of Design Strategy for HAS



Technology - centered

User/human - centered

Interaction - centered



Intelligent Adaptive Systems: An Interaction-Centered Design Perspective (Hou et al., 2014)

Interaction-Centered Design (ICD) Strategic Roadmap

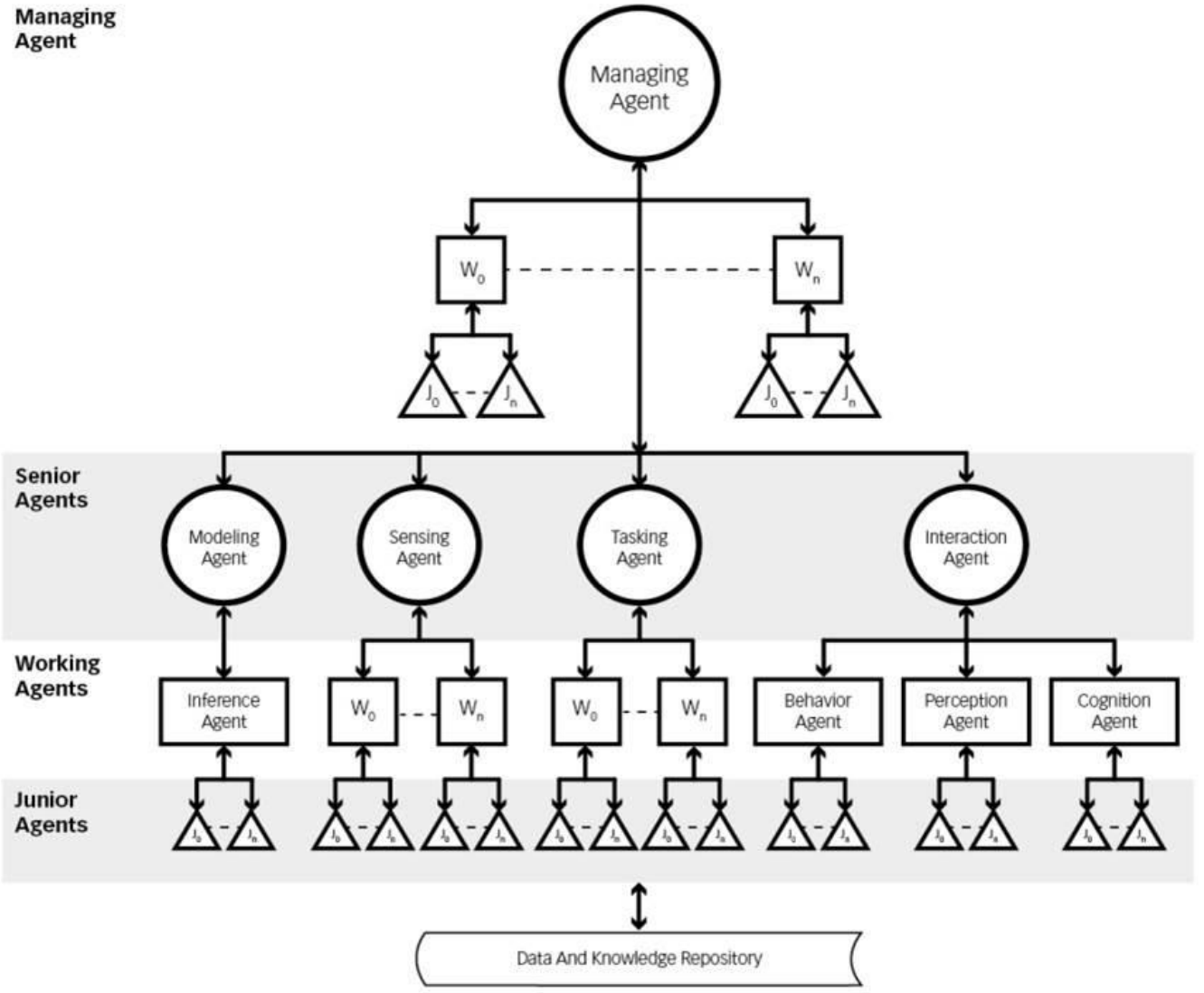
1. Conduct Taxonomic Analysis (Contextual Domain)
2. Select System Design Framework (for Interactions)
3. Select Analytical Techniques (System Requirements)
4. Select Design Methodology (Hardware and Software)
5. Select Operator-State Monitoring Approach (Meatware)
6. Comply with Design Guidelines (Domain Standards)

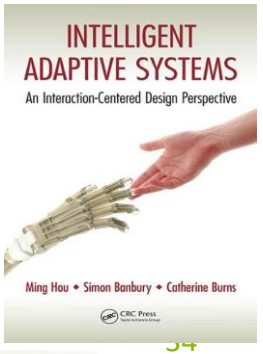
Hou, et al., 2007 (Journal of Cognitive Engineering and Decision Making)

Hou, et al., 2011 (IEEE Transactions on Systems, Man, and Cybernetics -- Part C)

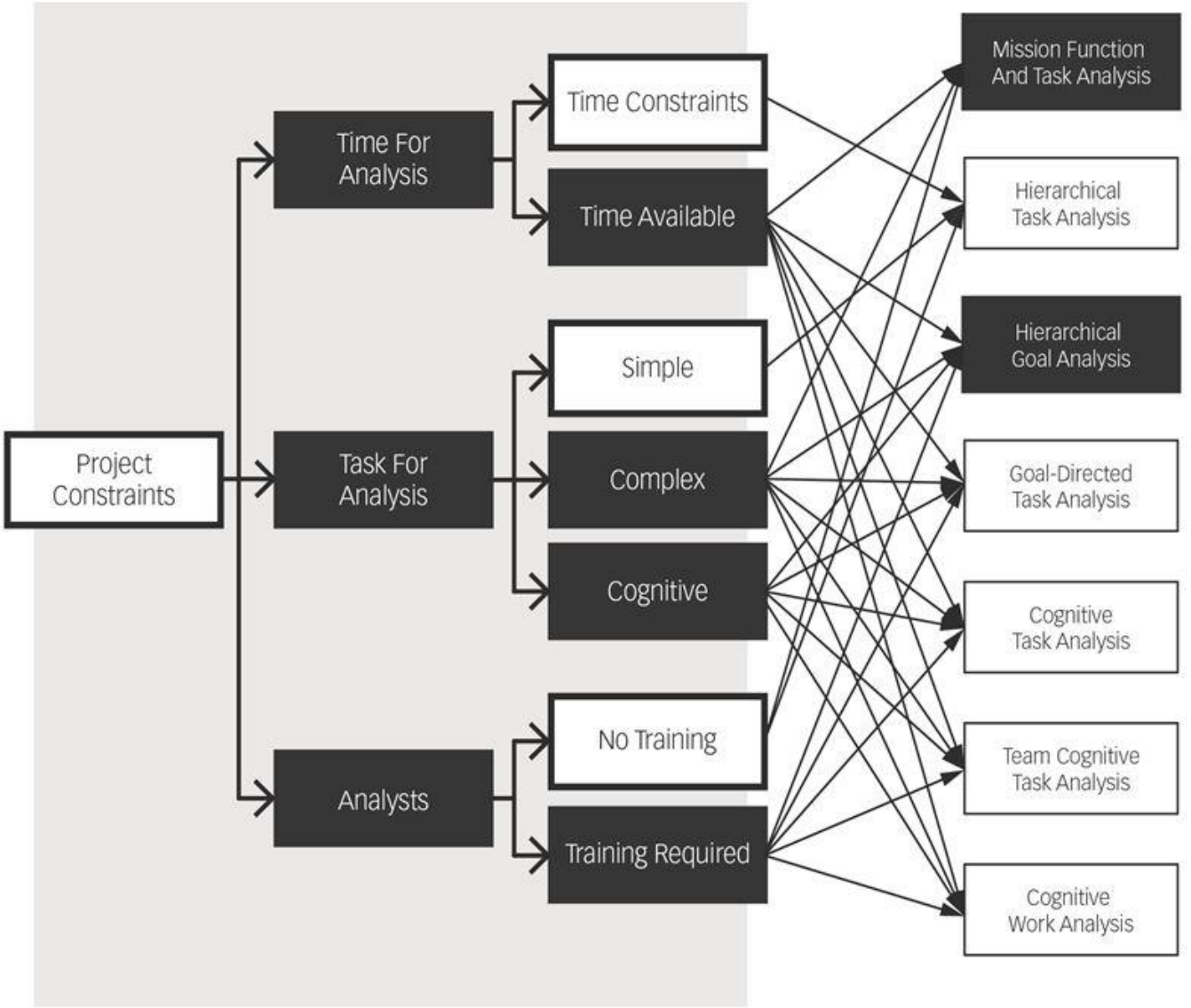
Hou, et al., 2014; 2020 (Intelligent Adaptive Systems – An Interaction-Centered Design Perspective)

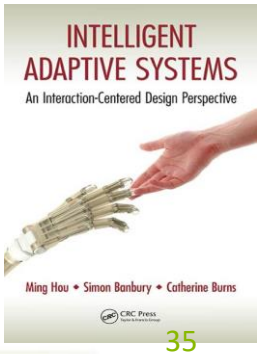
An Agent-Based Conceptual Framework for HAS Design





Selection of Analytical Techniques





Goal/Task Analysis for Functional Requirements

Microsoft Excel - Goal Tables.xls [Read-Only]

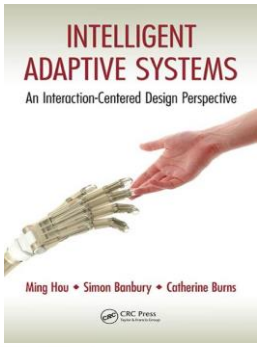
File Edit View Insert Format Tools Data Window Help Acrobat

Number	Level (Indent) & Goal/Objective and Subgoals/Subobjectives 1 2 3 4 5	IAI Candidate	Influenced Variable	Assignment	Completion Time (s)	Note
Top	I want to perceive the (...) conduct of the terrorist patrol mission					
7	... UAVs are employed					
7.1	... VTUAV is employed					
7.1.1	... VTUAV navigation is conducted					
7.1.1.1	... VTUAV route to the next operating area is planned	Yes	Route	UAV Plt	5 to 30	
7.1.1.2	... VTUAV search pattern is planned					
7.1.2.1	... selection of an appropriate search pattern	Yes	Pattern	UAV Plt	7 to 15	
7.1.1.2.2	... VTUAV expanding square search is planned					Not u
7.1.1.2.3	... VTUAV creeping line advance search is planned					Not u
7.1.1.2.4	... VTUAV track crawl search is planned					Not u
7.1.1.2.5	... VTUAV sector search is planned					Not u
7.1.1.2.6	... VTUAV random search is planned					Not u
7.1.1.2.7	... location of contact symbol is determined on tacplot		Location	UAV Plt	8	
7.1.1.2.8	... direction of movement of contact symbol		Direction	UAV Op	3	
7.1.1.3	... VTUAV waypoint is inserted	Yes	Waypoint	TACNAV	3 to 15	
7.1.1.4	... VTUAV endurance is calculated					Not u
7.1.1.5	... VTUAV destination is determined					Not u
7.1.1.6	... VTUAV time on task is calculated					
7.1.1.6.1	... location of potential VTUAV refuelling platforms		Location	UAV Plt	20	

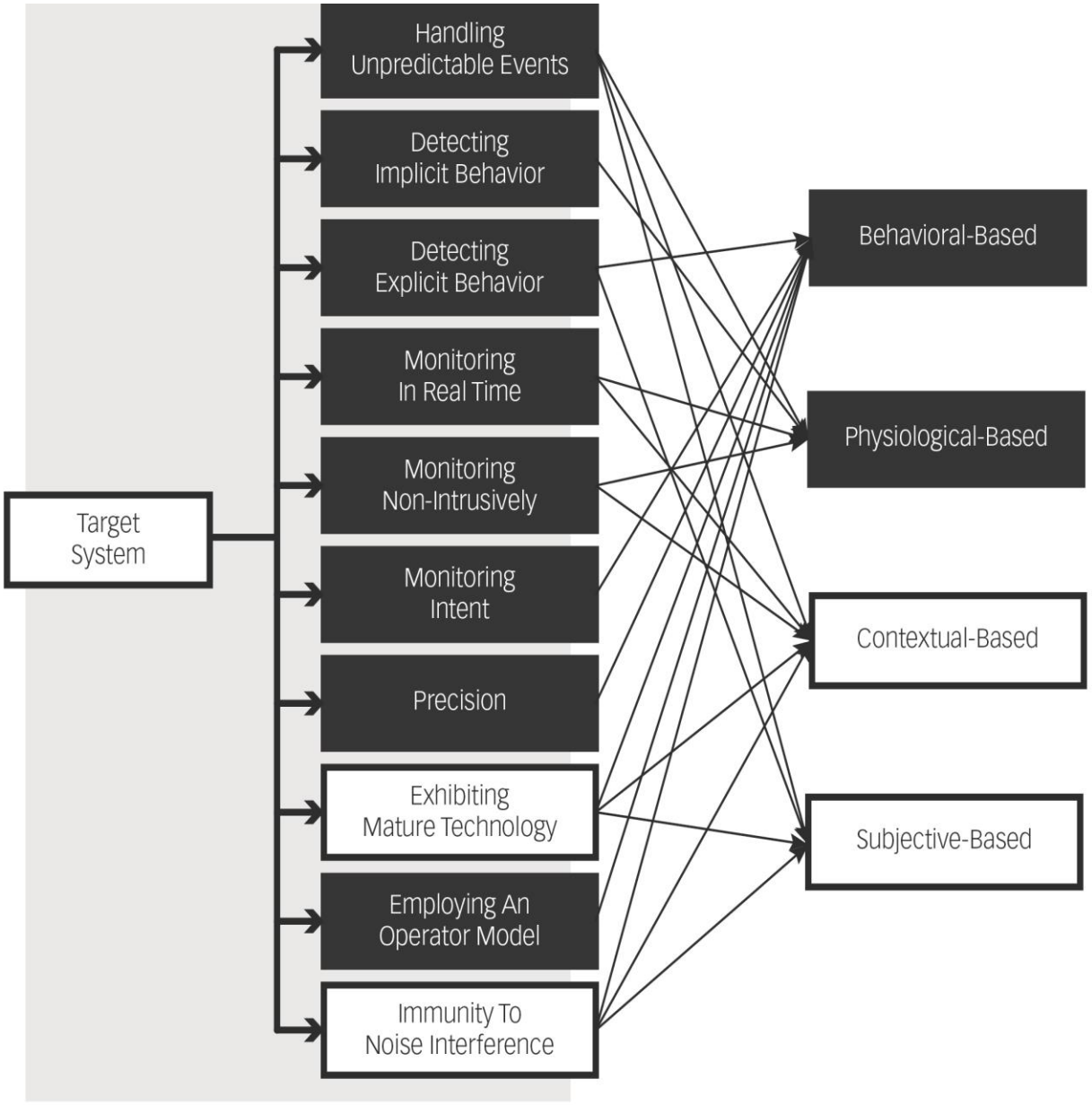
Top Goal | Goal 1 | Goal 2 | Goal 3 | Goal 4 | Goal 5 | Goal 6 | Goal 7 | Goal 8 | Goal 9 | Goal 10

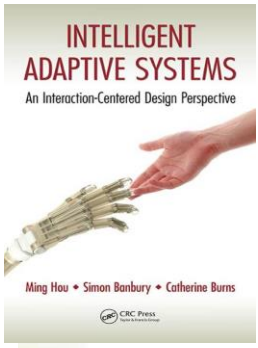
Ready

Start | Inbox - Microsoft O... | Calendar - Microsoft... | Telephone directory | Exploring - Final Re... | Microsoft Word | Microsoft Excel - ... | 5:52 PM



Selection of Operator Monitoring Approaches

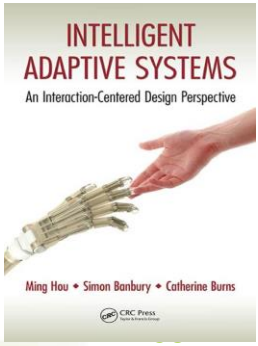




Design Issues to Address – W5+

for an Integrated Project Team (e.g., Project Manager, Design Engineer, and Systems Developer, etc.)

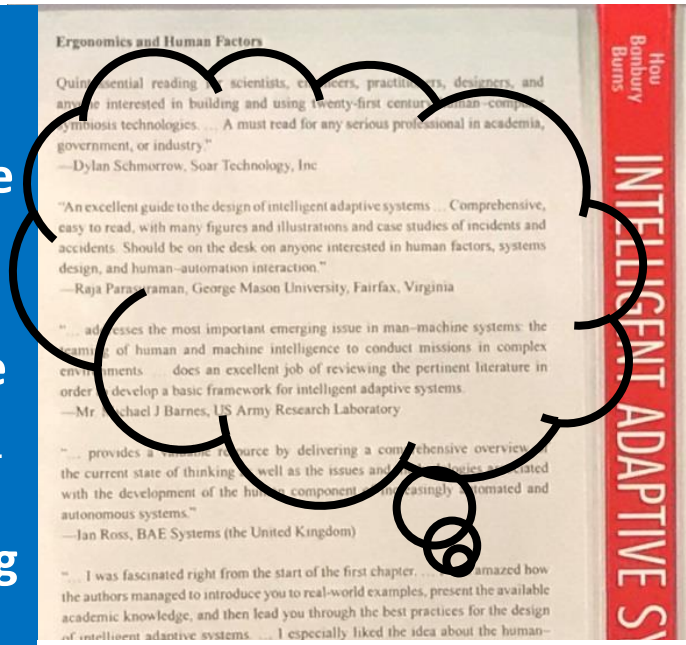
- Why (rationale, perceived need, perceived benefits)
- What (mission, task -- domain context)
- Where (working environment)
- When (timing, frequency, duration)
- Who (intended users/operators/stakeholders)
- How (use case/scenario – interaction)

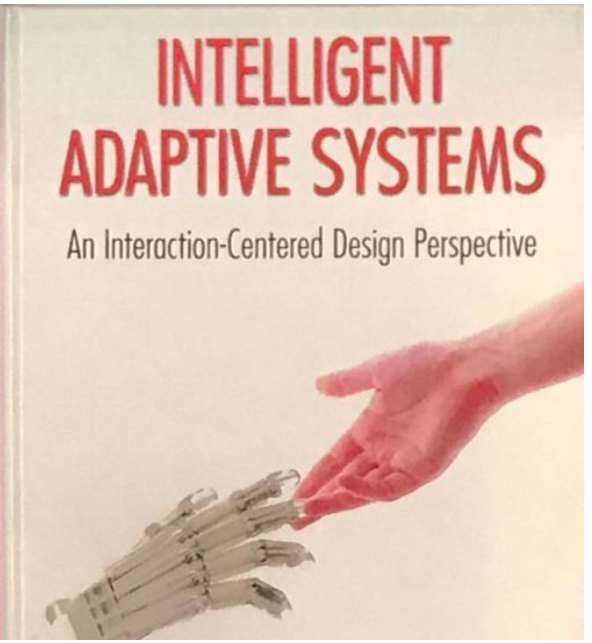


Peer Reviewed: Interaction-Centered Design (ICD) for A Collaborative Partnership (Human & Autonomy)

"…Setting the agenda for the coming years as Human Factors practitioners grapple with the demands that IAS will make on its operators and a clear statement of the importance of collaboration and partnership between Human and AI, and outlining how this can be achieved through interaction (centred) design…"

Book Review: Intelligent Adaptive Systems. C. Baber, University of Birmingham, Ergonomics, 2017, Vol. 60, No.10, 1458-1459.





A must read for any serious professional in academia, government, or industry, interested in building and using twenty-first century human-computer symbiosis technologies…

Dr. Dylan Schmorrow, Executive VP and Chief Scientist, Soar Technology, Inc., Former DARPA Executive

Adopted as Strategy, Principle & Guidance for NATO Standards Development

"… (HAS) capability development through building trust and process…"

LGen S. Kindsvater, NATO Deputy Chair of Military Committee

NATO Joint Capability Group Unmanned Aircraft Systems Spring Summit, 14/06/21


STANDARDIZATION RECOMMENDATION
RECOMMANDATION DE NORMALISATION

STANREC 4811

GUIDANCE ON SENSE AND AVOID FOR UNMANNED AIRCRAFT SYSTEMS

DIRECTIVES SUR LES SYSTÈMES DE DÉTECTION ET D'ÉVITEMENT POUR LES SYSTÈMES AÉRIENS SANS PILOTE

EDITION/ÉDITION 1
23 February/février 2018
NSO/0254(2018)JCGUAS/4811

NORTH ATLANTIC TREATY ORGANIZATION
ORGANISATION DU TRAITÉ DE L'ATLANTIQUE NORD

Published by the NATO STANDARDIZATION OFFICE (NSO)
Publiée par le BUREAU OTAN DE NORMALISATION (NSO)

© NATO/OTAN


STANDARDIZATION RECOMMENDATION
RECOMMANDATION DE NORMALISATION

STANREC 4685

HUMAN SYSTEMS INTEGRATION GUIDANCE FOR UNMANNED AIRCRAFT SYSTEMS

TITRE FRANÇAIS

EDITION/ÉDITION 1

2021

NORTH ATLANTIC TREATY ORGANIZATION
ORGANISATION DU TRAITÉ DE L'ATLANTIQUE NORD

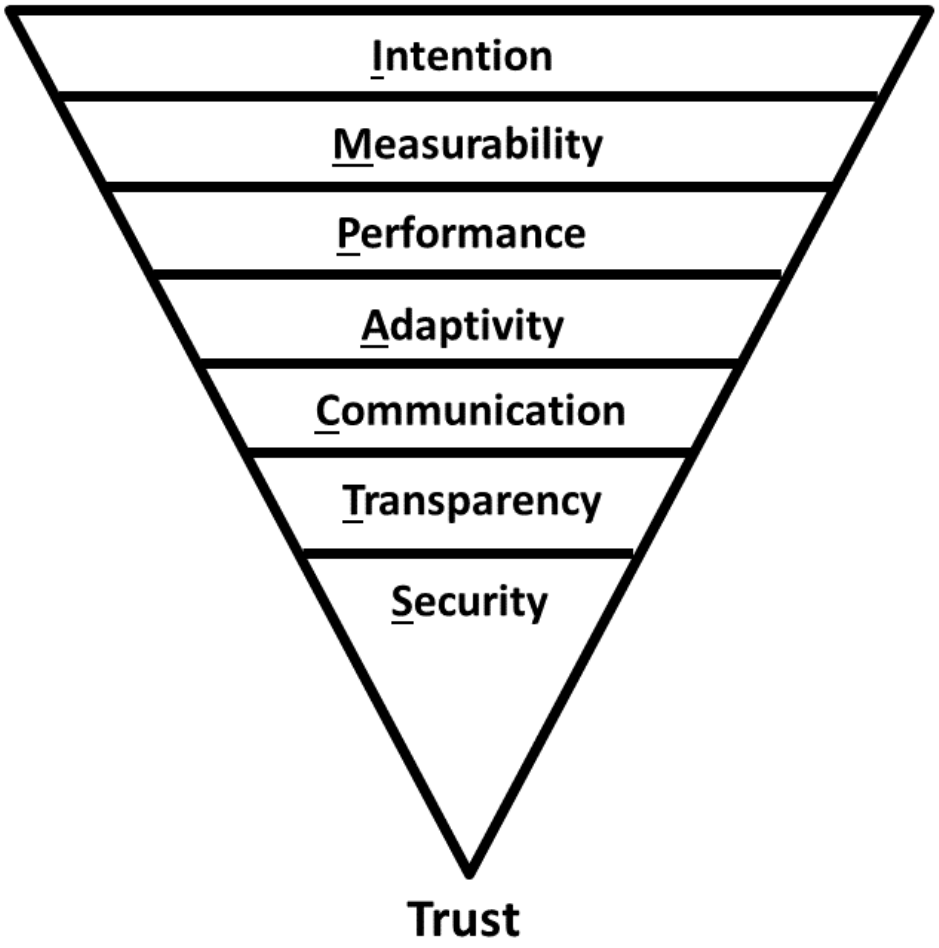
Published by the NATO STANDARDIZATION OFFICE (NSO)
Publiée par le BUREAU OTAN DE NORMALISATION (NSO)

© NATO/OTAN

STANREC 4811: Guidance on Sense and Avoid for Unmanned Aircraft System

STANREC 4685: Human Systems Integration Guidance for Unmanned Aircraft System

IMPACTS: A Trust Model for Human-AI Teaming and Human-Autonomy Interaction



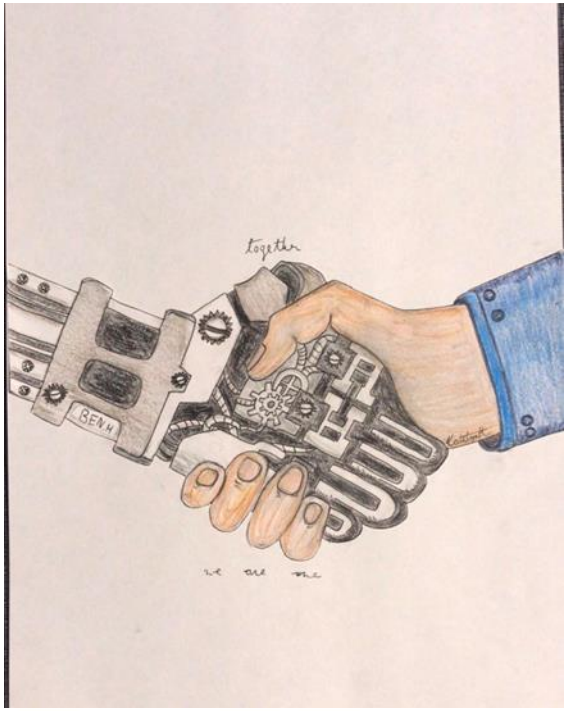
Hou, M., Ho, G., & Dunwoody, D. (2021), Special Issue on "Human-Machine Teaming in Military Contexts", Journal of Human-Intelligent Systems Integration.

Outline

- DRDC Toronto Research Centre (Who)
- Human-Autonomy Symbiosis Technology (WHAT)
- Interaction Issues with AI/Autonomy (Why)
- Interaction-Centered Design for Optimization (How)
- Application Examples of Interaction-Centered Design Approach (Where and When)

HAS Interaction Issues

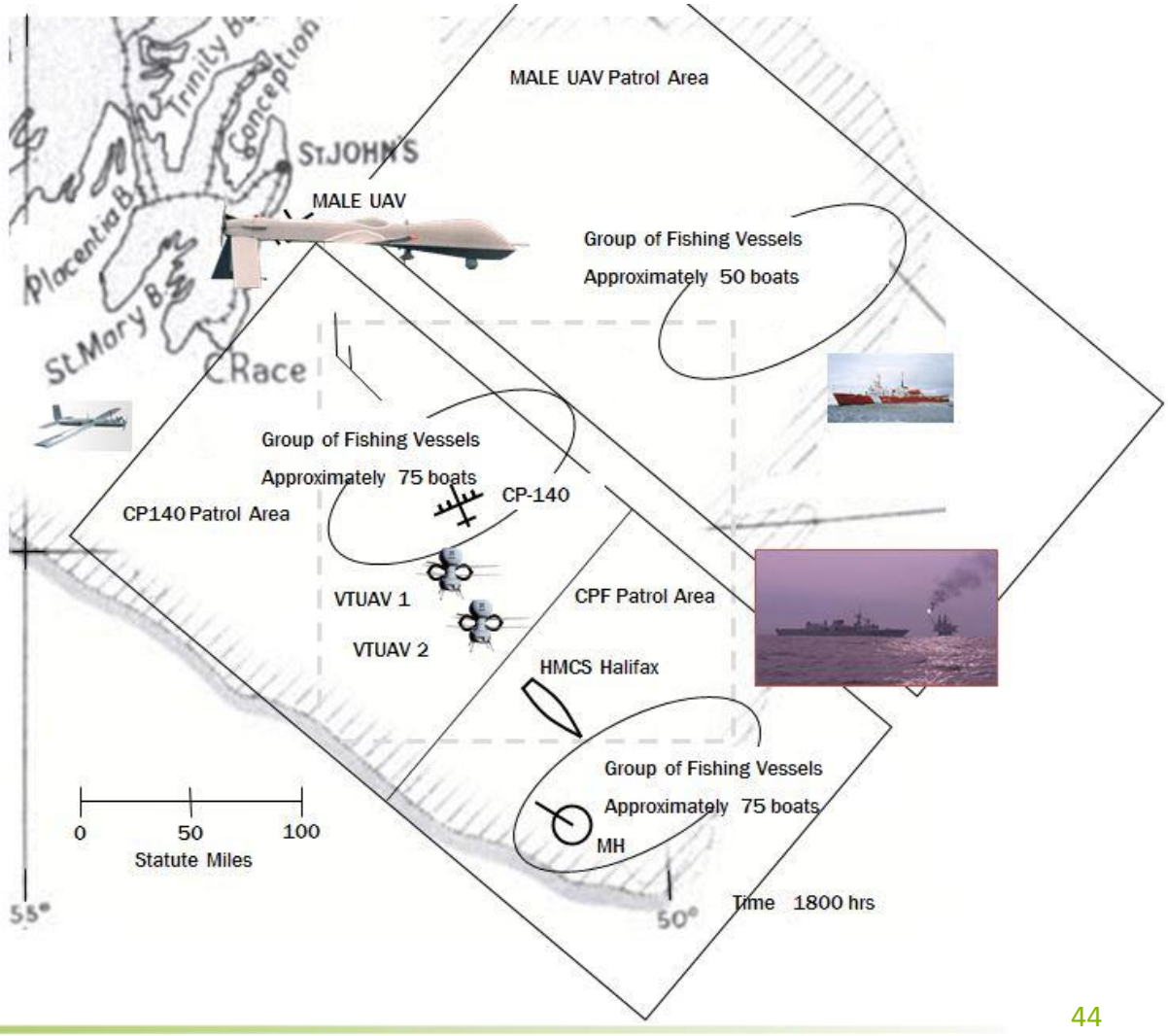
- Transparency
- Cognitive load
- Trust
- Accountability
- Legal and Ethical Aspects
- Policy and Regulations



Atlantic Littoral UAV Exercise

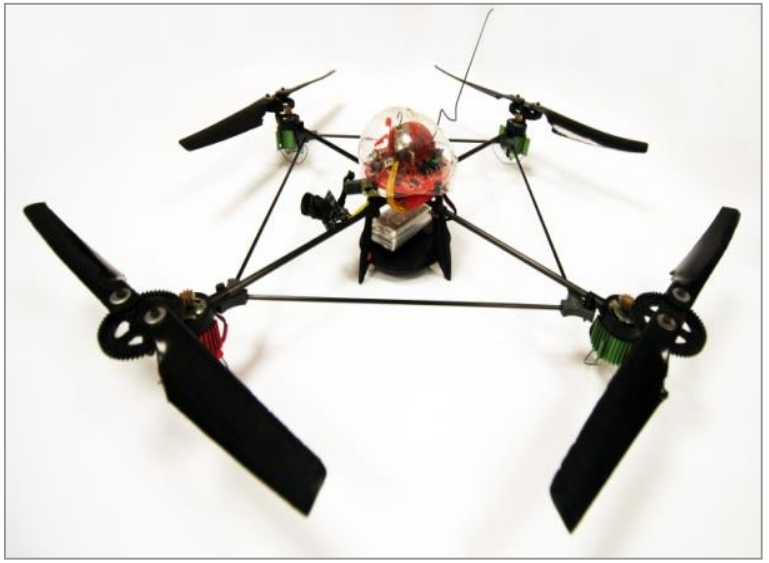
Multi-UAVs Searching for Terrorist Vessels

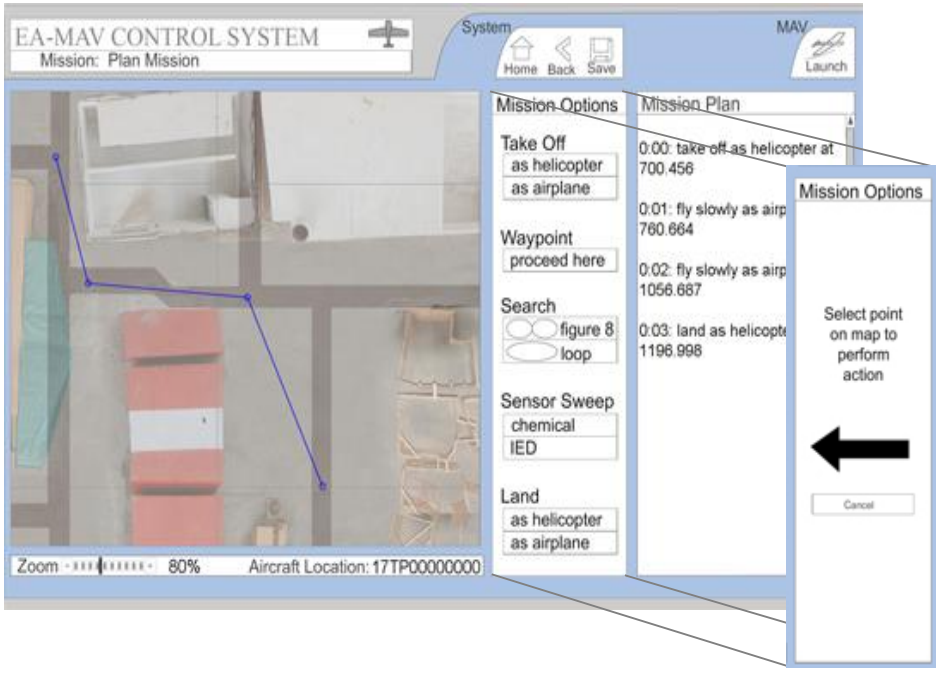
RCAF Maritime Patrol Aircraft CP140



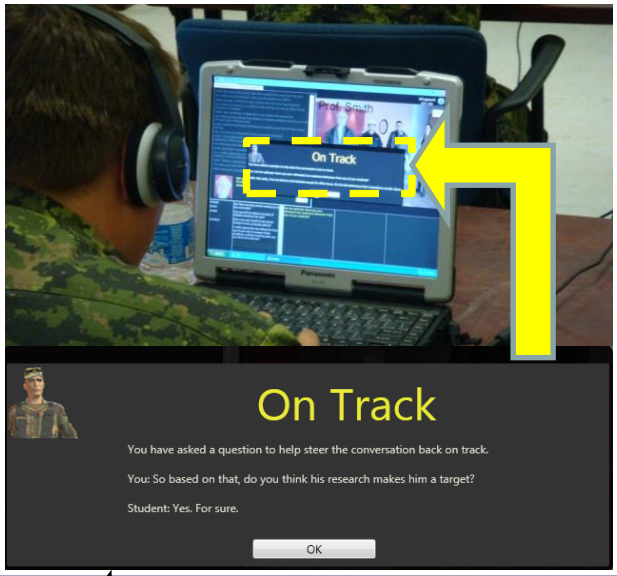
Information Overload

Soldier-Robot Teaming

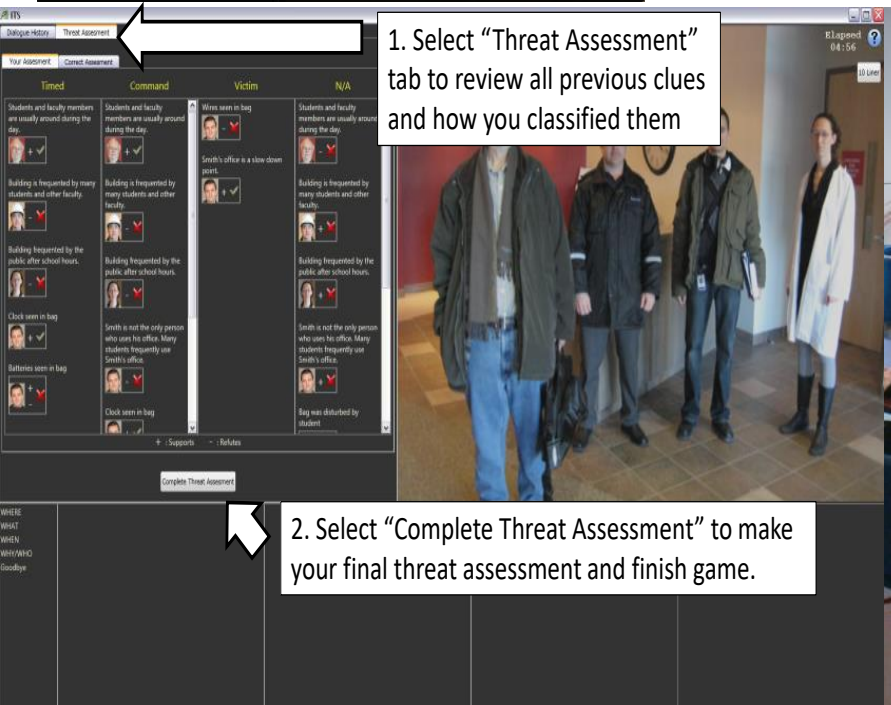




1st Canadian Intelligent Tutoring System (ITS)



- Students can question virtual witnesses and receive real-time and adaptive instructions based on their response and performance to learning context in Improvised Explosive Device (IED) disposal scenario.
- An integrated suite of ITS in Canadian Armed Forces (CAF) Counter-IED training course with improved efficiency, effectiveness (94%), and reduced cost.
- Patent application filed in both Canada and US.



Challenges of Weapon Engagement

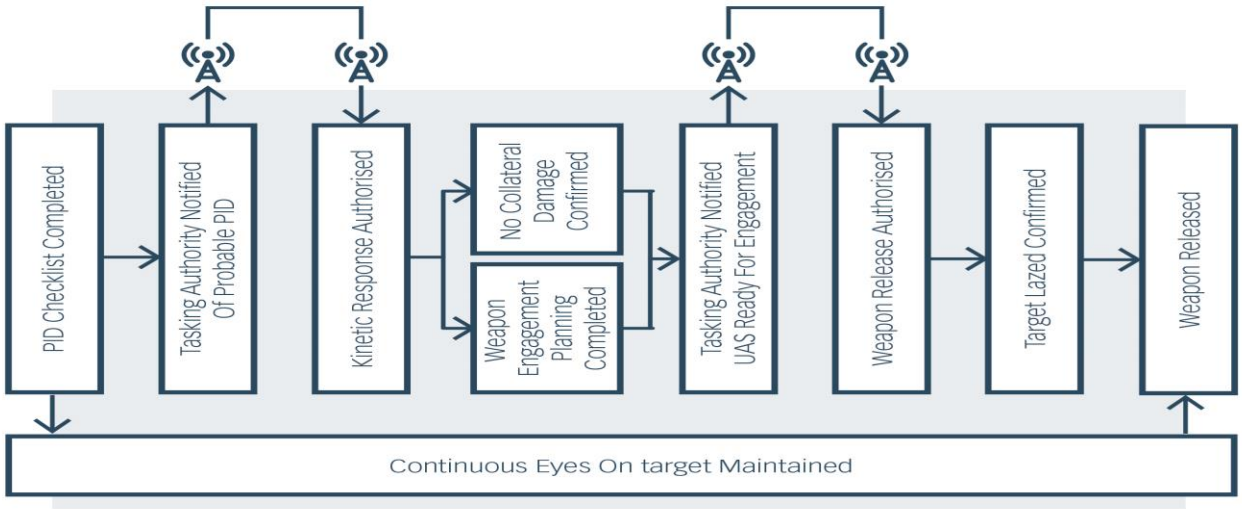
1. Lack of doctrinal knowledge on Rules of Engagement (ROEs), international Laws of Armed Conflict (LACs), and policies or regulations

2. Complex, lengthy, and error-prone target engagement processes significantly contribute to:

a. **loss of SA at individual and crew levels for both collocated and distributed teams;**

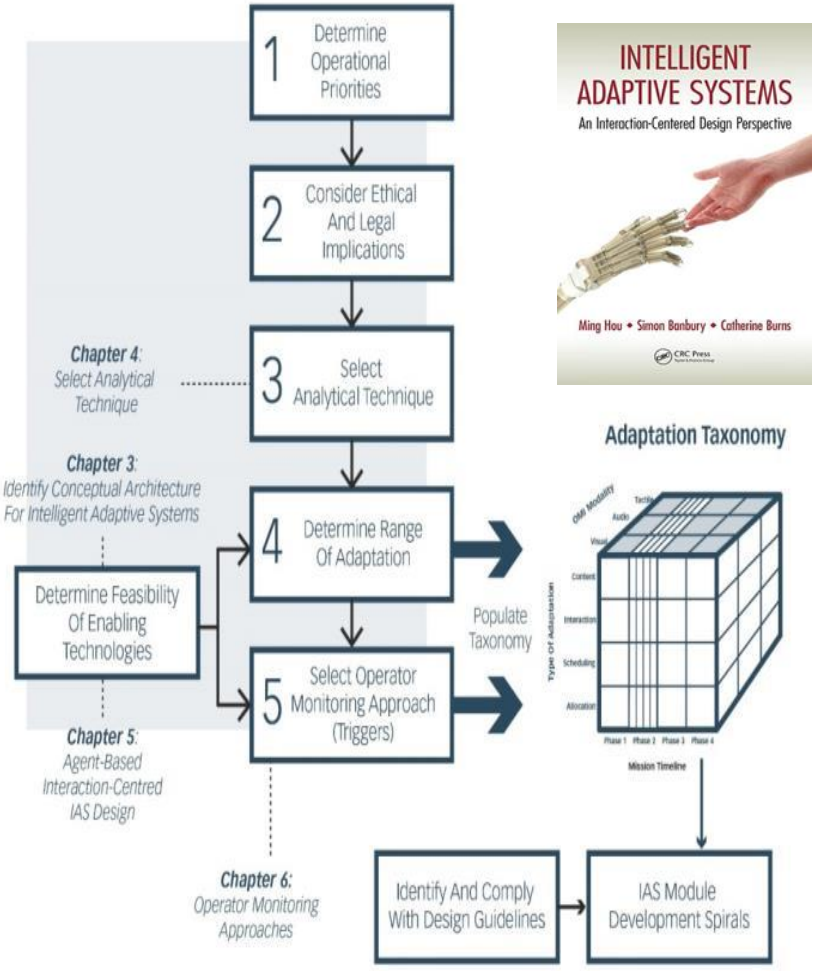
b. **Distrust of machine/technology partner (e.g., Autonomy/AI)**

c. **Potential mission failure**



Optimized Interaction for Trusted Human-Autonomy Collaborative Partnership

- The focus was on transparency for decision-making (intelligent and adaptive) with increased situational awareness and reduced workload to optimize transition of human interactions (with AI-enabled decision aid) between "on-the-loop" and "in-the-loop".



Authority Pathway for Weapon Engagement (APWE)

- An AI agent-based HAS with a weapon engagement stateboard for increased process transparency
- Ensures tasking authorities to follow ROEs, LACs, and target engagement procedures (legal, ethical, accountability, and trust).
- Automatically and dynamically updates the status of each step required to release a weapon, based on intelligence inputs from UAS crew, external authorities, and AI agents.
- Adapts its interface to UAS crew and other external users based on different ROEs and LACs information and communication requirements

Authority Pathway
WEAPONS ENGAGEMENT

Evaluation on Operational Effectiveness, Interoperability, & Military Utility

TTCP Autonomy Strategic Challenge
Autonomous Warrior 2018 Exercise

Operators' Feedback

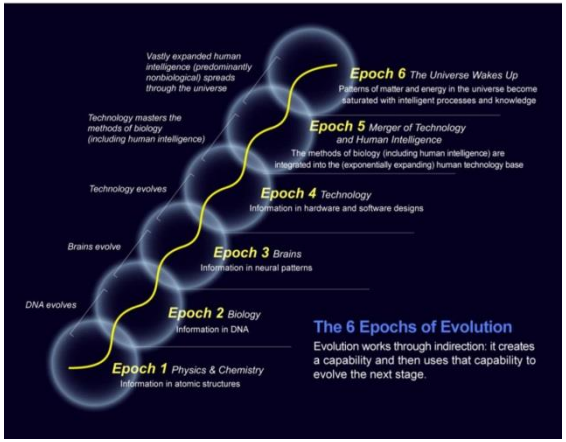
1. "… brilliant idea (to visualize engagement status for transparency) … "
2. "… massive fan; logical and great visibility of engagement status."
3. "… intuitive use for tasking authority to interact (with CDE and AI agents during target engagement process) …"
4. "… supports effective human-autonomy partnerships (teaming) …"

5. " … the most trustworthy of the whole thing (integrated C2 systems) because the increased SA and reduced workload (and potential human error) …"

Take Away

Interaction-Centered Design (ICD) strategy and methodologies are needed to mitigate potential risks given difficulty in designing fully fail-proof Intelligent Adaptive Systems (IASs) in the new era of

Human-Autonomy Symbiosis Technology



Page 29 "**Singularity is near**"

"We are what we repeatedly do.

Excellence, then, is not an act but a habit."

- Aristotle

A habit to comply with Regulations and Standards, follow established Processes, and then enable Safety and Trust.

THANK YOU and QUESTIONS?