Lowell
Scientific
Enterprises

IARIA

Voice in the Head: Prospects for Discrete Real Time Social Analytics

Dennis J. Folds, Ph.D.
Human Social Analytics (HUSO)
March, 2025

dennis.folds@lowellscientific.com

Chief Scientist,
Lowell Scientific Enterprises

Ph.D., Engineering Psychology, Georgia Institute of Technology
Atlanta, Georgia, USA (1987)

Fellow, IARIA

Forty-five years experience in conducting sponsored research for a variety of military and commercial customers

Career highlights: 35 years as Research Scientist at the Georgia Tech Research Institute (GTRI), Atlanta, Georgia, USA
Including 5 years as GTRI Chief Scientist and 1 year as Associate Director for Health & Human Systems Research

Current affiliations include:
Systems Engineering Research Center (SERC), Stevens Institute of Technology
Virginia Polytechnic Institute and State University (Virginia Tech)
Stephenson Technology Corporation, Louisiana State University
Problem Solutions, LLC (Chief Human Systems Scientist)

Evolution of Technology

Power

- Earliest tools depended on human power
- Engineers improved ways to harness gravity, wind, and other natural sources
- Service animals allowed improved performance of some functions
- Engineers created and improved the tools of *human-machine-animal systems*

Information Processing

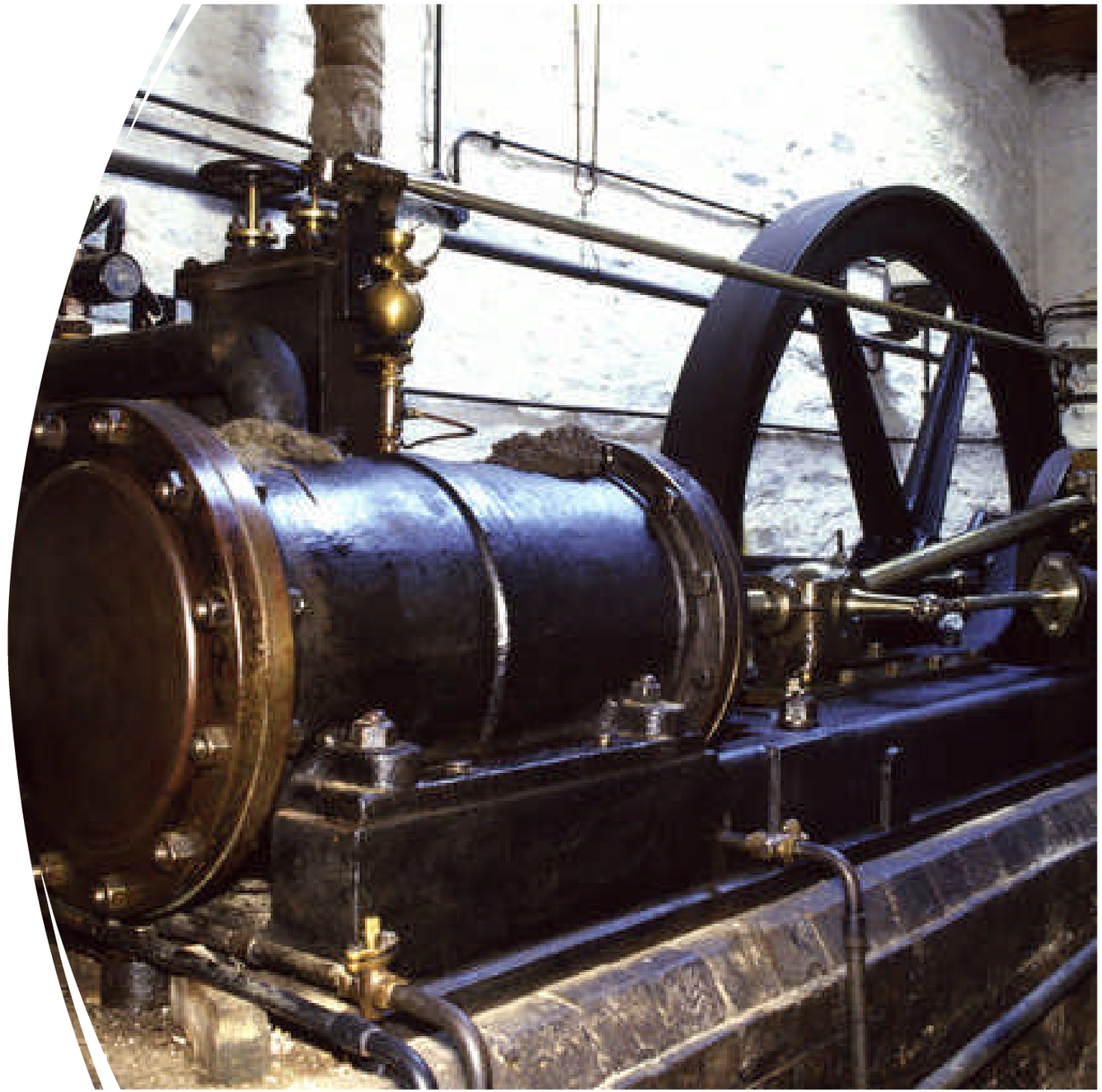
- Human intelligence performed all information processing in early systems
- Service animals provided some information processing capabilities, supervised by humans
- Improvements came from training and experience of humans and animals, and selective breeding of animals to improve certain traits.

Human Intelligence + Animal Power + (Relatively) Simple Machines

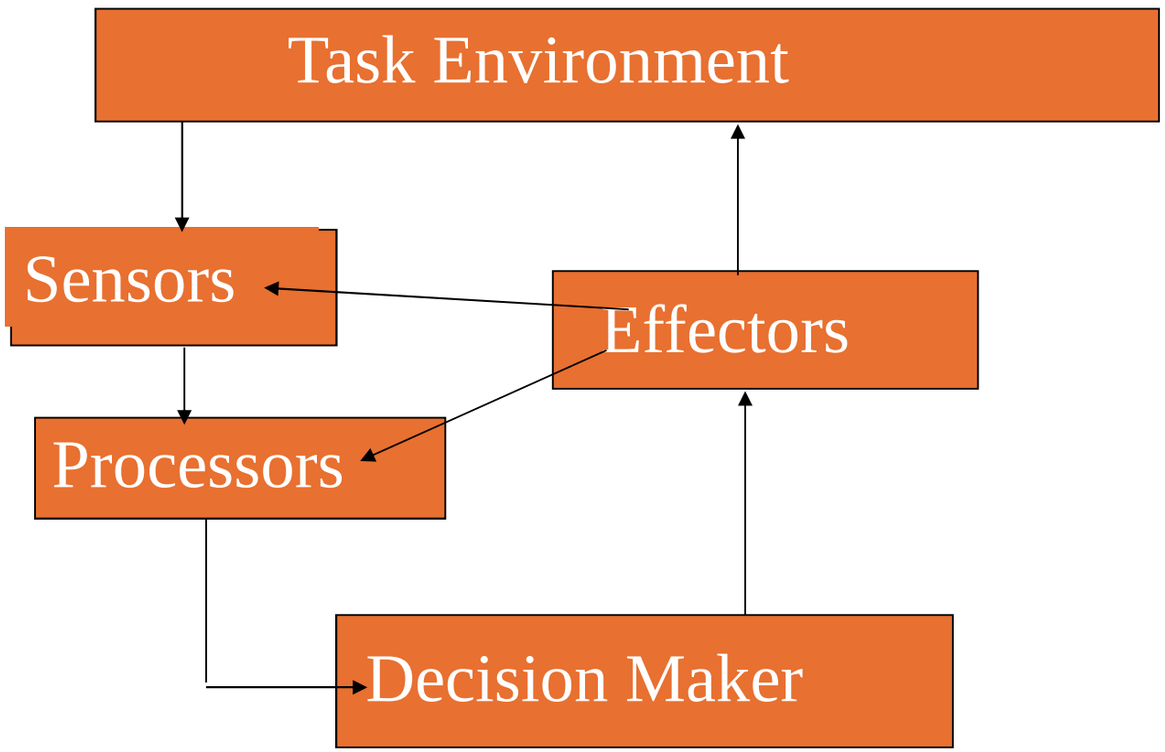
Human Intelligence +
Specialized Animal Skills
+ (More Complex)
Machines

The Age of Machines

- Shift to engineered sources of power
- Diminished reliance on animal power or animal intelligence
- Gradual development of capabilities machine-based information processing

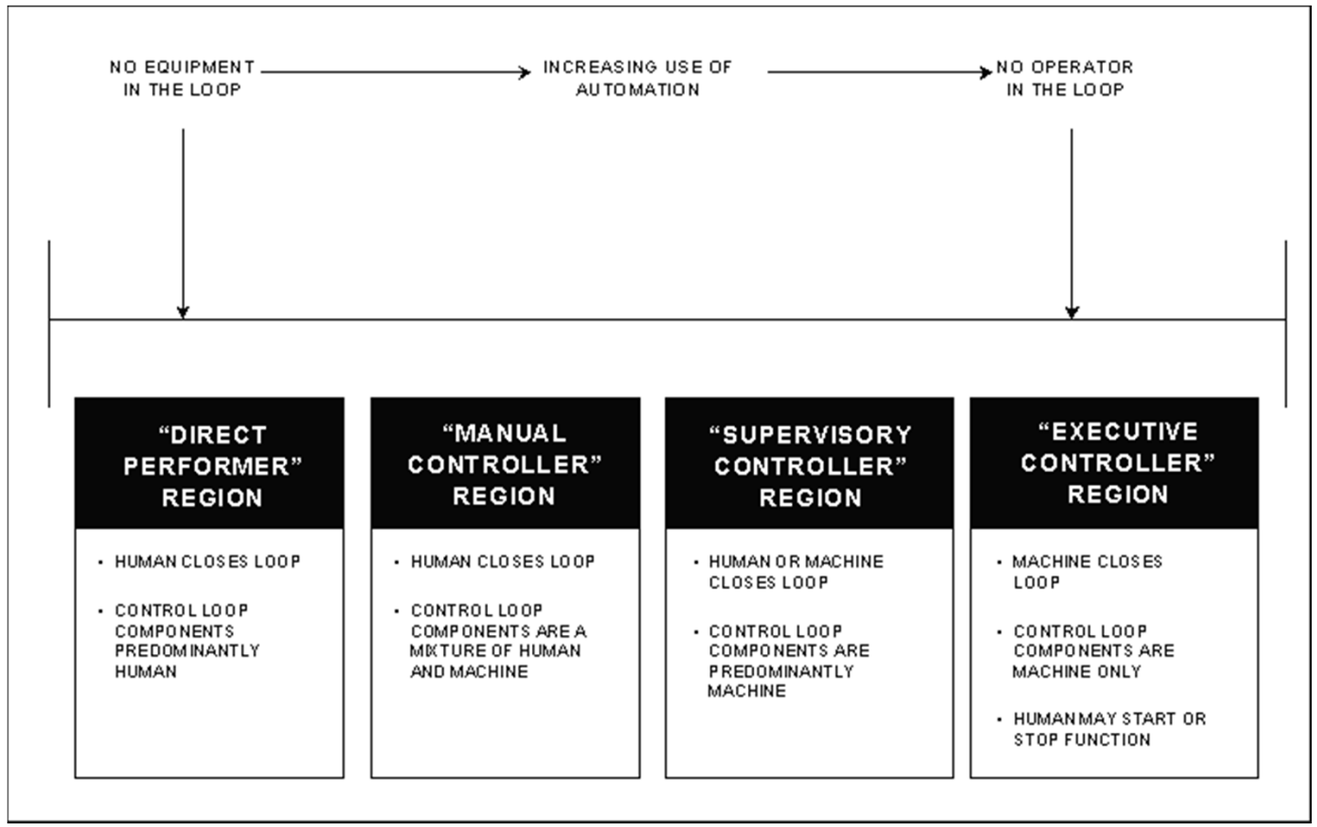


The Information Processing Loop



Human-Machine System

Continuum of Human Roles



Spiraling Complexity

- As machine capabilities mature, engineers put them together in new ways to create new functionality
  - Things that “just work” get used in new Direct Performer functions
  - Human uses a telephone to talk to someone else
- Combinations of things that rely on human decision making for coordinating their activity get used in new Manual Controller functions
- New machine-based decision making capabilities are implemented as Supervisory Controller functions
  - Until they get good enough for Executive Control

Trajectory

- Improves conditions for individuals, groups, and societies
  - Improved health and wellbeing
    - Nutrition, lower disease burden, improved security
  - Extended lifespan and life expectancy
    - Improved infant mortality rates
    - Increase in disability-adjusted life years (DALY)
  - Increased standard of living
    - Access to, and consumption of, goods and services
- Increase in leisure time
- Particular benefits for lower socio-economic status (SES) people

Challenge

- Create assistive AI/ML technology that can discretely deliver information that will:
  - Improve everyday life
  - Increase subjective well being
  - Compensate for functional limitations
- Use AI/ML and HUSO to support these functions:
  - Supporting education and life-long learning
  - Improving health and fitness
  - Aiding in job/task performance
  - Maintaining independent living across the lifespan
  - Facilitating treatment and management of acute and chronic conditions

Conversational Assistance

- Assisting, coaching, and mentoring generally involve verbal communications
  - Prompts
  - Explanation and elaboration
  - Interactive queries
- Private settings: normal conversational style
- Public settings: discrete dialog, generally imperceptible to others nearby

Strategic Goal: Voice in the Head (VITH)

- Leave the ear canal open, not occluded by headphones
  - Normal perception of ambient sounds and voices
  - Especially important for users who are blind
  - Also important in threatening environments
- Leverage bone conduction technology
- Mobile devices allow integration into everyday experiences
  - AR glasses / goggles & bone conduction (BC) headphones
    - Possibly combined into single device
  - Locally controlled by personal smartphone (etc.)
  - Augmented by other wearables (e.g., smartwatch)

Psychophysics – Research Agenda

- Equal loudness contours for BC transmission
  - Different mounting points have different properties
  - Best mounting points may not be compatible with integration into glasses
- Speech comprehension at near-threshold presentation levels
  - Could affect word choice as some phonemes are conducted more effectively (e.g., some unvoiced consonants have poor conduction)
- Need algorithms for adaptive presentation in dynamic ambient conditions
  - Detection vs. comprehension in nearby listeners

Application Examples

- Experiential Learning Maven in Real World Conditions
- Social and Cognitive Assistant for People with Functional Limitations
- Cognitive Load Manager in High Stress Situations

Experiential Learning: Language

Context: User is acquiring and developing command of a foreign language

- Learning includes interaction in virtual reality (VR) with gradual acquisition of vocabulary, grammar, diction, prosody, and idiom, guided by a language maven
- **VITH Support:** In real world conditions, the maven monitors ambient conversations and provides cues, prompts, and suggestions via discrete VITH channel
  - Primary goal is to promote learning and increased fluency in the new language

Note that this is different than simply having an AI-based translator

Experiential Learning: Place and Navigation

Context: User is learning about a new place (e.g., intended tourist destination)

- Learning includes exploration in VR, guided by a maven
- Acquiring knowledge about attractions, and navigation in the destination

- **VITH Support:** In real world conditions (at the destination), the maven provides prompts and other assistance via VITH channel
  - Assists in walking, driving, and public transportation
  - Provides reminders about significant cultural sites, etc.

Note this is different than a tour guide or turn-by-turn route guidance

Functional Limitations: Autism Spectrum

Context: A person with ASD struggles with reading social cues.

- **VITH Support:** AI provides real-time **conversation coaching**, such as:
  - Recognizing facial expressions and tone.
  - Suggesting appropriate responses or exit strategies.

- **Example:** *“John just crossed his arms and leaned back—he may be losing interest in the topic. Consider asking him a question about his interests.”*

Functional Limitations: Memory Loss

- **Context:** A person with mild cognitive impairment forgets details about people they meet.
- **VITH Support:** AI functions as an **augmented memory**, discreetly providing names, past interactions, and personal details.
- **Example:** *“This is Dr. Smith—you met at last month’s conference. He has two kids and enjoys sailing.”*

Functional Limitations: Visual Impairment

- **Context:** A blind individual moves through a complex environment.
- **VITH Support:** AI integrates **computer vision and environmental data** to provide real-time navigation.
- **Example:** *“You’re approaching an intersection. The pedestrian light is red—wait to cross.”*

Functional Limitations: Speech Impairment

- **Context:** A person with a degenerative condition struggles with speech fluency.
- **VITH Support:** AI **predicts and suggests words**, subtly prompting the user to maintain fluid conversation.
- **Example:** *“You might say: ‘I’d love to hear more about that.’”*

Cognitive Load Manager: Medical Emergency

- **Context:** A paramedic performs under pressure.
- **VITH Support:** AI provides **situational reminders, alerts, and decision support** while keeping cognitive overload low.
- **Example:** *“The patient’s O2 level just dropped—consider supplemental oxygen.”*

Cognitive Load Manager: Negotiations

- **Context:** A diplomat at a tense negotiation session.
- **VITH Support:** The AI analyzes speech patterns and body language to detect deception or shifts in tone.
- **Example:** The AI quietly alerts, *“The minister’s tone suggests reluctance. Try a softer approach.”*

Cognitive Load Manager: Tactical Support

- **Context:** A field operative requires critical information without looking at a screen.
- **VITH Support:** AI provides **silent, real-time mission updates, enemy positioning, or extraction routes**.
- **Example:** *“Thermal scan from UAV detects movement ahead. Approach with caution.”*

Ethical Considerations and Design Issues

- **Privacy & Trust**
  - Ensure **AI only activates when needed** to avoid unwanted intrusions.
  - User retains control over AI prompts.
- **Adaptive Learning**
  - AI **personalizes guidance** based on **user preferences, cognitive load, and social context**.
- **Latency & Edge Computing**
  - Real-time processing must be **fast and reliable**, possibly leveraging **edge AI** for offline scenarios.
- **Multimodal Integration**
  - VITH AI could **combine speech with haptic feedback** (e.g., subtle vibrations on wrist to indicate urgency).

Ongoing and Planned Research Program

- Psychophysical experiments with BC technologies
  - Dissertation and other graduate student research
  - Focus on perceptual effects

- First VITH Prototype
  - Military officer training program
    - Physical Fitness
    - Academic Course Requirements
    - Focus on content

Conclusion

- The trajectory of technology development led to great improvements in human conditions
- Potential improvements are identified from using AI/ML supplemented by HUSO
- Discrete delivery of conversational prompts through VITH channels
- Significant research challenges should be tackled

Questions?

- For more information, contact

- Dr. Dennis Folds
- dennis.folds@lowellscientific.com