Take-Aways from WS#1

Kevin Sullivan
University of Virginia

Overview

- Need for radical improvements in system performance, requiring audacious new research aims, methods, and communities
  - (see written comments in separate document)

- A picture of emergent areas of interest at Workshop #1

A View of Workshop #1

- People in the loop - at all scales
- Models, mechanisms, and methods
- From smart to learning systems
- Design, architecture, implementation, infrastructure
- Challenges in sensing, data, and analytics
- Trust in and trustworthiness of cyber-social learning systems
- Simulation
- Bias, ethics, empathy
- Research methods
- Notional research projects
- Overall gestalt of workshop #1

People in the loop

- New methods of studying people and social phenomena at scale
  - From individual physiological and psychic states to mass social movements
- Affective computing, sensing everything from physiology to motivation
- Models for interventions at many scales: personal, dyadic, team, market
  - E.g., social-cognitive theory; role of failure, credit incentives, responsibility, accountability, blame, enforcement
  - E.g., precision psychology: using deep machine learning to learn what aspects of particular therapies work for specific people, contexts, and problems
- Scaling (e.g., for suicide prevention) is a tremendous challenge
- Governance. Oversight. Co-production.
- Social engineering: Who decides social good and thus optimization goals?

Models, methods, mechanisms

- What models, mechanisms, and methods to use in a given situation?
  - E.g., sense-analyze-respond vs. probe-sense-respond vs. act-sense-respond
- How will we challenge, reflect on, validate, learn limits of, and evolve models?
- E.g., models of stakeholders, relationships, how decisions impact them
- How to understanding stakeholders and their needs at all levels
- How do stakeholder groups fit together, including governance and regulators?
- What are characteristics of stakeholders; how do they constrain solutions?

From smart to learning systems

- Not just smart and connected (static control) but learning (evolution of knowledge, practice, and performance)
  - Single- versus double-loop learning.
- What does it mean for a socio-technical system to be “built to learn?”
- Learn about past but also project future, then check to see how predictions worked out and self-calibrate (learn to predict): automated agents to do this
- New kinds of training. The system being trained, the humans being trained: where in the system, and what are the forms of conditioning and training?
- Reinforcement learning, e.g., for training of autonomous system components
- Inverse reinforcement learning to ascertain objectives from observations of good practice, e.g., to understand stakeholder objectives

Design, architecture, implementation, infrastructure

- Design, develop, and deploy of an emerging class of cross-platform, service-integrated, technology products to enhance performance and/or create a platform for economic development in cities and communities
- How do we design, prototype, develop, operate, evolve CSLS?
- Architecture emphasizing learning loops within service delivery systems
- Architecture emphasizing independent oversight & related roles
- What are the structures and dynamics, layers, modules, stakeholders, etc.
- *Designing for emergent behavior* in large-scale CSLSs

Sensing, data, analysis

- Data, information, decision-making: inputs, rationale, provenance
- Complete and consistent data collection
- Filtering and qualification of information, disinformation, misinformation
- Collecting right data is often hard, so we collect what is easy to collect
- Measuring what’s easy to measure can produce bad optimizations; how do we get smart about designing proxy measures?
- How will we know what's unexpected, new stabilities and instabilities, re-tune and resettle in face of uncertainties, instabilities, unintended consequences.
- Sense-making models, where data precede framework for understanding
- Visualization and understanding of data and results
- Data are backwards looking, but do not reveal what could or should be

Trust and trustworthiness

- Safety-preserving simulation, predictive modeling
- Trustworthiness, architected: e.g., role of governance in fostering trust.
- Design and engineering of participant experience to foster trust
- Factors: explanation and transparency; perspicuity of metaphors; experience with system; governance; business, legal, regulatory factors (permits, etc.)
- Interaction of learning, thus adaptation, with trust and trustworthiness
- Challenges: rapid adaptation, diffused governance and decision making; data quality and training bias; ability to measure and characterize uncertainties
- Architecture and oversight: proposal (e.g., permits), continuous (e.g., inspections of factories), and retrospective (e.g., accident analysis)
- Validating system properties:, safety, privacy, compliance, etc.
- Assurance methods for CSLS

Simulation

- How might simulation augment observation or experimentation
- When they’re impractical in the real world
- As a way of predicting what might be about to happen
- To project intended and unintended consequences of interventions

Bias, ethics, empathy

- Forms and sources of bias
  - Sampling
  - Statistical tradeoffs (bias vs. variance; overfitting, accuracy)
  - Ethically adverse (discriminatory) bias
- Reduce socially adverse bias in CSLS
- Role of policy and ethics in this regard

Research methods

- We are using old techniques, statistical techniques, modelling techniques, research techniques, and methods
- What is the role of testbeds and simulation environments in CSLS research
- Experimental methods and enabling policy, technology, people, etc
    - E.g., is a hospital system willing to open itself up as an experimental testbed

A few notional research topics

- Languages for modeling, defining, automating CSLS
- Machine learning methods that account for relational context
- Dealing with non-IID characteristics of data from complex social environments
- Programming not in objects, but in lines coming together (Ingold)
- Predictive models operating for individuals, dyads, communities, society
- Interventions: what-if analysis (prediction)
- Interactions: how do independently operating CSLSs compose when they act on shared aspects of the environment (initiatives can reinforce and conflict)
- ML: causal models, social inference, non-IID data
- Social science: psych + social models + big data + ethnographic (small data)
- New systems engineering/science for how people, data, systems interact
- Ethics of AI, autonomy , and in system design and operation
- Deeper Integration of big data into (revolutionary) social sciences

Overall gestalt from Workshop #1

- “... there was a growing sense of enthusiasm and discussion within the room that it was starting to feel like we're not in Kansas anymore...once you start putting all of these pieces into the same design constellation we were looking at something that to us felt pretty different.”
- There are common abstractions that go across a broad number of the most compelling cyber-social systems domains
- Need for better mental models of CSLS, what we want CSLS to accomplish, and why, in more precise terms, e.g., what are societal drivers and risks and how will CSLS address them
- Need for better, more useful scenarios, case studies, success/failure cases