Class Reading Assignment 5: Problems of Social Computing Systems

Course: CS 6474 / CS 4803 Social Computing

Grade: 7% of overall course grade (70 points total)

Due Date: the last class of instructional period on April 27 | *Earlier submissions encouraged | Earlier submissions will be graded sooner and feedback provided ~2-3 weeks from the time of submission*
What to hand in: Submit as a single PDF on Canvas

Formatting Guidelines:
• Length: approximately 5-6 pages single-spaced, 1-inch margins
• Font: at least 11pt, readable serif or sans-serif

Grading Emphasis:

- Specific use of evidence from papers (not generic summaries)
- Clear articulation of mechanisms and causal reasoning
- Thoughtful critique of measurement and methods
- Ability to identify tensions or limitations across studies
- Concrete and feasible design ideas (where applicable)
- Clarity and organization of writing

Collaboration Policy:
This is an individual assignment. You may discuss high-level ideas with classmates, but all submitted work must be your own. You may not share written responses.

AI Use Policy:
You may use AI-based tools only for proofreading or improving clarity. You **may not** use AI tools to generate ideas, arguments, or structure. Responses should reflect your own reasoning and engagement with the readings and lectures.

This assignment builds directly on Week 9-10 lectures and discussions on the problems of social computing systems, that spanned several topics like online abuse and harassment, mis- and disinformation, polarization, and informational bias.

Question 1: Mechanism Tracing in Studies

This question requires you to trace a specific mechanism across papers, rather than broadly compare domains.

Part A: Situational vs. Structural Drivers (20 pts)
Refer to Cheng et al. (trolling), Vosoughi et al. (false news diffusion), and Nyhan et al. (like-minded exposure). Pick *any one* of these three papers, to answer the following questions:

i. Identify one clearly defined mechanism in the paper you choose (be precise), such as:
   - "negative mood increases trolling probability"
   - "novelty increases retweet likelihood"
   - "algorithmic exposure reduces cross-cutting content"

ii. For the identified mechanism:
   - What is the unit of analysis? (user, content, network, etc.)
   - What is the causal pathway (input → process → outcome)?
   - What evidence supports it (experiment, model, descriptive pattern)?

Part B: Cross-Paper Tension (10 pts)
Identify one pair of findings that are in tension across the paper readings in Weeks 9 and 10. An example of a valid tension could be: "*Exposure to opposing views increases polarization (Bail et al.) versus reducing like-minded exposure has no attitudinal effect (Nyhan et al.)*." Explain:

i. What exactly conflicts
ii. Why both findings might still be valid (depending upon context, platform design, population, outcome differences etc.)

Question 2: Interrogating Measurement and Operationalization

This question asks you to engage with how constructs are measured, not just what is found in the papers.

Part A: Construct to Measurement Mapping (15 pts)
Choose *one* construct from any of the papers assigned during Weeks 9-10 (e.g., "trolling", "misinformation", "polarization"). Answer the following questions:

i. How is the construct operationalized? Below are two examples:
   - "Flagged posts as proxy for trolling"
   - "Retweet cascades as diffusion"

ii. What are two limitations of this operationalization?

iii. What type of measurement error might result? Specifically focus on: a) False positives/negatives; b) Construct validity issues; c) Bias in labeling or sampling

Part B: Measurement Consequences (10 pts)

For the same chosen construct in Part A of Question 2:

i. How might the main result change if the construct were measured differently in the corresponding paper?
ii. What alternative measurement would you use? Would the effect likely increase, decrease, or disappear? Why?

Part C: Designing a Better Measurement Strategy (15 pts)

Propose an improved way to measure the chosen construct from Parts A-B of Question 2. Your answer must include: i) Data needed; ii) Method (e.g., annotation, behavioral logs, multimodal signals); iii) One tradeoff (e.g., scalability vs. validity).