

NETS 213:

Crowdsourcing and

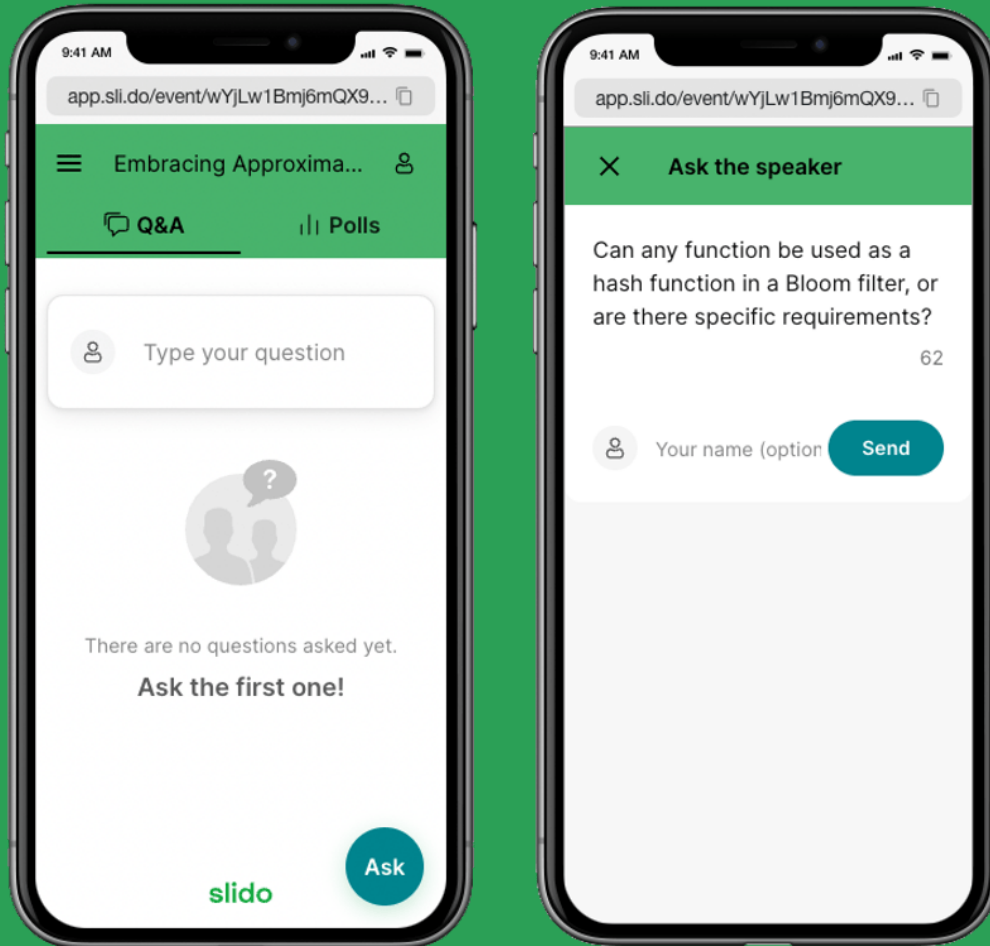
Human Computation

Fall 2025

Instructor: Jérémie Lumbroso

Ask Questions with Sli.do

1. Ask new questions (**anonymously** or **not**)
2. Upvote questions
3. Vote in polls



Administrative

Course Context

This is a BETA course iteration

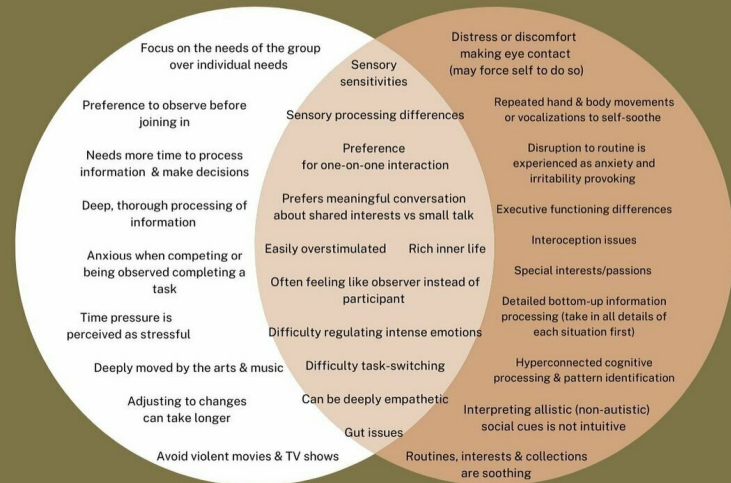
- Taking over from Chris Callison-Burch
- Incorporating major developments in AI since 2020
- Your patience and feedback are appreciated

Self-presentation



@STEPHANIEGARDNERWRIGHT

Highly Sensitive or Autistic?



Highly Sensitive Traits

Autistic Traits

Course Overview

What We'll Cover:

- Crowdsourcing platforms and mechanics
- Programming for human computation
- Economics and incentive design
- Quality control and aggregation
- Machine learning integration
- Ethical implications
- Building crowd-powered systems

Course Requirements

Weekly Work

- Programming assignments (Python)
- Platform analysis
- Academic paper reviews
- Crowd work experience

Major Projects

- Replicate one research paper
- Design original crowd system
- Final project (groups of 3-4)

Course Details

Grading formula

- Weekly assignments (50%)
- Final project (45%)
- Participation (5%)

Necessary resources

- Spend \$20-\$40 on crowdsourcing credits
- Assumed/expected to have a generative AI license
- Will provide all readings

Websites and tools

- We will be using Gradescope for assignments
- Ed Discussion for class forum
- The course website (to be launched) is:
 - **<https://crowdsourcing.wiki>**

Who Should Take This Course?

Ideal for:

- CS students interested in HCI/AI
- Data science students wanting human-in-the-loop experience
- Business students exploring platform economics
- Social science students studying digital labor
- Anyone interested in the future of work
- Anyone interested in how data for AI is made
- Anyone interested in the future of decision-making

Prerequisites:

- Basic programming (Python helpful)
- Curiosity about human-computer interaction
- Critical thinking about technology's impact

Introduction

The Hidden Humans in "AI"

Amazon's Just Walk Out Technology

- Marketed as "AI-powered checkout"
- Reality: 1,000+ workers in India reviewing transactions
- 70% of transactions required human review (as of 2024)

// "AI" often means "Actually Indians" - or more broadly, invisible human labor

Does Amazon's cashless Just Walk Out technology rely on 1,000 workers in India?



Betty Lin-Fisher
USA TODAY

Published 5:23 p.m. ET April 4, 2024 | Updated 4:28 p.m. ET April 5, 2024



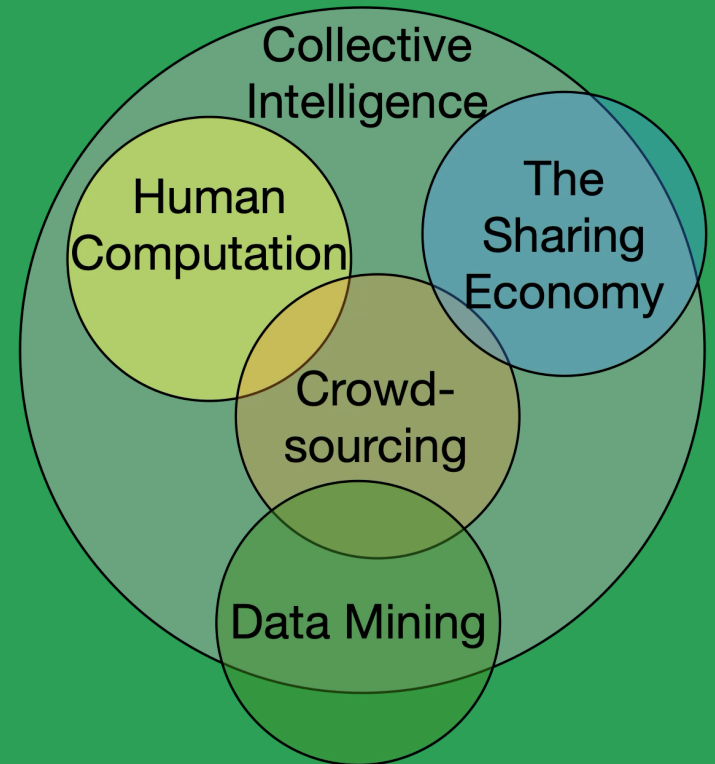
Key Concepts and Definitions

Collective Intelligence: Enhanced capacity from groups working together to solve problems none could achieve alone

Human Computation: Humans performing tasks computers cannot efficiently solve, integrated into computational processes

Crowdsourcing: Obtaining services/content from large groups via open calls on digital platforms

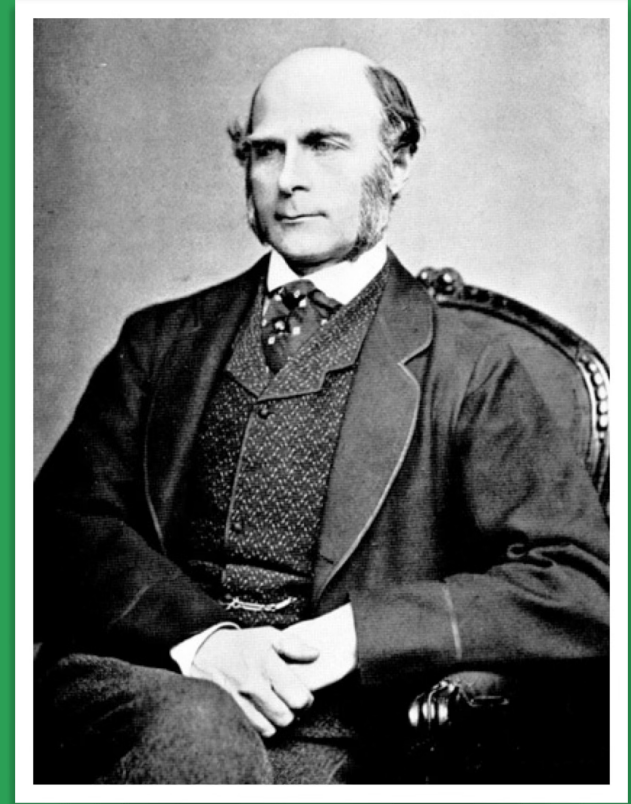
Platform/Sharing Economy: Peer-to-peer sharing of underutilized resources via digital marketplaces



The Wisdom of Crowds

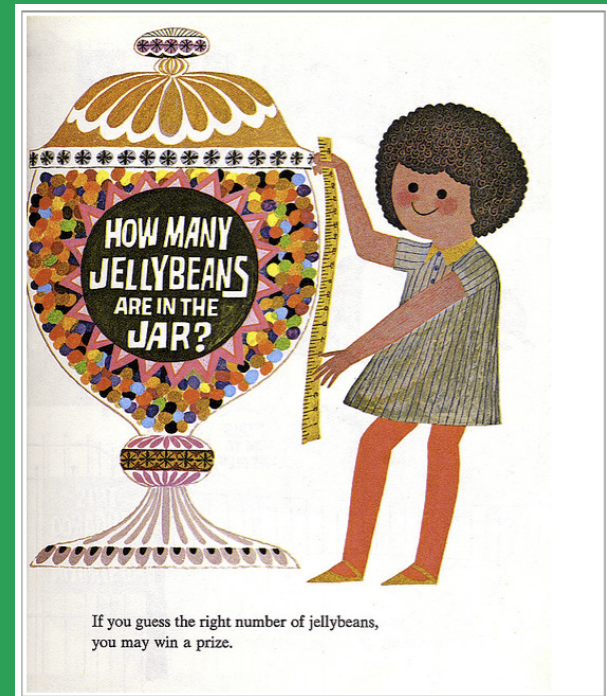
Classic Example: Francis Galton (1906)

- Plymouth livestock fair
- 800 people guessed ox's weight
- Individual guesses varied wildly
- **Median guess: 1,207 lbs**
- **Actual weight: 1,198 lbs**
- Crowd error: < 1%



Modern Wisdom: Jelly Bean Experiments

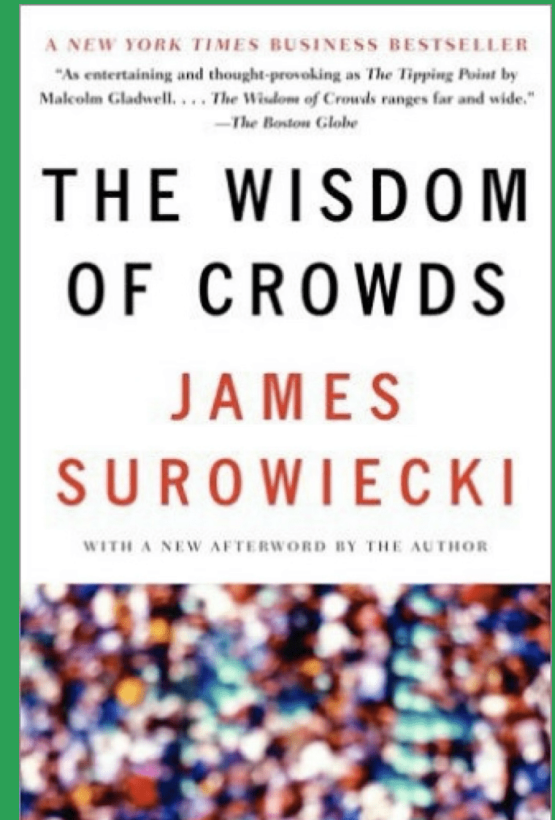
- Repeated studies show consistent patterns
- Average guess typically within 3% of actual
- Key insight: Errors cancel out when aggregated
- But only under specific conditions...



Requirements for Crowd Wisdom

(Surowiecki, 2004)

1. **Diversity of Opinion** - Each person has private information
2. **Independence** - Opinions aren't determined by others nearby
3. **Decentralization** - No one is dictating answers
4. **Aggregation** - Mechanism to combine private judgments



When Crowds Fail: Groupthink

Definition

A psychological phenomenon where desire for harmony results in irrational or dysfunctional decisions

Examples:

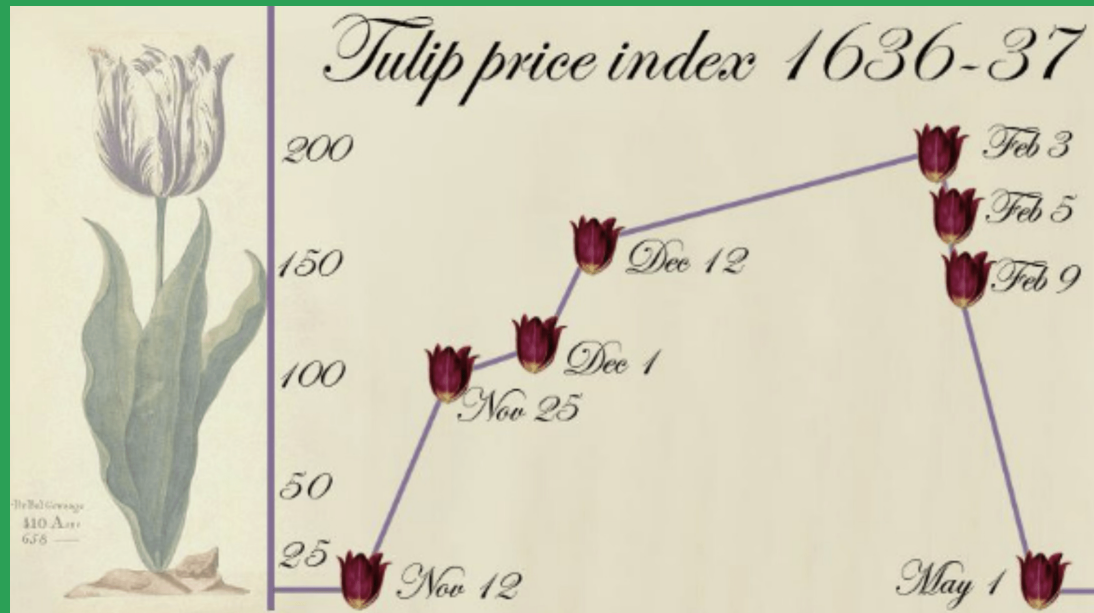
- Bay of Pigs invasion (1961)
- Challenger disaster (1986)
- 2008 Financial crisis
- Theranos board decisions



When Crowds Fail: Information Cascades

Historical Bubbles

- Tulip Mania (1636-1637) - Single bulb = 10 years' wages
- South Sea Bubble (1720)
- Dot-com Bubble (1995-2001)
- NFT Bubble (2021-2022) - Bored Apes from \$400K to \$40K
- Meme Stocks (2021) - GameStop, AMC



When Crowds Fail: Mob Justice

Boston Marathon Bombing (2013)

- Reddit "detectives" analyzed photos
- Falsely accused Sunil Tripathi (missing student)
- Family harassed during tragedy
- Actual bombers remained free longer
- Key lesson: Crowds lack accountability



Success Story: Wikipedia

The Numbers (2024)

- 60+ million articles
- 300+ languages
- 100,000+ active contributors
- 2 billion+ unique devices monthly
- More accurate than Britannica (Nature, 2005)

Key Success Factors:

- Clear policies and guidelines
- Transparent editing history
- Dispute resolution mechanisms
- Balance of openness and control



WIKIPEDIA
The Free Encyclopedia

Success Story: Crisis Response

Haiti Earthquake (2010)

- **Problem:** Responders didn't speak Kreyol, maps were outdated
- **Solution:** Ushahidi platform + global volunteers
- **Process:**
 - Haitians texted 4636 for help
 - Diaspora translated messages
 - Volunteers mapped locations
 - Responders received actionable intel
- **Result:** Thousands of lives saved

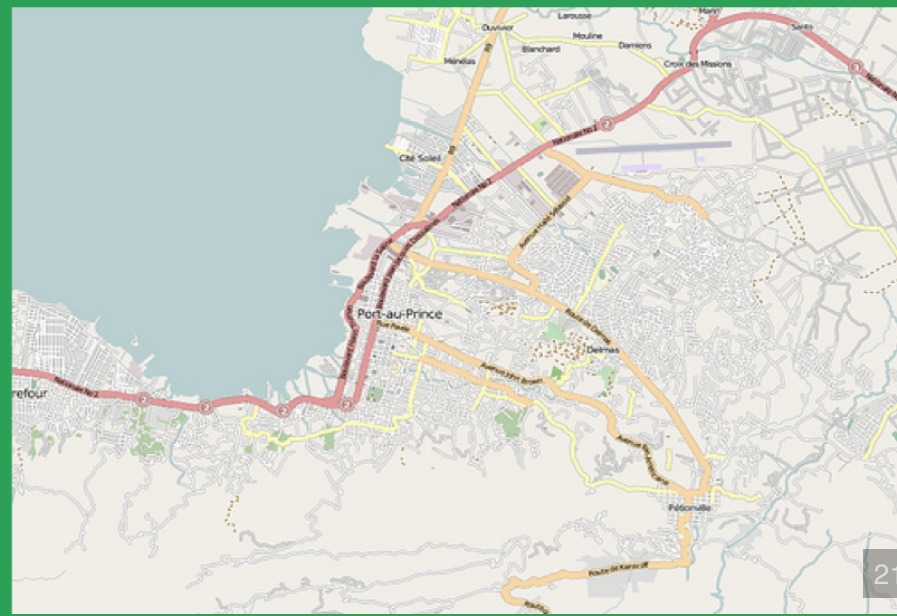
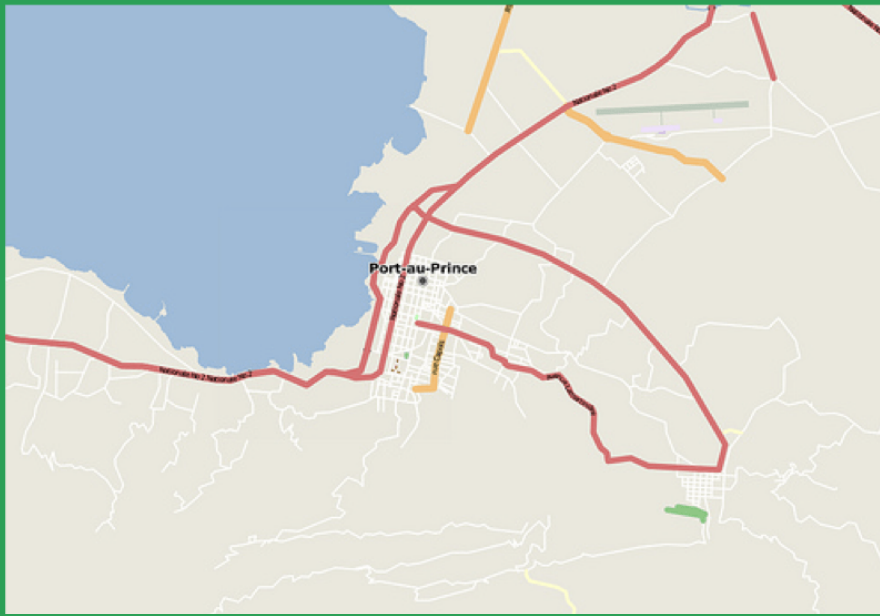


Success Story: OpenStreetMap Response

Before: Incomplete maps

After (48 hours): Detailed street-level data

- 600+ volunteer mappers
- Used satellite imagery
- Local knowledge via diaspora
- Now standard disaster response protocol



Modern Platform: Amazon Mechanical Turk

The Marketplace (2024)

- 500,000+ registered workers
- 100+ countries represented
- Median wage: ~\$2/hour (after unpaid time)
- Used by: Academia, AI companies, startups

The Hidden Truth:

Most "AI training" happens here

The AI-Human Labor Pipeline

Training ChatGPT/Claude/Gemini requires:

1. **Data Cleaning:** Humans filter toxic content
2. **Annotation:** Humans label training examples
3. **RLHF:** Humans rate AI responses
4. **Red Teaming:** Humans find failure modes
5. **Ongoing:** Humans handle edge cases

The Paradox:

"AI" that eliminates jobs... creates different invisible jobs

ARTIFICIAL INTELLIGENCE

The AI Was Fed Sloppy Code. It Turned Into Something Evil.

29 |

The new science of “emergent misalignment” explores how PG-13 training data — insecure code, superstitious numbers or even extreme-sports advice — can open the door to AI’s dark side.



Recent Success: Crowdsourced Fact-Checking

X/Twitter Community Notes (2021-present)

- Users collectively fact-check posts
- Requires consensus across political divides
- Algorithm weights "surprising agreement"
- Measurably reduces misinformation spread
- One of few successful anti-misinformation tools

Recent Success: Citizen Science at Scale

iNaturalist (2024 stats)

- 150+ million observations
- 3+ million users
- 430,000+ species documented
- Used in 5,000+ scientific papers
- Discovered 100+ new species
- Critical climate change data

Recent Failure: GameStop/WallStreetBets

January 2021

- Reddit crowd coordinated buying
- Stock price: \$17 → \$483 → \$40
- Some made millions, many lost savings
- Revealed market fragility
- Raised questions about market manipulation

Lesson: Crowd coordination ≠ crowd wisdom

The Generative AI Revolution

What Changed (2020-2024)

- Large Language Models (GPT, Claude, Gemini)
- Image Generation (DALL-E, Midjourney, Stable Diffusion)
- Code Generation (Copilot, Cursor)
- Video/Audio synthesis

The Hidden Truth:

**All trained on human-generated content All
require ongoing human feedback**

Human Computation in the AI Age

The New Pipeline:

Visible: "AI" products and services

Hidden:

- Data labelers in Kenya (\$2/hour)
- RLHF raters in Philippines (\$3/hour)
- Content moderators in India (\$1.50/hour)
- "Prompt engineers" in Venezuela (\$4/hour)

We haven't eliminated human work, we've made it invisible + better integrated

Crowdsourcing Platforms Today

Microtask Platforms

- Amazon MTurk, Clickworker, Microworkers
- Scale AI, Surge AI, Toloka

Freelance Platforms

- Upwork, Fiverr, 99designs

Innovation Platforms

- InnoCentive, Kaggle, TopCoder

Citizen Science

- Zooniverse, iNaturalist, eBird

...

- Zooniverse, iNaturalist, eBird

Applications in Research

Computer Vision

- ImageNet: 14M images, 1000+ object classes
- Enabled deep learning revolution
- Every photo labeled by humans
- Took 2.5 years, cost \$100K+

Natural Language Processing

- Every chatbot trained on human conversations
- Translation systems need bilingual validation
- Sentiment analysis requires cultural context

Applications in Business

Content Moderation

- Facebook: 15,000+ moderators
- YouTube: 10,000+ reviewers
- TikTok: 20,000+ globally
- Psychological toll documented

Data Entry & Verification

- Medical record digitization
- Receipt transcription
- Address verification
- Business card scanning

Dark Patterns: The Attention Economy

Your Data as Unpaid Labor

- Every Google search improves their algorithm
- Every LinkedIn like trains recommendation systems
- Every CAPTCHA solves digitizes books/trains AI
- Every YouTube view optimizes engagement

You are the product AND the worker

Dark Patterns: Algorithmic Management

Gig Economy Control Systems

- Uber/Lyft: Surge pricing manipulates driver behavior
- DoorDash: "Top Dasher" gamification
- Amazon Flex: Competitive block grabbing
- TaskRabbit: Rating anxiety

"Flexibility" masks algorithmic control

Ethical Considerations

Core Issues:

1. **Fair Wages:** Median \$2/hour globally
2. **Invisible Labor:** "Ghost work" props up AI
3. **Power Asymmetry:** Platforms hold all power
4. **Global Inequality:** Exploits wage differentials
5. **Psychological Impact:** Trauma from content
6. **Skill Development:** Does crowd work build careers?

The Future: Human-AI Collaboration

Near Term (2024-2026)

- AI handles routine, humans handle exceptions
- Humans provide judgment, AI provides scale
- Creative work becomes prompt engineering
- Quality control becomes more critical

Key Question:

Who captures the value created?

Welcome to the Future of Human Computation

Remember:

Every AI breakthrough stands on invisible human shoulders

Our Mission:

Make that human intelligence visible, valued, and ethical

Let's build better systems together