When Incentives Backfire, Data Stops Being Human



Data is Invaluable

Progress in AI ~ Amount of High Quality Date

- Deep Learning Era
 Human Annotations
- Language Models Era ← Internet Data



Shortcuts Donkey does not like walking, so it reaches out its neck to grab the carrot



t	a	
S		

But Human Data is in Crisis

Generative LLM Era has caused a crisis:

- Annotations: Workers use LLMs to shortcut.
- Internet: Flooded with LLM content.



Human Factors

- In Quality: High quality data depends on how engaged people are when producing data. Modulated by giving them financial incentives.
- In Quantity: High quantity data depends on making tasks parallelizable and repeatable. Done by atomizing bigger tasks / task fragmentation.

Era of Experience

- Designing in the Middle: Systems which help users carry out their daily tasks, but also collect useful (non-personal) data for machine learning.
- Deployed Products: Examples include Roomba, Tesla, Waymo, ReCAPTCHA. Requires high trust.

Sebastin Santy, Prasanta Bhattacharya, Manoel Horta Ribeiro, Kelsey Allen, Sewoong Oh

Data Quality has become Sacred

Quantity available, but Quality has become elusive.

- Technical Approaches: DataComp, better models.
- But what about quality at the source of data?

Problem with Quality

- Overjustification Effect: No reward or unexpected reward outperforms expected + leads to sustained motivation in future iterations. Explained by Self-Perception and Self-Determination Theory.
- Workers optimizing for financial incentives leads to overjustification impacting data quality.

Games

- Offer structure, but also organic engagement. Has actions, strategies, decisions, all useful for AI.
- Games with a Purpose: ESP Game, Peek-a-boom, QuickDraw, Iconary, HumanOrNot, ROFT.
- Create new games, or re-purpose, or else, gamify.



How do we bring back high-quality human data?

We ask, what motivates people to contribute data in the first place?



Fatigue

Donkey likes walking, but walking is not worth the reward given anymore

Problem with Quantity

- Alienation: With fragmentation, workers lose sight of the broader purpose, eventually disengaging them from the product of their labor.
- Regression in Maslow's Needs: Work is no longer fulfilling. Done only for survival, data stops being human: in its quality, and in its effects on workers.

High Trust Design

- Identifiable Data: Gameplay data, not user data.
- Incentive Manipulation: Avoid psychological hooks. Find universal motivators to avoid bias.
- Rethinking Compensation: Pay for value of data, not as a lever. Post-hoc and delayed recognition.

University of Washington | A*STAR | Princeton University | University of British Columbia