An Optimisation View of Online Attention Markets

Lexing Xie, School of Computing, The Australian National University
In collaboration with Marco Cheung and Haiqing Zhu

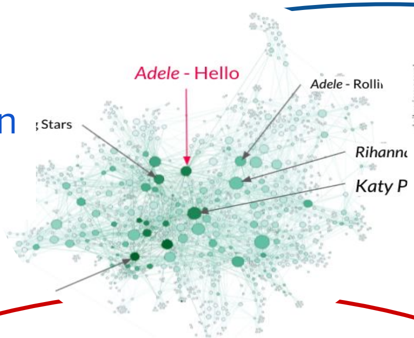
Integrated AI Network https://ai.anu.edu.au/
Computational Media Lab http://cmlab.dev

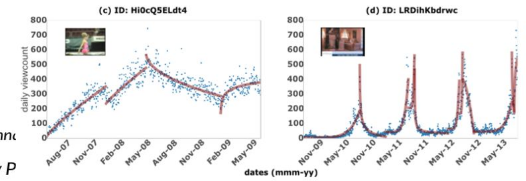
@REIKEN AIP, June 2025

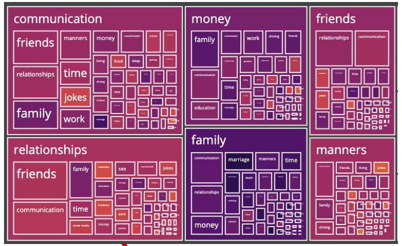
computational media lab @ ANU

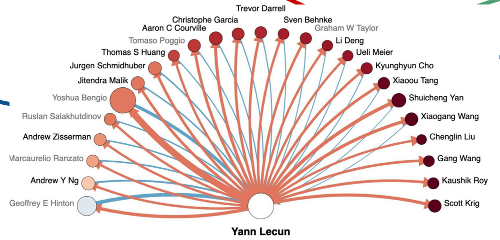
http://cmlab.dev

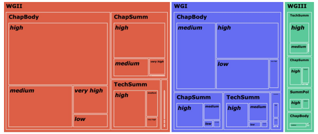
Machine learning, games and optimisation in attention markets

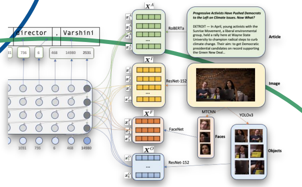


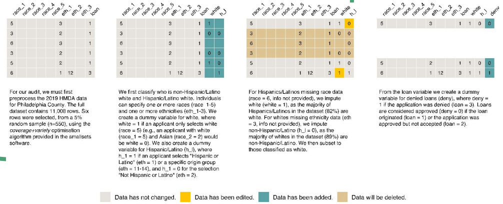










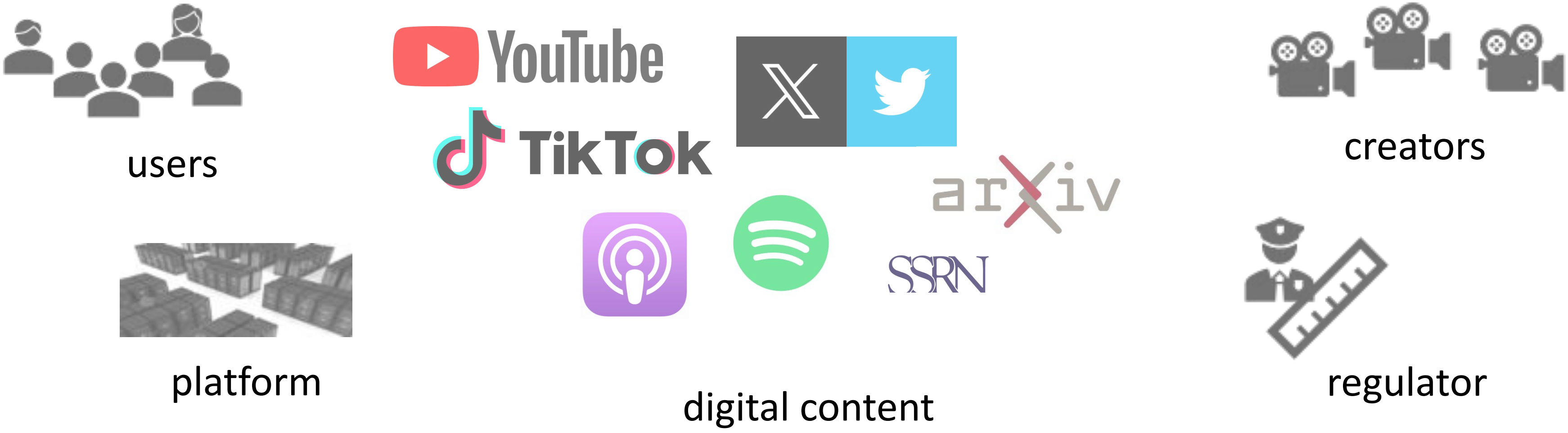


Human-centered AI with a purpose: bridging info gaps in climate and environment, understanding daily moral dilemmas

Inter-disciplinary AI, interactive visualisations, vision and language … …

Design + teach: {ML, algorithm, games} view of **Network Science**

Attention is a scarce resource

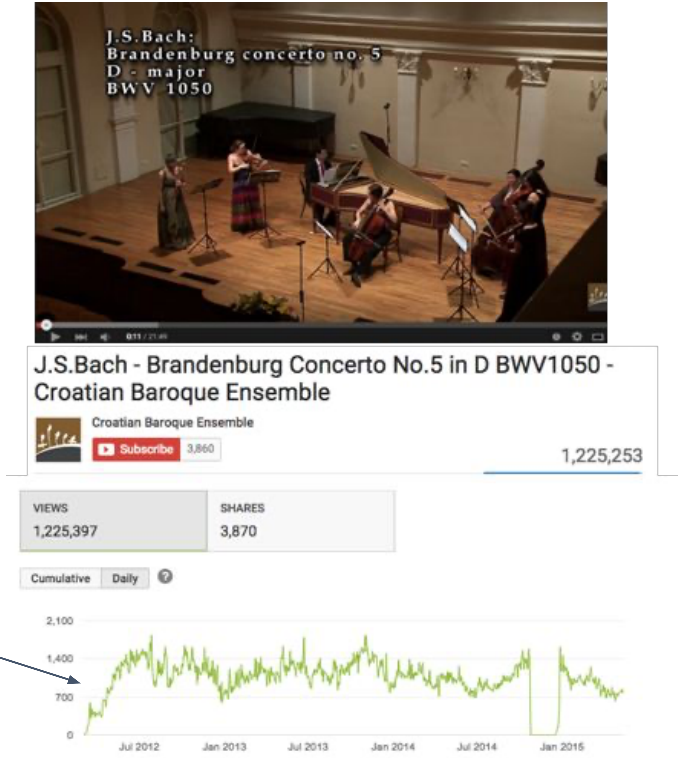


Abundance of content ➔ scarcity of attention [Simon 1971]

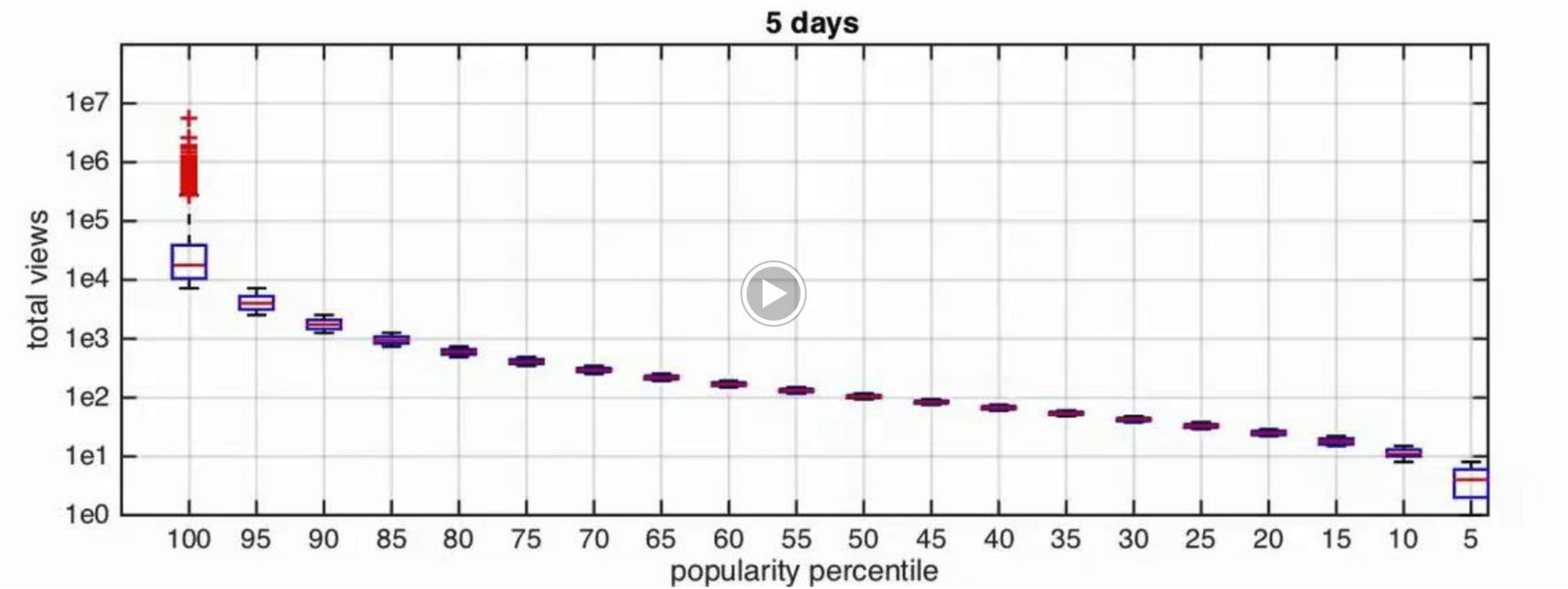
- How do attention evolve, what drives it? (with publicly available data)
- What are the properties of the market system involving content, users, and platforms?

What is item popularity?

daily # of views

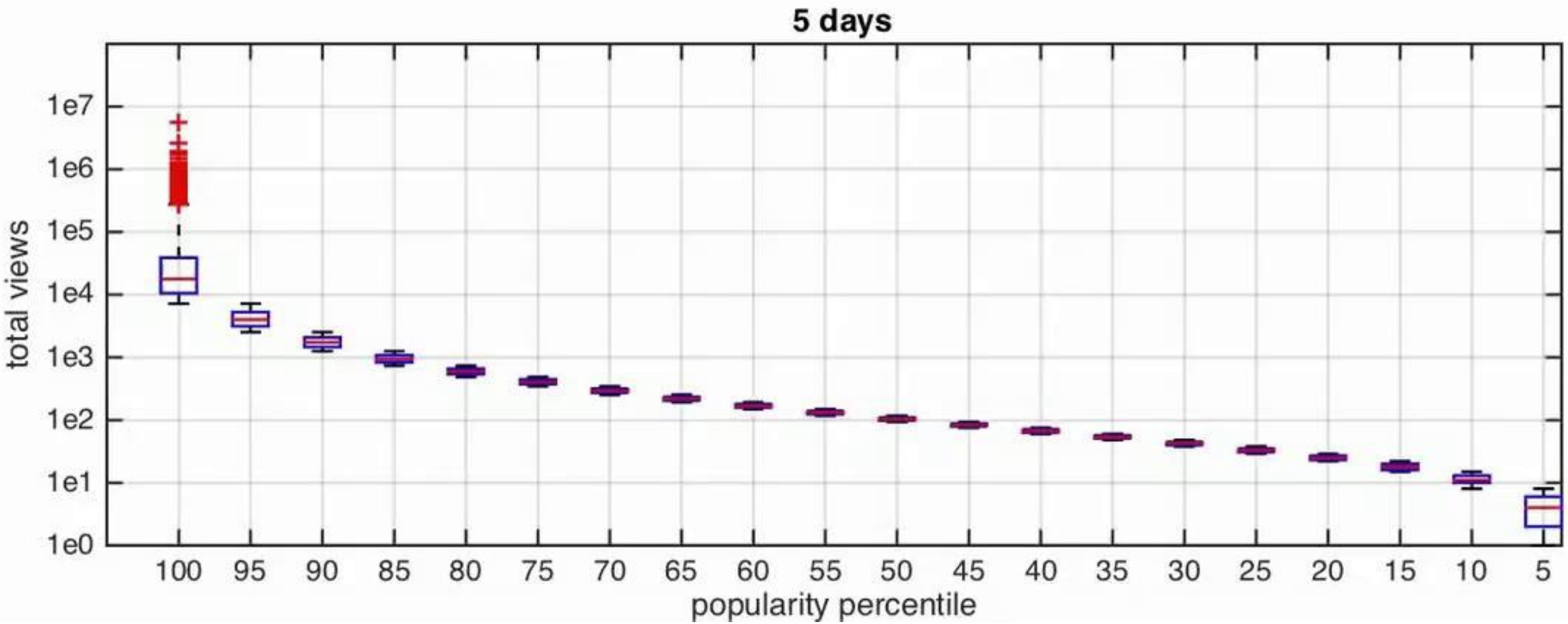


Popularity scale over time



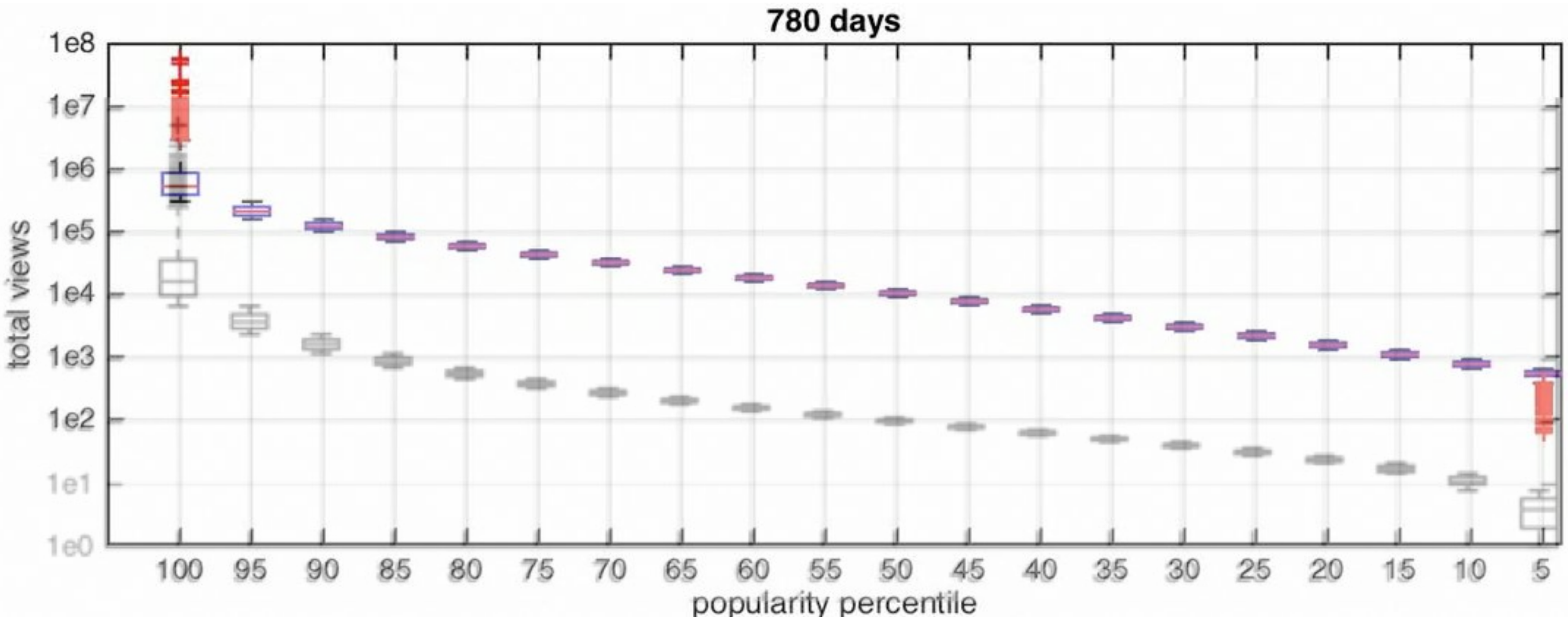
- 172K+ videos, >8K in each 5% *popularity bin*;
- videos in the middle bins are within 1.3x of each other's view-count.
- < 1% videos has 1M views after 2 years

Popularity scale over time



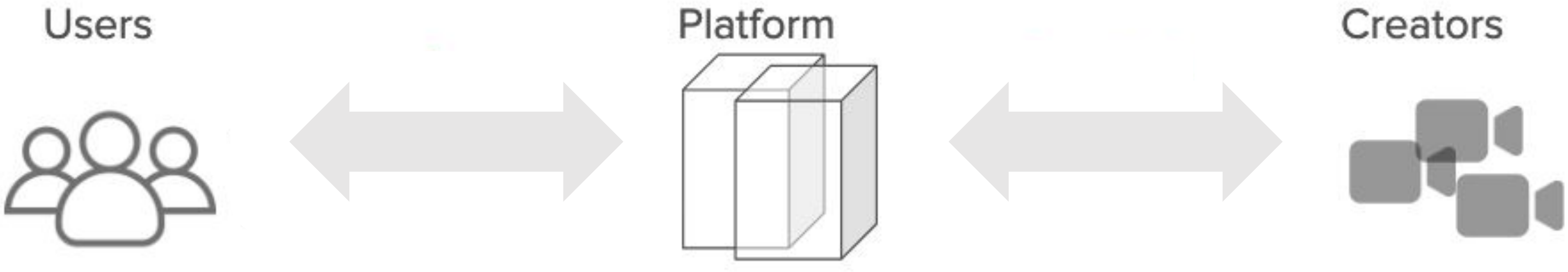
- 172K+ videos, >8K in each 5% *popularity bin*;
- videos in the middle bins are within 1.3x of each other's view-count.
- only 1% videos has 1M views after 2 years

“Rich-get-richer” as videos age



Videos (of the same popularity percentile) has ~100x in views over 2 years

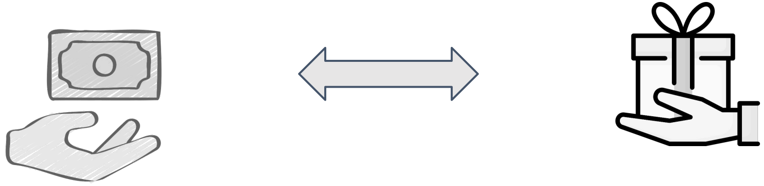
Talk outline



- How to describe **Attention as a marketplace**?
- Are there implicit potential (Lyapunov) functions for the underlying dynamics?
- Do the distributed interactions have equilibria – are they reachable? Are they good?
- Can we incentivize or influence quality of production?

Aside: morals, LLMs, influence flowers

Attention as a marketplace

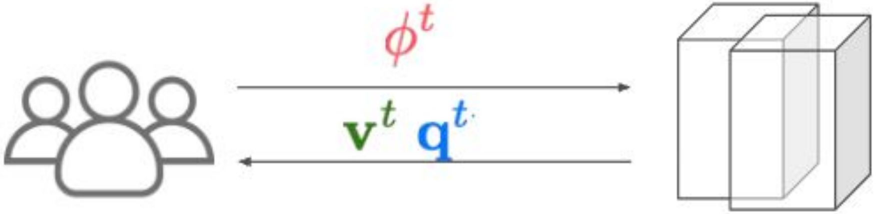


Time

Information, entertainment, education …

users

platform



Popularity update

$$\phi^t = \mathcal{P}(\mathbf{v}^t, \mathbf{q}^t, \phi^{t-1})$$

Markets	"Traditional" e.g. Arrow-Debreu	Attention
Supply	limited	∞
Price	modulate supply-demand	constant (per unit time)
Scarcity	$	attention/time
Market maker	minimal power	recsys/reward

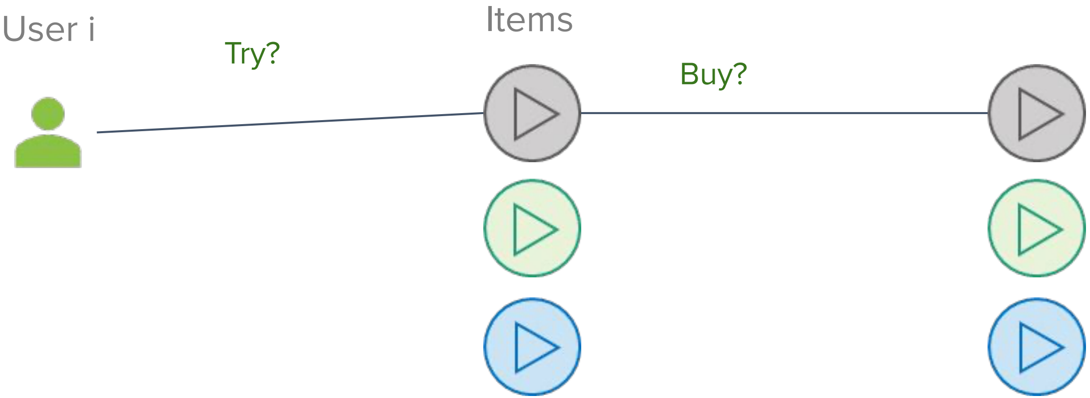
The Musiclab Experiment

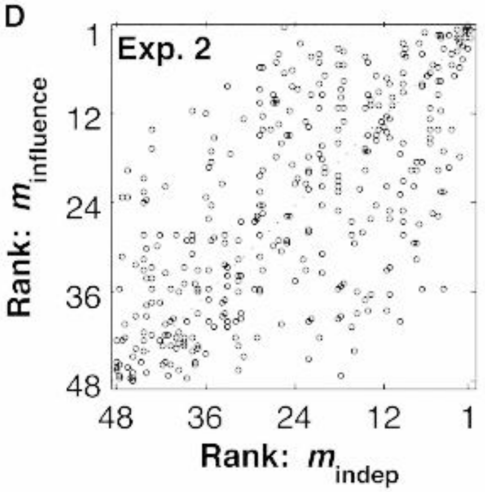
"[Experimental study of inequality and unpredictability in an artificial cultural market.]" Salganik, Dodds, and Watts. *Science*, 311:854-856, 2006.

14.3K participants, 48 unknown songs from unknown bands; 8x2+1 “worlds”

- popularity signal and ranking each plays a role
- unpredictable market shares observed across 8 separate “worlds”

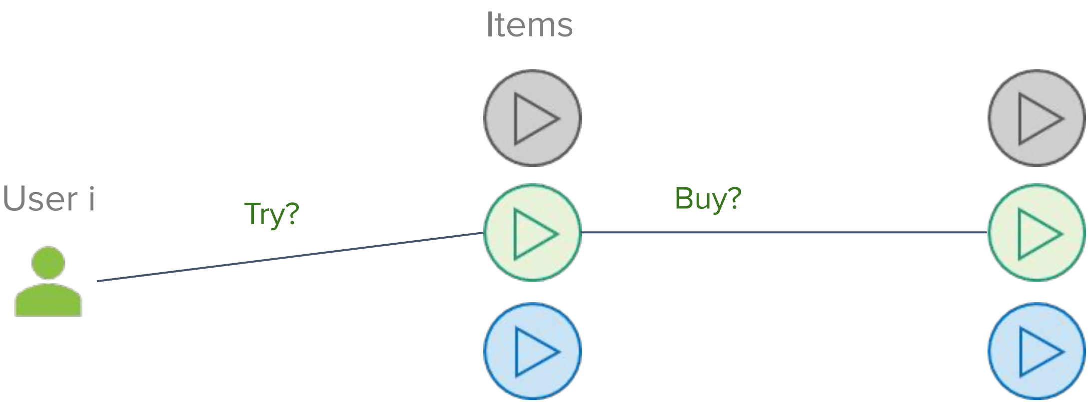
> “Success was also only partly determined by quality: The best songs rarely did poorly, and the worst rarely did well, but any other result was possible.”





Musiclab as a Trial-offer Market

[Salganik et al. 2006, Krumme et al 2012, Maldonado et al. 2018]



Probability to ***try*** item *j* -
multinomial logit choice

Probability to ***buy*** item *j*:
binomial coin-flip

$\propto v_{ij}(\phi_j^t)^r$ $\quad \boldsymbol{\phi}^t \quad q_{ij} \in [0,1] \quad \boldsymbol{\phi}^{t+1}$

Visibility (platform controls this)

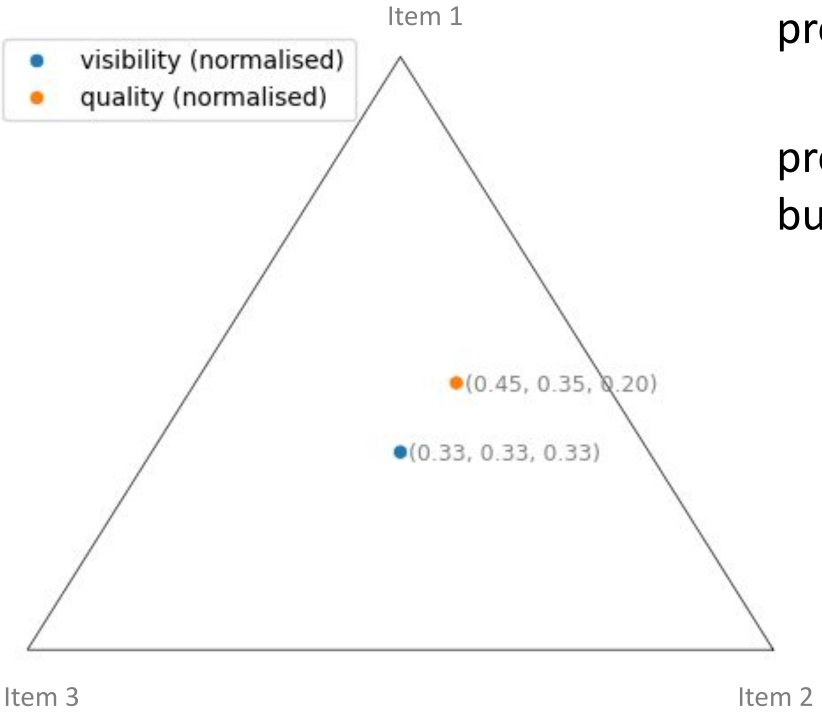
Market share (after round *t*)
$\phi \in \Delta$ n-dim simplex

Quality (intrinsic of each item)

- Trial-offer market with choice model describes the musiclab experiment.
- There is at least one fixed point in market share.

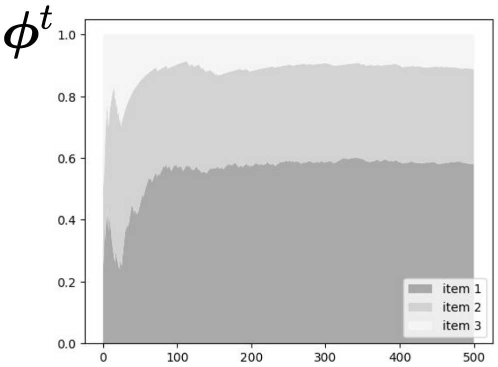
What is the dynamic and interactions between visibility, market share and quality?

Toy example: attention market with 3 items

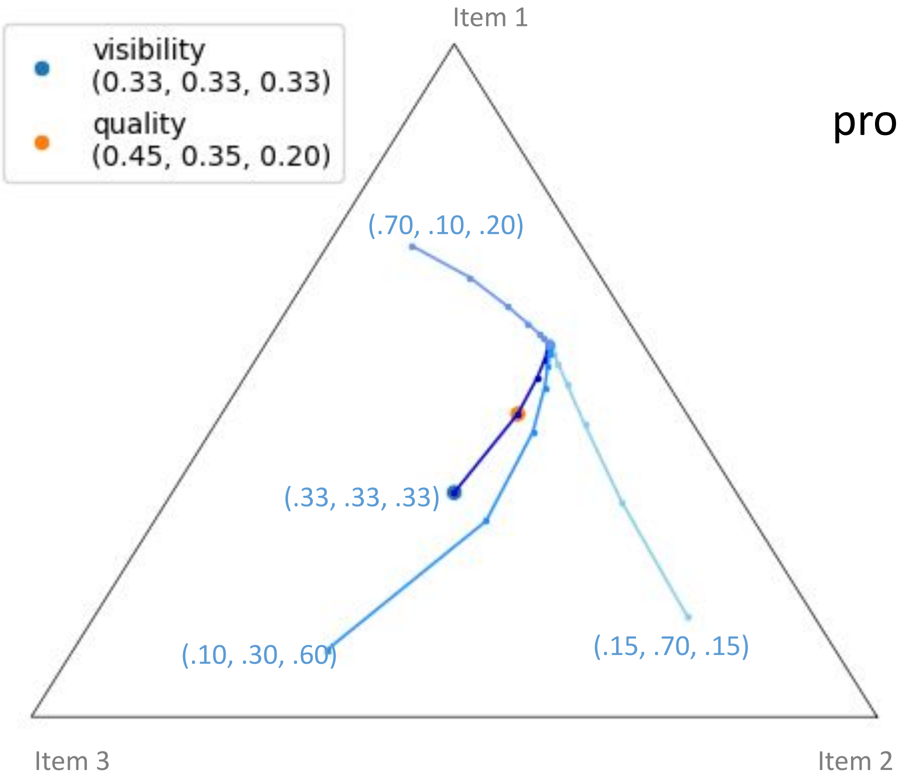


prob. to try item j $\propto \; v_j(\phi_j^t)^r$

prob. to try and then buy item j $\phi_j^{t+1} \; \propto \; q_j \; v_j(\phi_j^t)^r$



Toy example: how does quality affect market share?



prob. to buy item j $$\phi_j^{t+1} \propto q_j \, v_j (\phi_j^t)^r$$

set r = 0.5

Does this dynamical system have an objective function?

Will this stochastic process converge?

Two potential functions

let $\bar{q}_j = v_j q_j$

Trial-offer update $$\phi_j^{t+1} \propto \bar{q}_j (\phi_j^t)^r$$

Total utility

$$\max \quad \sum_{j=1}^{|\mathcal{I}|} \bar{q}_j \phi_j^r,$$

subject to $\boldsymbol{\phi} \in \Delta$.

Log utility **regularised by** **entropy**

$$\max \quad \Psi(\boldsymbol{\phi}) := \sum_{j=1}^{|\mathcal{I}|} \left(\phi_j \log \bar{q}_j - (1-r)\phi_j \log \phi_j\right),$$

subject to $\boldsymbol{\phi} \in \Delta$.

Does it naturally get to the optimal?

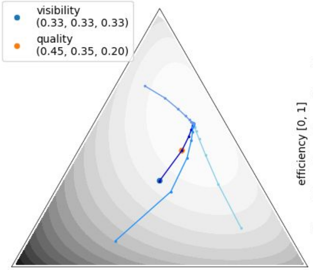
Mirror descent update with KL div

$$\min_{\phi \in \Delta} \left\{ \sum_{j=1}^{|\mathcal{I}|} \left(-\log \bar{q}_j + (1-r)(1 + \log \phi_j^{t-1})\right) \cdot \phi_j + \left(\phi_j \log \frac{\phi_j}{\phi_j^{t-1}} - \phi_j\right) \right\}$$

stationary point for the argmin

$$\phi_j^* \propto (\bar{q}_j)^{1/(1-r)}$$

Unique equilibrium



Attention market with personal preferences

Change of variable $\phi_{ij}^t \leftarrow \mathrm{normalize}(b_{ij}^t)$ $b_j^t = \sum_i b_{ij}^t$

$$b_{ij}^t = w_i q_{ij} \frac{v_{ij}(\phi_j^{t-1})^{r_i}}{\sum_k v_{ik}(\phi_k^{t-1})^{r_i}} = w_i q_{ij} \frac{v_{ij}(b_j^{t-1})^{r_i}}{\sum_k v_{ik}(b_k^{t-1})^{r_i}}$$

Fraction of the population w preference q_{ij}

Positive feedback loop: higher market share begets more attention

Attention market with personal preferences

Change of variable $\phi_{ij}^t \leftarrow \mathrm{normalize}(b_{ij}^t)$ $b_j^t = \sum_i b_{ij}^t$

$$b_{ij}^t = w_i q_{ij} \frac{v_{ij}(\phi_j^{t-1})^{r_i}}{\sum_k v_{ik}(\phi_k^{t-1})^{r_i}} = w_i q_{ij} \frac{v_{ij}(b_j^{t-1})^{r_i}}{\sum_k v_{ik}(b_k^{t-1})^{r_i}}$$

Positive feedback loop: higher market share begets more attention

Proportional response in Fisher Markets

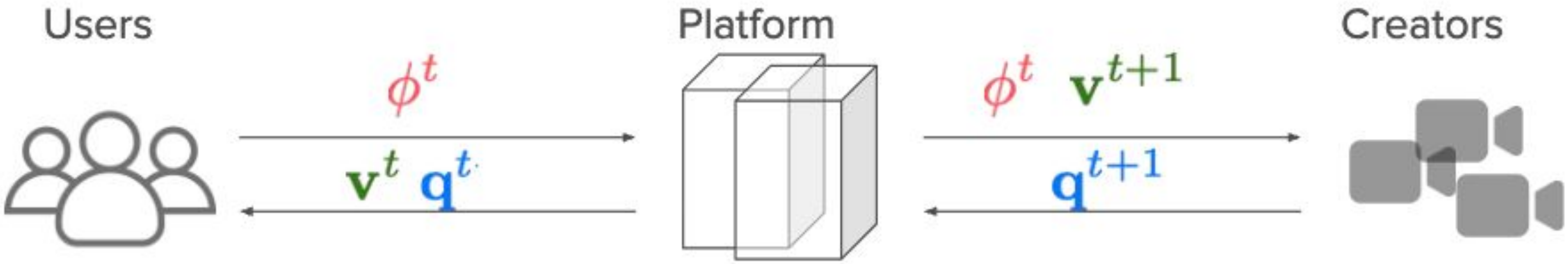
[Zhang2007,Birbaum2011, Cheung2018]

$$b_{ij}^t = w_i \frac{v_{ij}(b_{ij}^{t-1}/b_j^{t-1})^{r_i}}{\sum_{k=1} v_{ik}(b_{ik}^{t-1}/b_k^{t-1})^{r_i}}$$

Negative feedback loop: higher price drives down consumption

- Overall objective function is similar to Nash social welfare
- The probabilistic response dynamic is stochastic **mirror descent**

Attention Markets are Two-sided



Popularity update

$$\phi^t = \mathcal{P}(\mathbf{v}^t, \mathbf{q}^t, \phi^{t-1})$$

Quality update

$$\mathbf{q}^{t+1} = \mathcal{Q}(\mathbf{v}^{t+1}, \phi^t; \mathbf{c})$$

$$\phi_j^{t+1} \propto q_j \, v_j (\phi_j^t)^r$$

“Best response” by creators

Creators act by adjusting their q to maximize their own utilities.

$$\text{maximize} \quad u^t(q_j) = q_j \cdot \underbrace{\frac{v_j^{t+1}(\phi_j^t)^r}{\sum_i v_i^{t+1}(\phi_i^t)^r}}_{} - c_j(q_j).$$

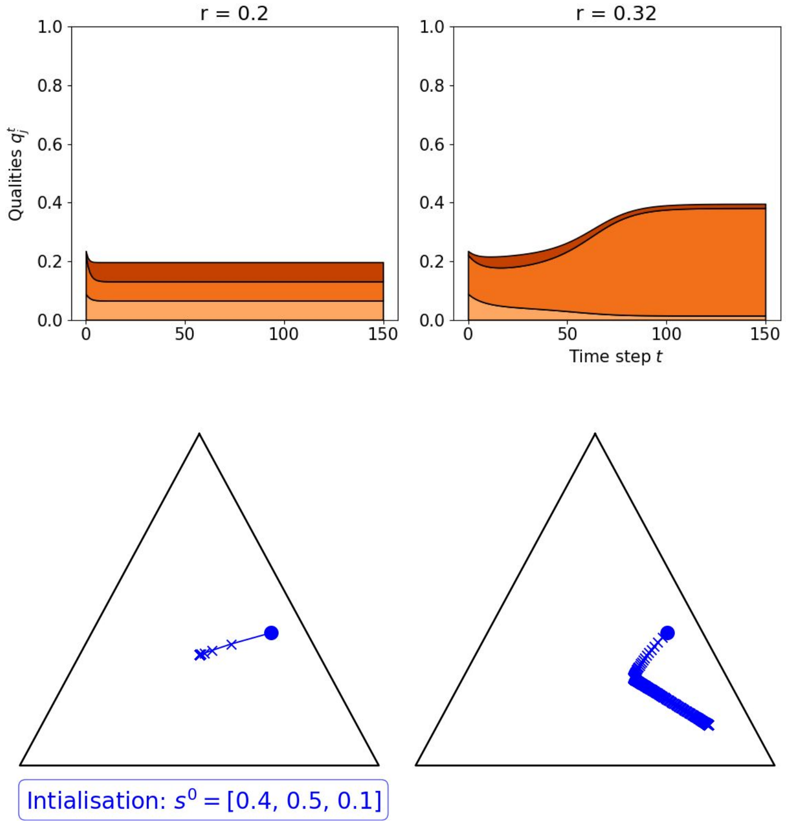
$$c_j'(q_j) = s^t_j$$

: Exposure - prob. of item j being shown at time t

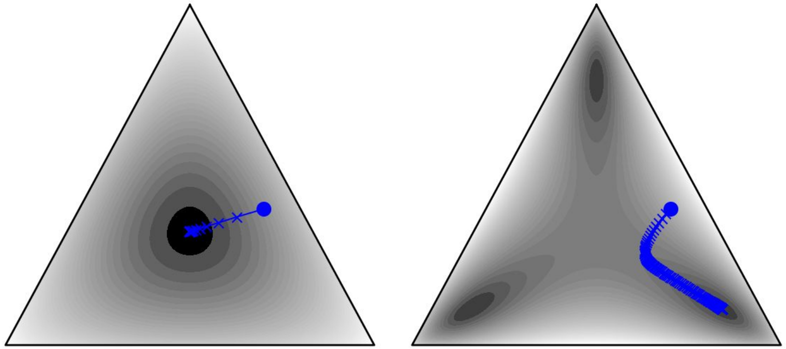
“Best response” by creators

Larger r

- stronger feedback from market share.
- Can stimulate creators to improve quality and hence be more competitive.
- Subject to randomness in initial q



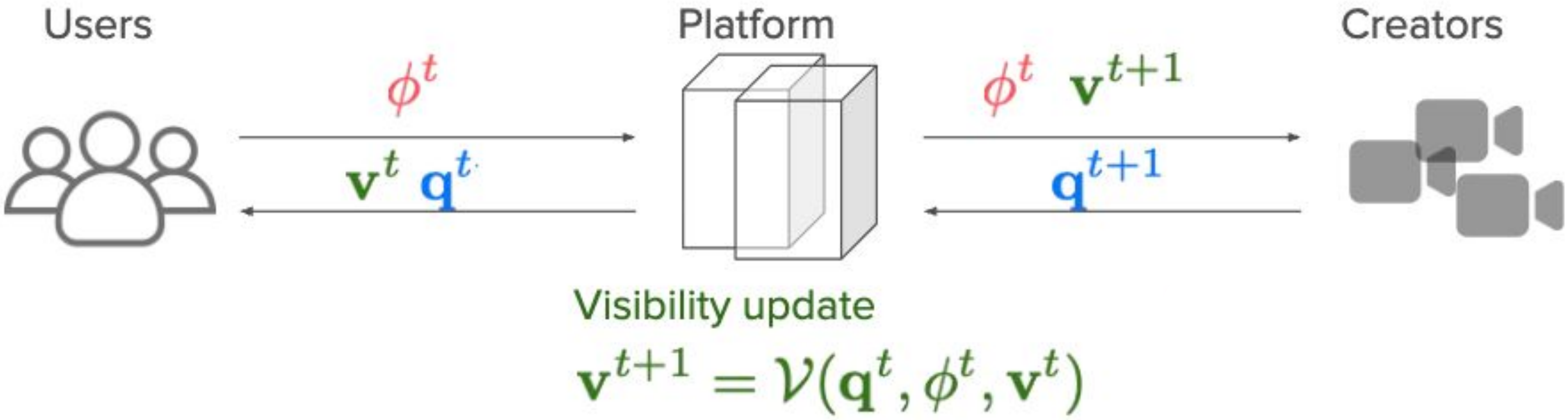
Potential function



Intialisation: $s^0 = [0.4, 0.5, 0.1]$

$$\max \sum_j \left(\underbrace{s_j \log v_j}_{\text{Alignment between exposure and recsys}} + \underbrace{(r-1) s_j \log s_j}_{\text{Entropy}} + \underbrace{r \int_0^{s_j} \log (c_j')^{-1}(z)\, \mathrm{d}z}_{\text{Production cost}} \right)$$

What can the platform do?



Mixed recommendation: to balance among quality q, popularity ϕ, signal from any other data μ

$$\frac{\mu_j (q_j^t)^\alpha (\phi_j^t)^r}{\sum_i \mu_i (q_i^t)^\alpha (\phi_i^t)^r}$$

r: market signal strength
α: quality signal strength

Potential functions – generalised

Mixed recommendation strategy - a, b, are constants defined by α, r and learning rate.

$$\max \sum_j \left(s_j \log v_j + (\;b\;) s_j \log s_j + a \int_0^{s_j} \log(c_j')^{-1}(z)\ \mathrm{d}z \right)$$

Constant recommendation

$$\max \sum_j \left(\underbrace{s_j \log v_j}_{\text{Alignment between exposure and recsys}} + \underbrace{(r - 1) s_j \log s_j}_{\text{Entropy}} + \underbrace{r \int_0^{s_j} \log(c_j')^{-1}(z)\ \mathrm{d}z}_{\text{Production cost}} \right)$$

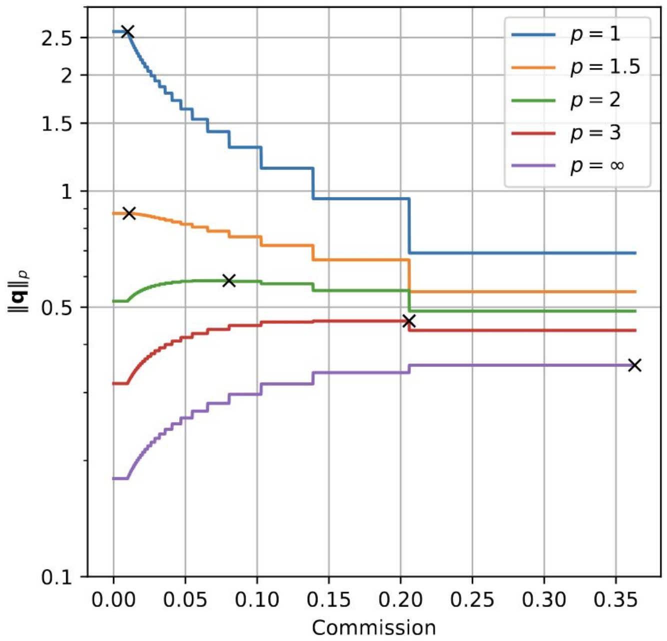
[Zhu, Xie, Cheung, in submission]

Entry fees vs quality

Can one discourage or prevent production of low-quality content?

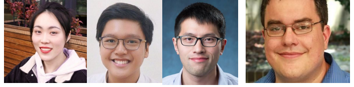
- "Structural" … existing participants prevents new participants with worse cost functions from entering
- Strategic: Platform charge a commission and redistribute the income to creators

Different reward strategies for two-sided markets is worth further investigation.



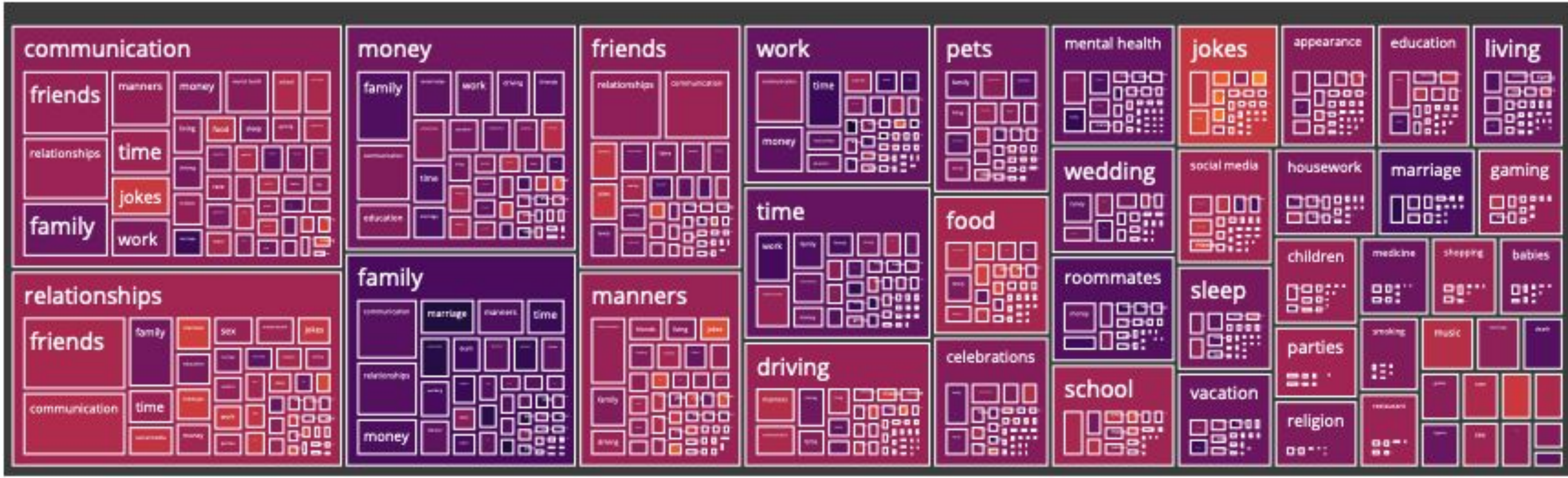
Q: What can we say about other people's work on creator incentives? In fact, we should have a related work slide, what should it say?

Aside 1 of 3 - Mapping 100,000 real-life moral dilemmas



[Nguyen, Chen et al, ICWSM 2022, 2024, in submission]]

- /r/AmITheAsshoe - what moral issues do people grapple with?
- 47 topics found, people perceive them in pairs
- Empirical philosophy meets moral psychology - judgements are malleable
- NLP method: prompting right is the key to labeling moral+value relevance

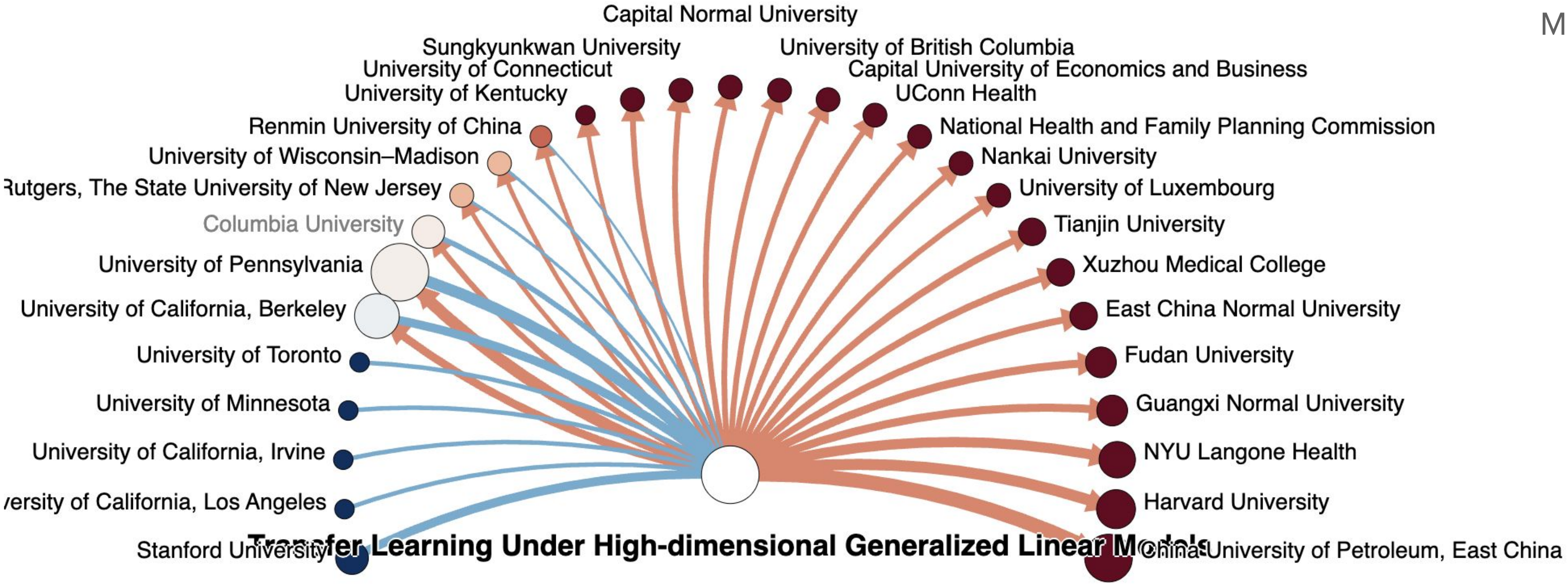


 # What is an influence flower?

See arxiv.org/2105.14328 this flower

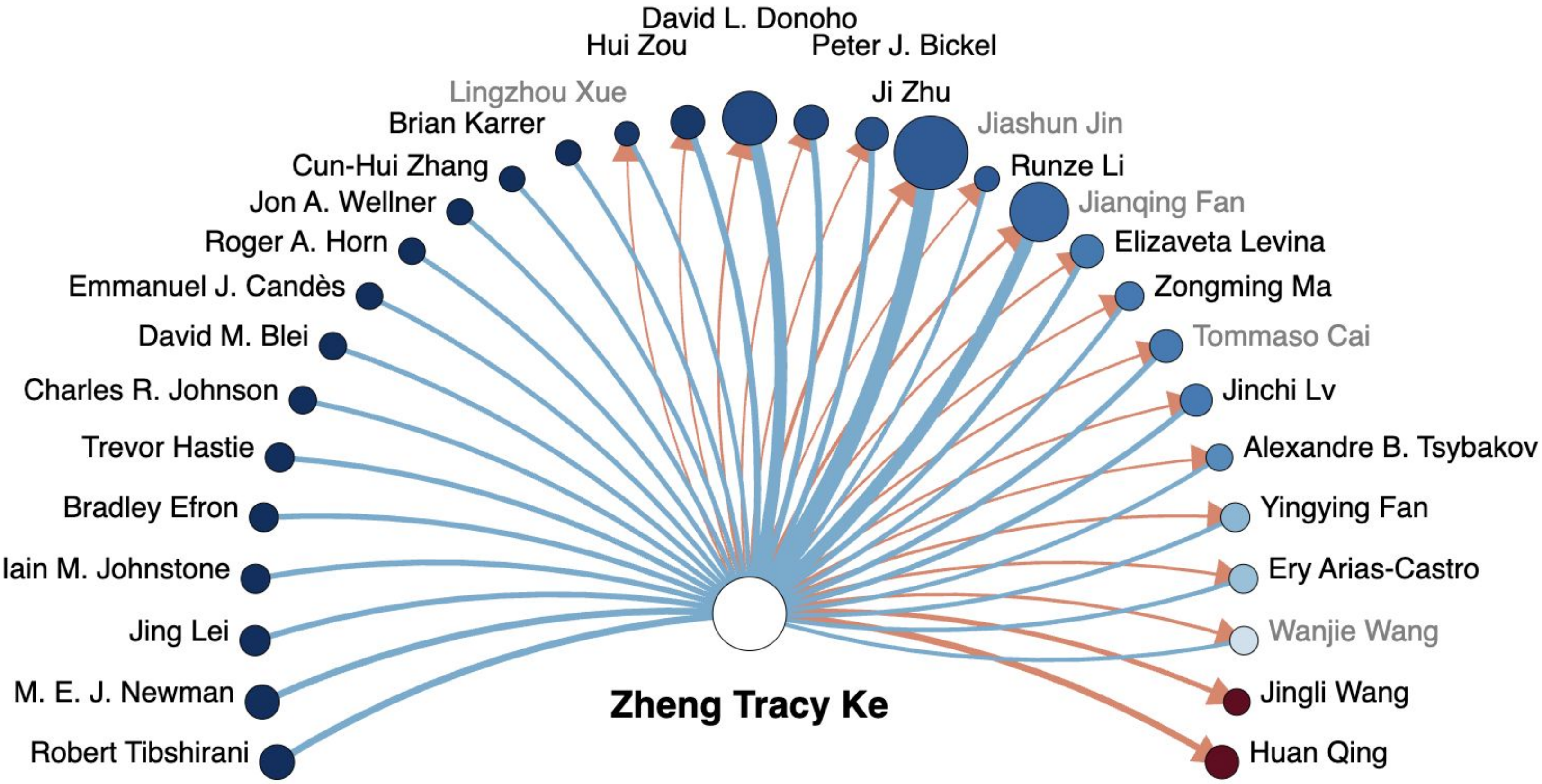
A qualitative visualisation and webapp to profile the incoming and outgoing intellectual influence among academic entities.

Minjeong Shin



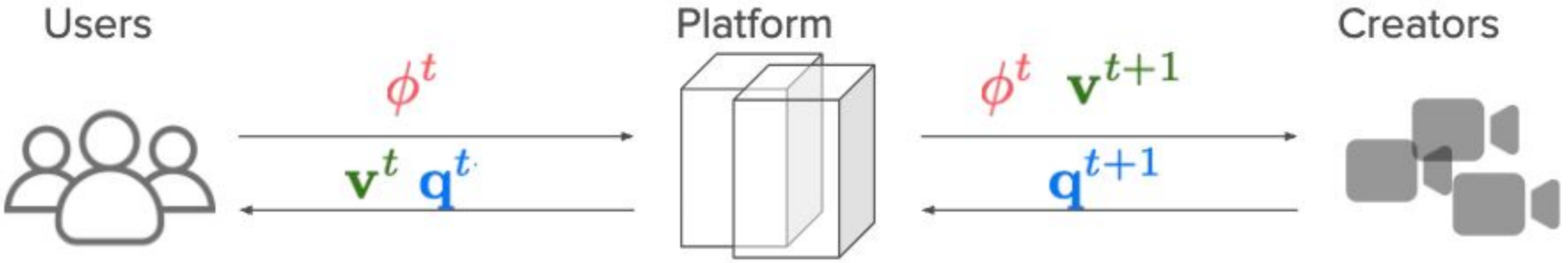
Blue arcs denote *incoming* influence from the authors to the paper, with their thickness proportional to the number of *references* made.
Red arcs denote *outgoing* influence from the paper to the authors, with their thickness proportional to the number of *citations* received.

Aside 2 of 3 - Influence Flower for Academic Entities



t

Summary and outlook



- ➔ Distributed interactions among Users, creators, platform is a market with positive feedback loop
- ➔ We uncover a series of underlying potential functions
- ➔ Natural interaction dynamics correspond to mirror descent on this landscape
- ➔ Structural and strategic barriers can incentivize creators

Ongoing work: fairness of the attention ecosystem, attention market in science