Expertise and Dynamics within Crowdsourced Musical Knowledge Curation: A Case Study of the Genius Platform

Derek Lim & Austin R. Benson

dl772@cornell.edu, arb@cs.cornell.edu

Code & data online

Cornell University

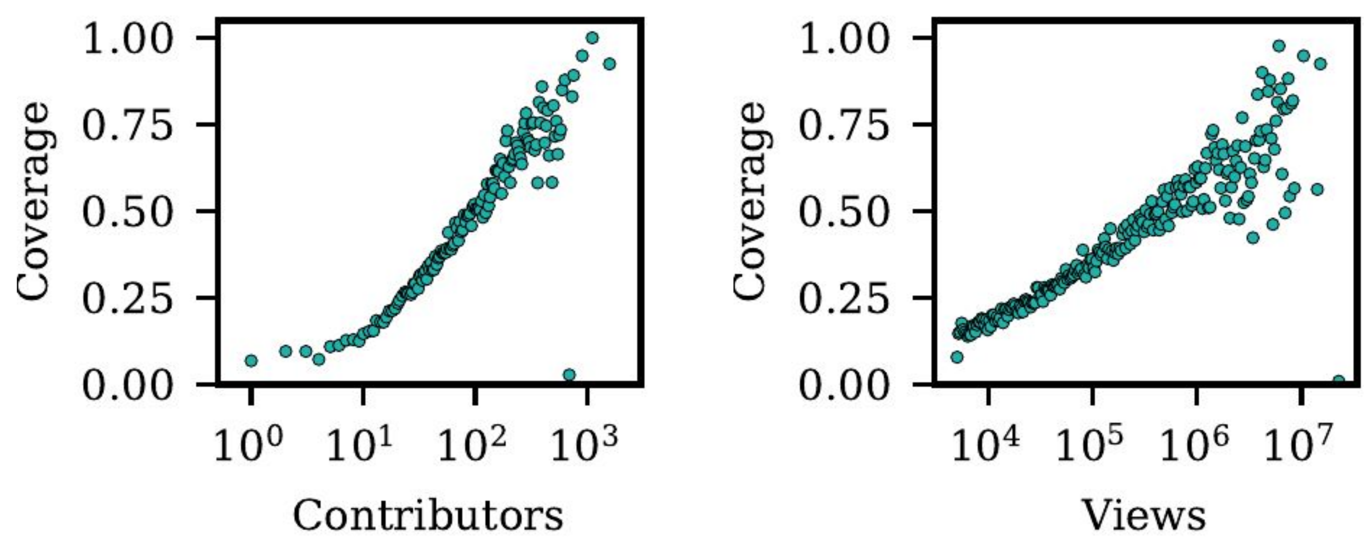
ICWSM

Genius differs from other crowdsourced info sites

- genius.com: widely-used platform primarily for crowdsourced annotations of song lyrics
- Informal content with slang, jokes, profanity (contrast w. Wikipedia's encyclopedic tone)
- Each lyric segment is an implicit prompt: "What is the meaning / what is interesting about these lyrics?" (contrast w. Q&A sites with explicit prompts)

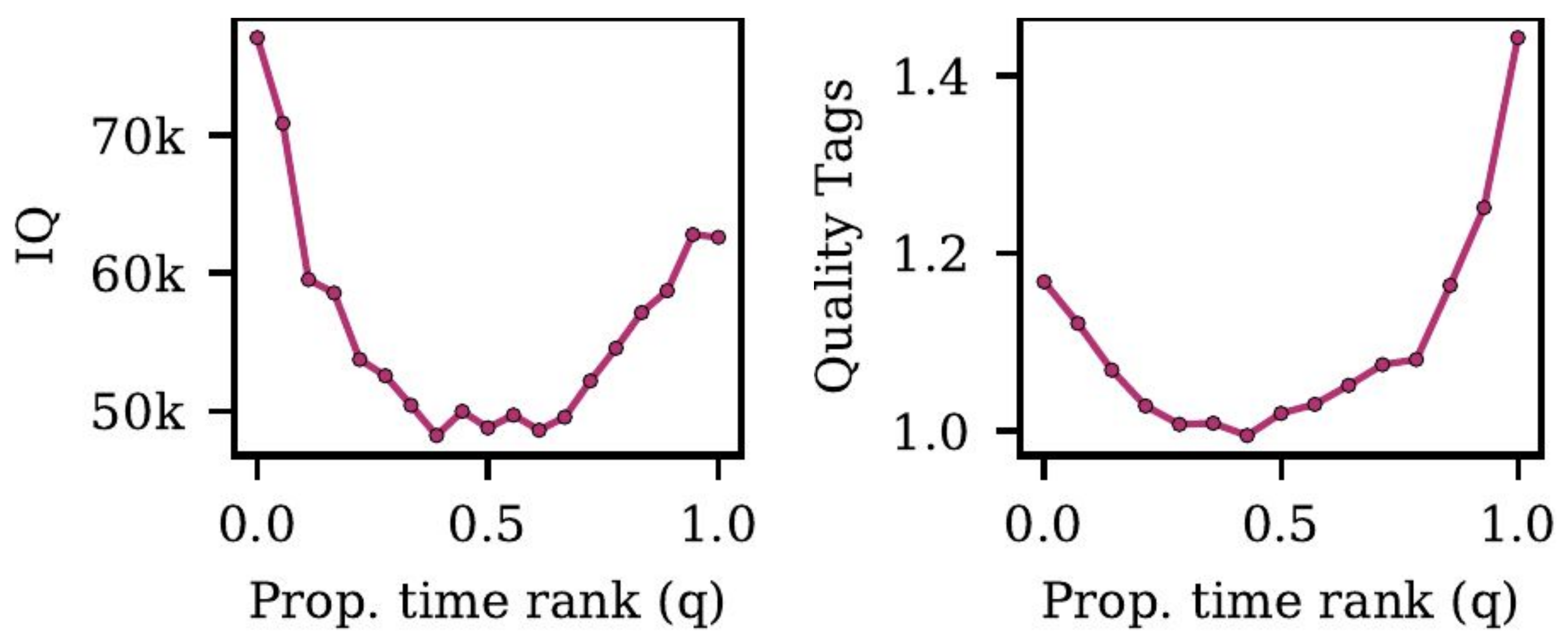
We define Genius-specific metrics

- Measure annotation quality by quality tags e.g. <img>, <iframe>, <blockquote>, <twitter-widget>
- Measure lyric originality by quantile functions and word frequencies
- Measure annotation coverage: proportion of lyrics of a song that are annotated



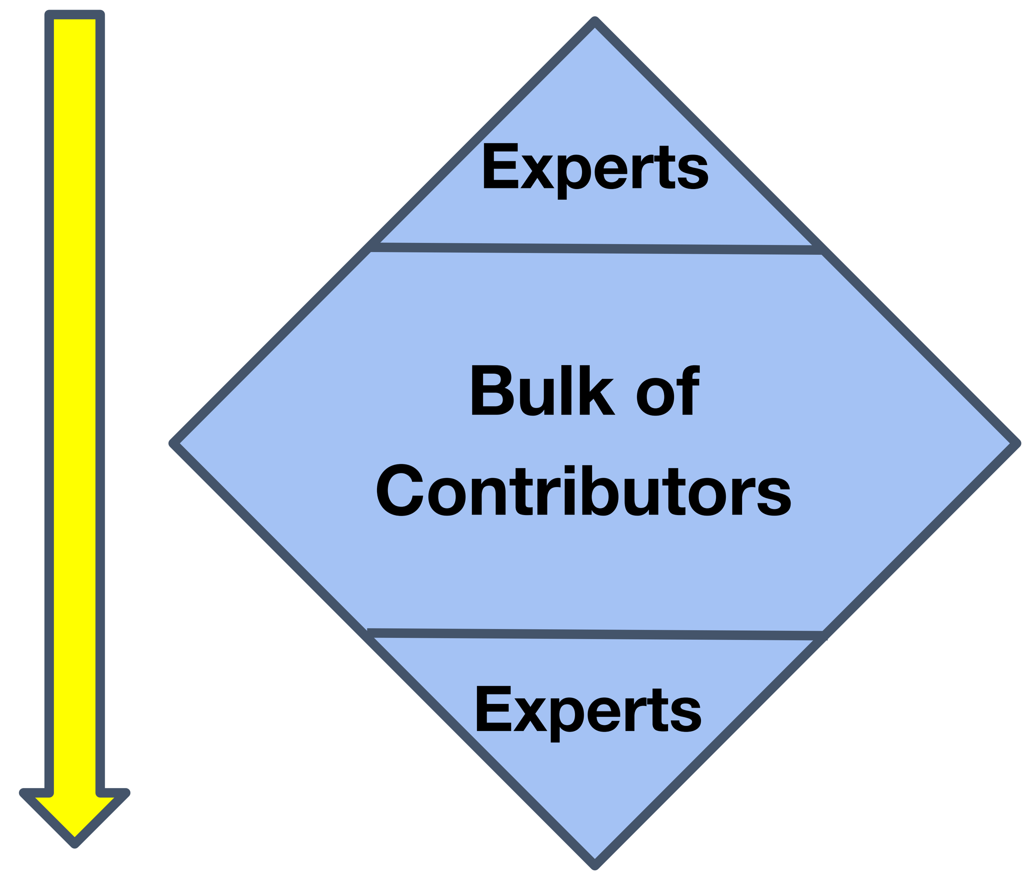
U-shaped temporal dynamics are unique to Genius

- Annotator expertise and annotation quality higher for early and late annotations on a song
- Differs from Stack Overflow (experts in early answers) and Yelp/Amazon Reviews (quality reviews early)



IQ Diamond Model

- Conceptually, lyrics processed and annotated by experts, then bulk of contributors, then experts again
- Contrast with "reputation pyramid" on Stack Overflow (expertise decreases for later answers)



A user utility model with both network effects and congestion

- Network effects: #annotation ↑ then utility ↑ because more popular songs give more reach and points
- Congestion effects: #annotation ↑ then utility ↓ because cannot annotate taken lyric segments, so less choice

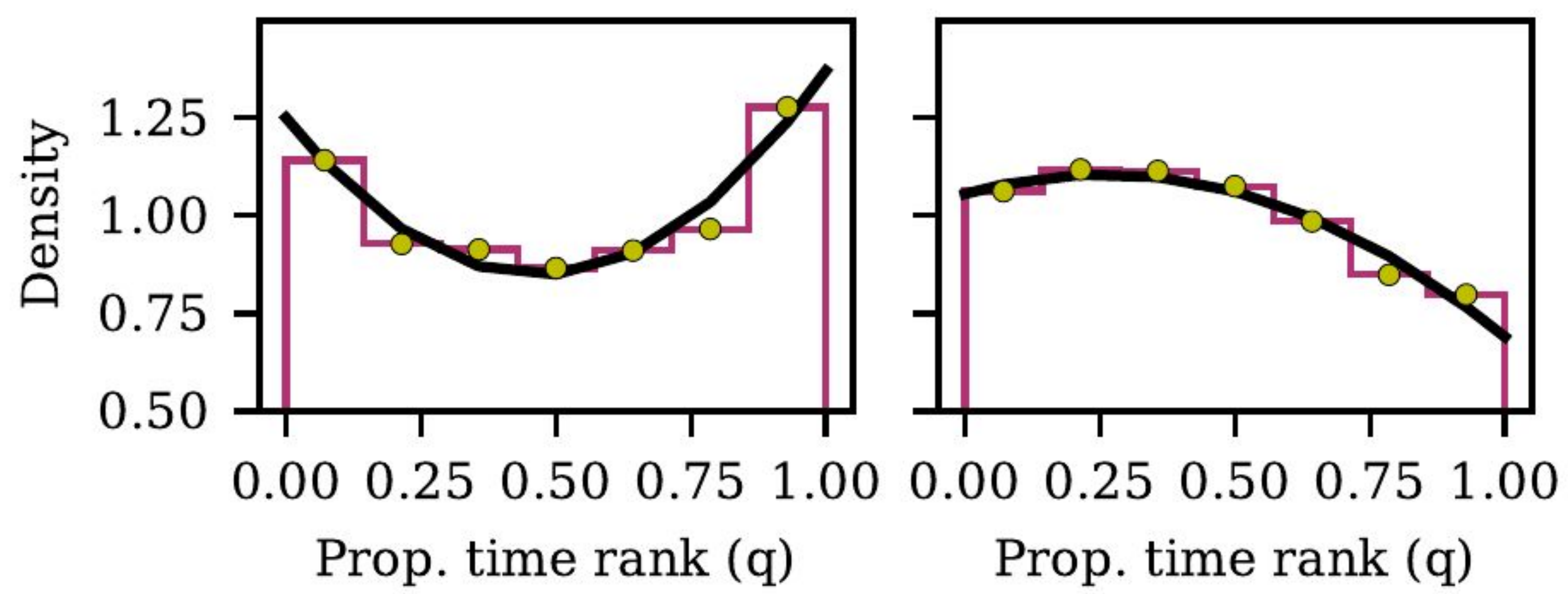
$$u_k(\rho) = \underbrace{b_k}_{\text{a priori utility}} + \underbrace{f_k\left(\sum_{j\neq k}\rho_j\right)}_{\text{network effects}} - \underbrace{g_k\left(\sum_{j=1}^{N}\rho_j\right)}_{\text{congestion effects}}$$

ρ_k = proportion of lyrics on song annotated by user k

Fit utility functions for high-expertise users and lower-expertise users

Experts (left) U-shaped utility over time

Other users (right) Inverted U-shape utility curve



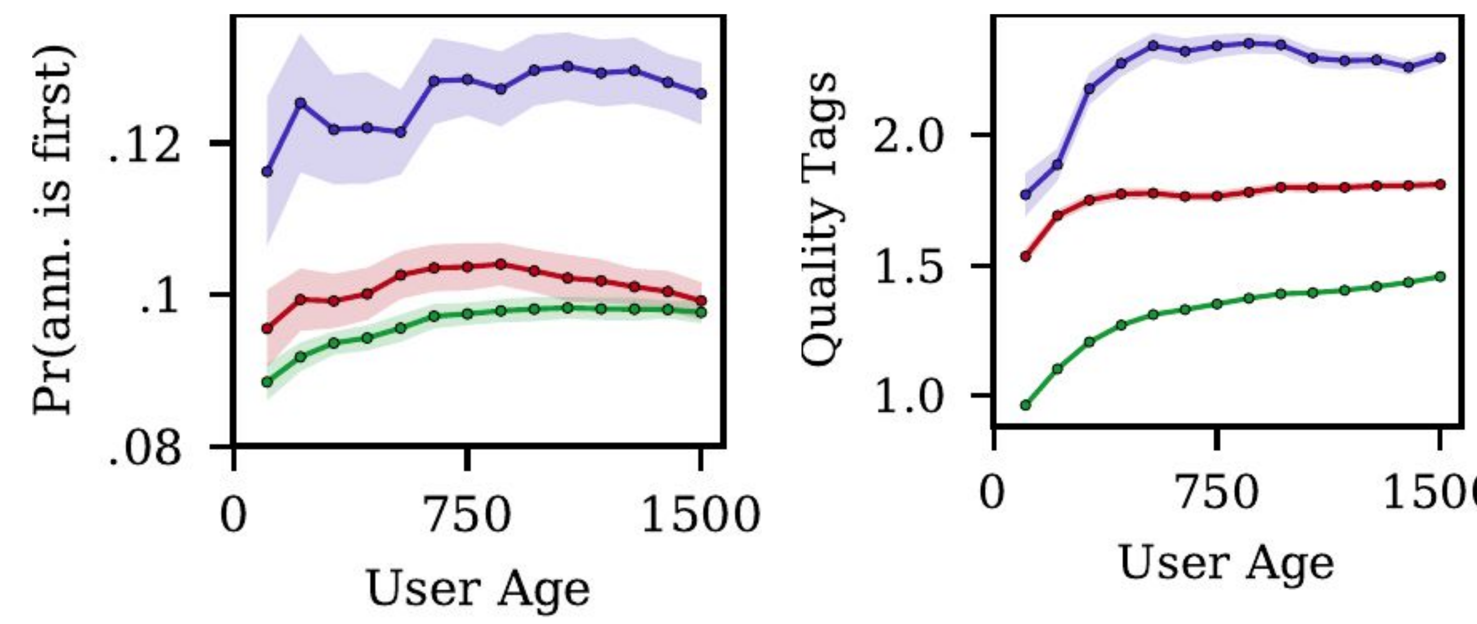
Acknowledgements. This research was supported in part by NSF Award DMS1830274, ARO Award W911NF19-1-0057, ARO MURI, and JPMorgan Chase & Co.

Users change, but early contributions of eventual experts and distinct

In first contributions, eventual experts (blue):

More often make first annotation on song

Make higher quality annotations



Successful early prediction of super-experts by annotations and edits

- Using features derived from IQ Diamond model and user evolution findings, we fit logistic regression models that substantially beat a majority-class baseline in prediction of super-experts (99.8 percentile expertise users)

α_1	mean # of quality tags in first 15 annotations
α_2	mean time between first 15 annotations
α_3	# of first 15 annotations that are a song's first
α_4	mean originality on songs for first 15 annotations
e_1	mean time between first 15 edits
e_2	# of first 15 edits that are an annotation's first

Predictors	Accuracy	AUC
$\alpha_1, \alpha_2, \alpha_3, \alpha_4, e_1, e_2$	$.673 \pm .029$	$.748 \pm .030$
$\alpha_1, \alpha_2, \alpha_3, \alpha_4, e_1$	$.671 \pm .029$	$.744 \pm .030$
$\alpha_1, \alpha_2, \alpha_3, \alpha_4$	$.659 \pm .030$	$.733 \pm .031$
$\alpha_1, \alpha_2, \alpha_3$	$.659 \pm .028$	$.727 \pm .031$
α_1, α_2	$.652 \pm .030$	$.715 \pm .033$
α_1	$.616 \pm .032$	$.674 \pm .034$