

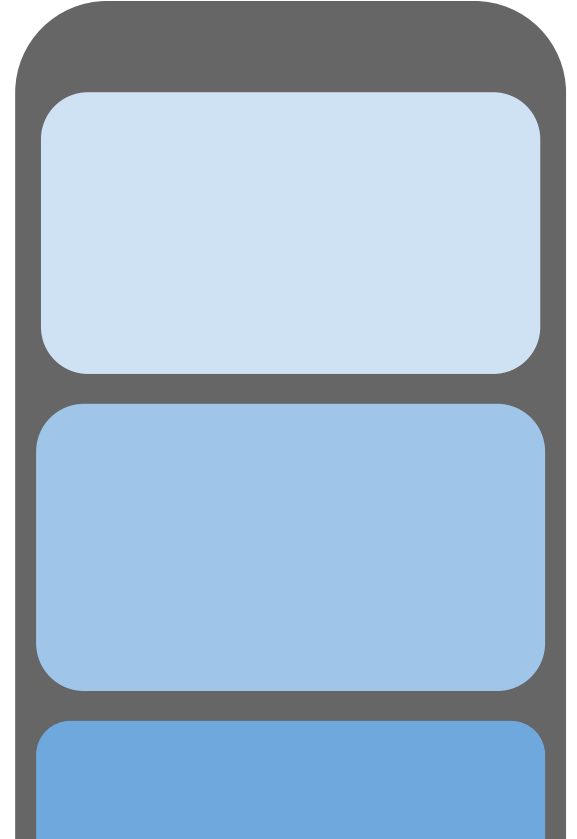
Purpose-Driven Feed Elicitation



Lindsay, Xiyuan, Charlotte, Tiziano, Michael

Content

1. Problems with Current Feeeeeeeds
2. Interviews -> Feed Elicitation
3. “Purpose-Driven” Feed Elicitation
4. System Demo
5. Experiment & Evaluation Plan



Engagement-oriented feed algorithm provides very limited information available to the users on what they see, and why.

- Learned from passive signals of user behavior
- For maximising engagement and time spent



Anxiety

Depression

Emotional Dysregulation



addiction

doom-scrolling



Polarisation

antagonism

Addiction
Attention Issues



How do we create a better feed FOR USERS?

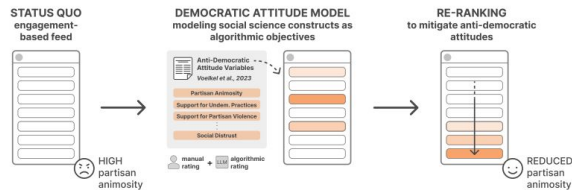
Alexandria: A Library of Pluralistic Values for Realtime Re-Ranking of Social Media Feeds

Akaash Kolluri^{*1}, Renn Su^{*1}, Farnaz Jahanbakhsh², Dora Zhao¹,
Tiziano Piccardi¹, Michael S. Bernstein¹

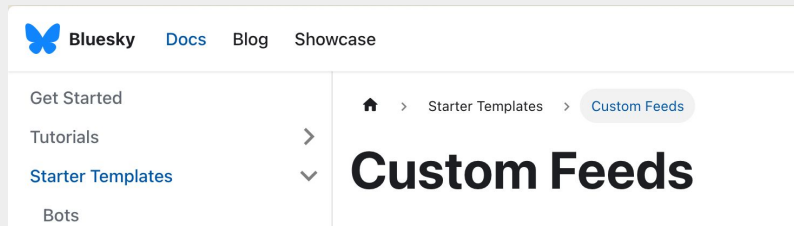
“Fixing the feed for users” via re-ranking /
new objective functions

Embedding Democratic Values into Social Media AIs via Societal Objective Functions

CHENYAN JIA^{*}, Stanford University, USA; Northeastern University, USA
MICHELLE S. LAM^{*}, Stanford University, USA
MINH CHAU MAI, Stanford University, USA
JEFFREY T. HANCOCK, Stanford University, USA
MICHAEL S. BERNSTEIN, Stanford University, USA



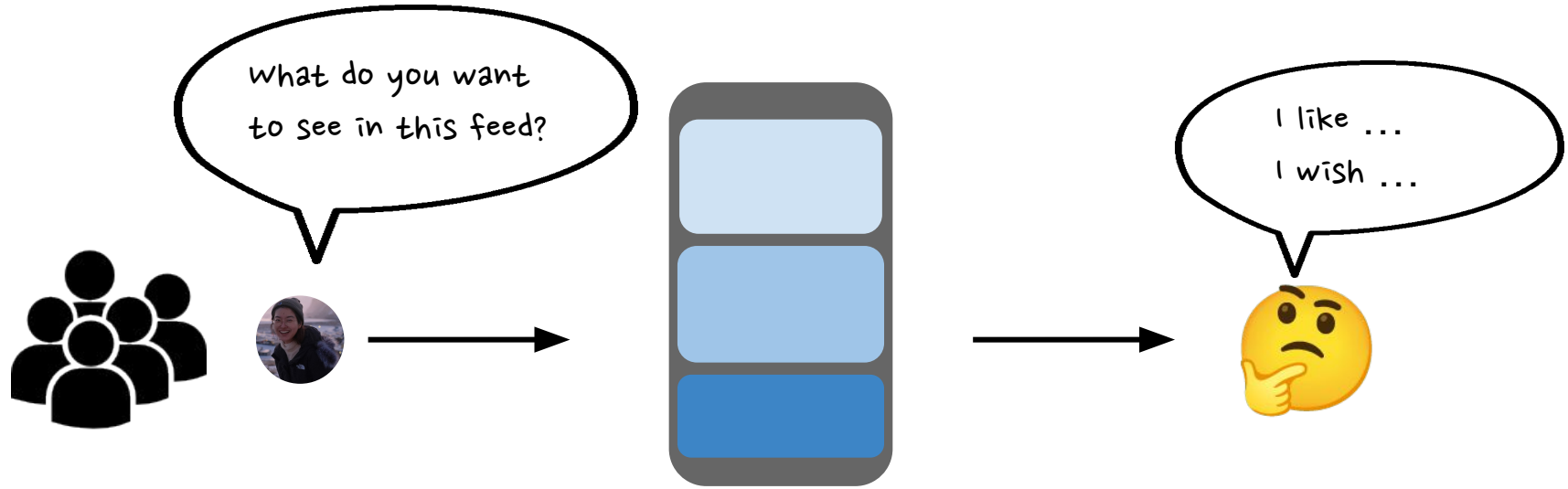
How do we allow users to create a better feed THEMSELVES?



Yet,

How do the users
imagine and articulate
their desired feed?

We tried feed elicitation through interviews.



We interview our lovely
labmates about their
ideal feed

Create feeds based
on interview

Ask for Feedback

Problem: People are fixated on content / topics. But they lack specification and changes rapidly.

What do you want
to see in this feed?



I want to see jokes
and funny videos!



Dark humor that goes
too far. 🦴



Sophisticated chemistry
jokes that nobody
understands and makes
you feel stupid. 🧪



Silly songs that loop
endlessly and get stuck
in your mind. 🤯



Exclude dark humors
and NSFW!



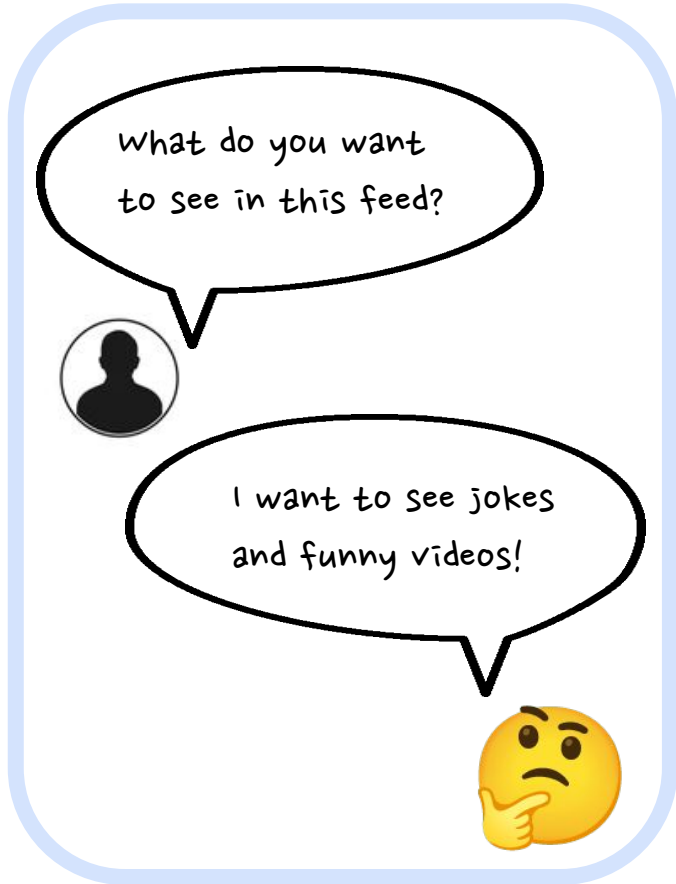
IF inner jokes,
prefer CS related;



Just something light
during quick work breaks !



Existing tools fixate our minds to only think of content.



- Our minds are wired to adapt to existing tools e.g. keyword search, hashtags, not thinking about other potential dimensions
- It takes effort to think through the possibilities



Insights: Beyond just Topics?
Purpose Driven -> 3 dimensions.

Our Proposed Method

Purpose

- Goal
- User context
- Experiential outcome

Topics

- Desired topics/content
- Seemingly relevant but undesired content

Quality

- Ideal qualities
- Quality Exclusions

We automated the **3 stage interview** process with GPT and guide users to think through **why they are creating this feed, and what does it look like**

Interview Principles

1. Only ask one question at a time.
2. Provide examples or suggestions of things users may want to consider including or excluding beyond their existing specifications
3. Aim to understand the user's decision boundary by probing what they want to see and what they do not want.
4. ...

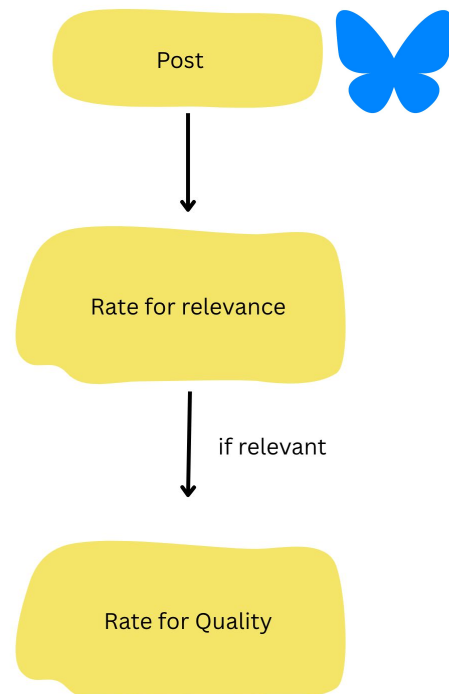
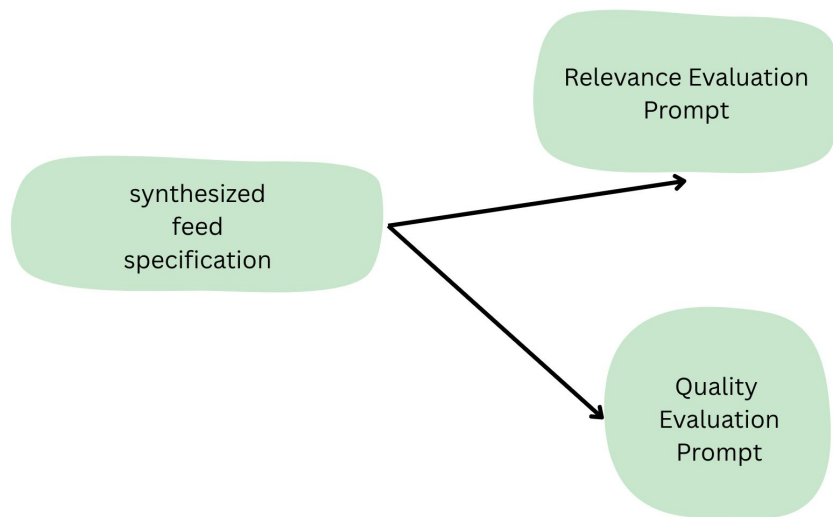


Live Demo

+ Sample feeds

Feed spec comparison

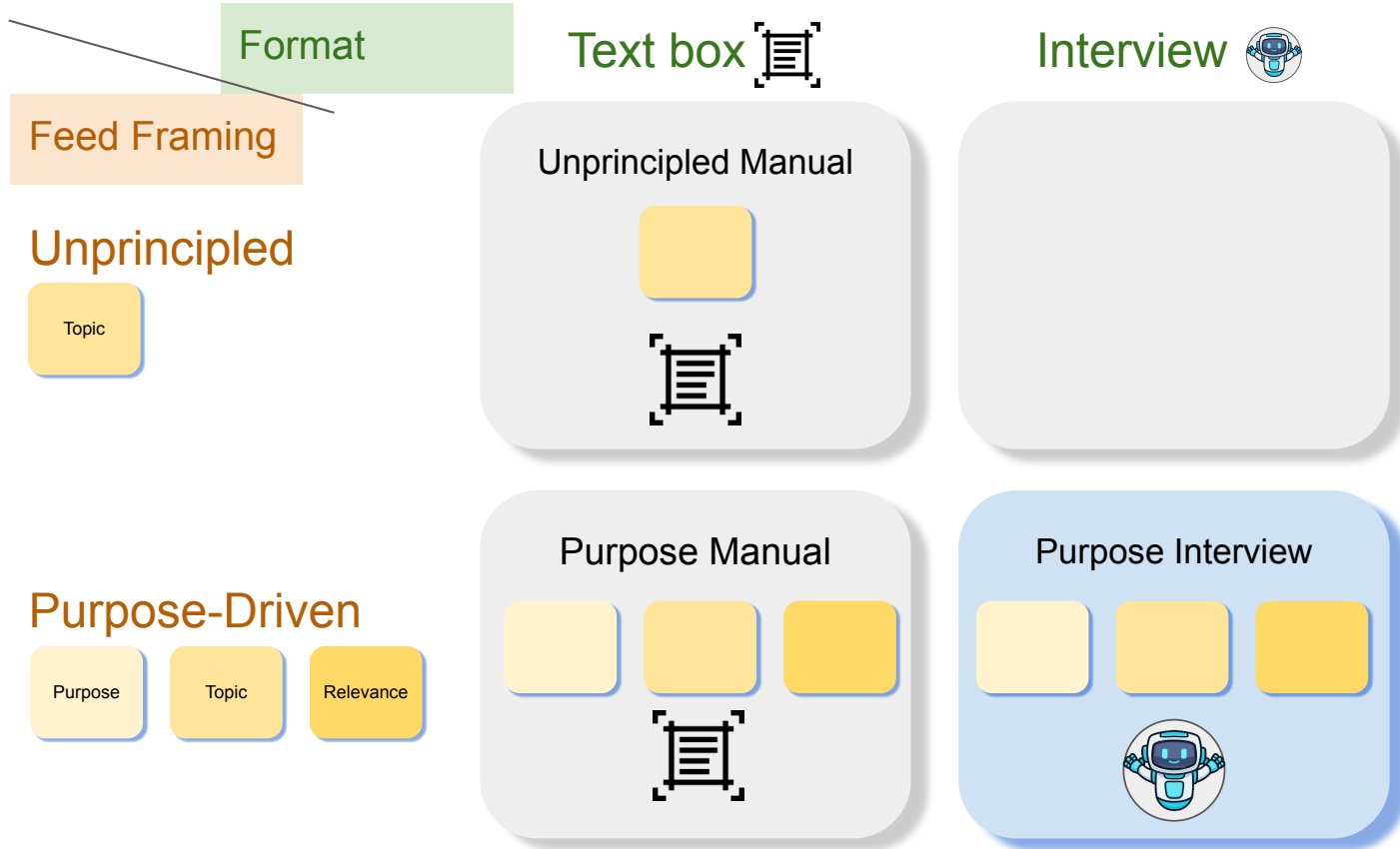
Posts are rated following 2 steps : relevance + quality



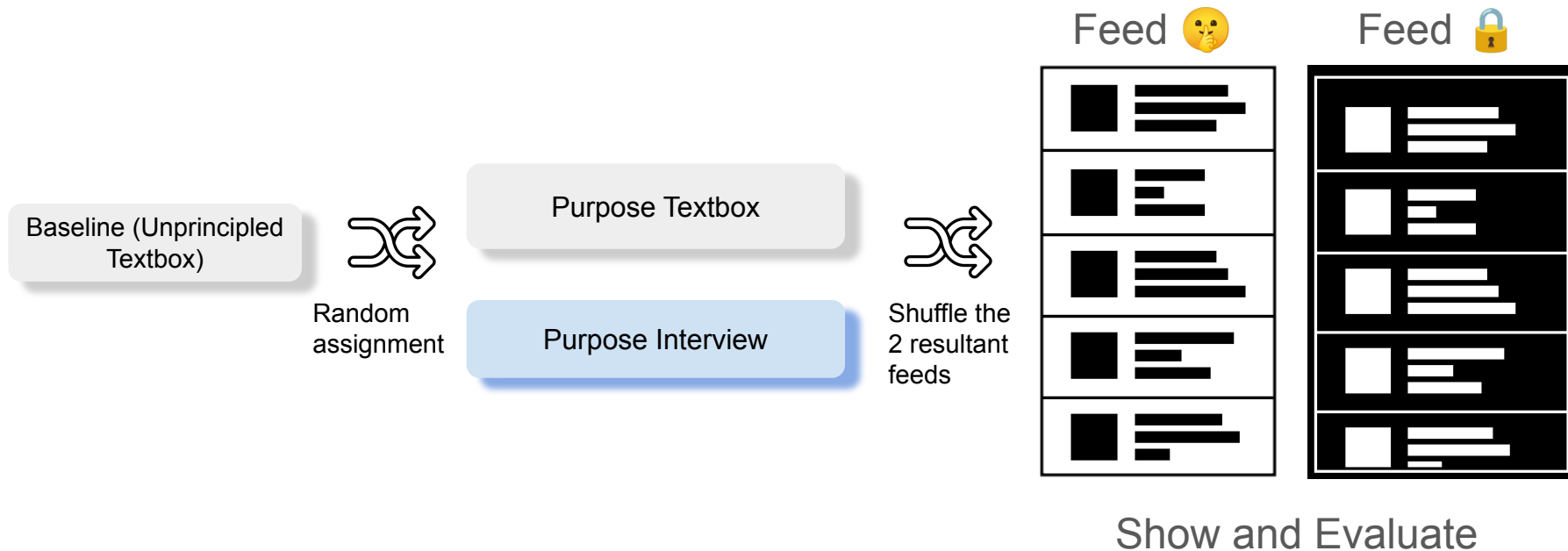
WHY?

- Information Retrieval tradition - distill a smaller model for filtering
- GPT underperform - if assign a holistic score for relevance and quality, well-written irrelevant posts would still be ranked highly

2X2 Experiment Conditions



Experiment Process



Evaluation

Quantitative

- Top 20 posts per feed → user rates: *Approve / Neutral / Disapprove*
- Mark if post was *anticipated* or *unanticipated*
- Hypothesis: ↑ unanticipated good, ↓ unanticipated bad

Qualitative

- User interviews on overall feed creating experience

Future Directions

- Vectorize each post to enable fast retrieval, replacing API-based relevance checks.
- Distill a more compact rating model instead of relying on GPT
- Run real-time feed generation experiments on existing social media platforms.
- Support other dimensions of a post beyond textual - picture/links/follower count/engagement features/if followed by user

THANK YOU

