RecSys - Past, Present, and Future (Challenges) Course

Markus Reiter-Haas, Elisabeth Lex

ESSAI Summer School 2026

July 2026

Course Website:
https://aisocietylab.github.io/recsys-ppf-course/

Part III: Beyond Optimizing Predictions

RecSys - Past, Present, and Future (Challenges) Course

Markus Reiter-Haas, Elisabeth Lex

July 2026

Course Website: *https://aisocietylab.github.io/recsys-ppf-course/*

About today's class

- Ethical challenges in Recommender Systems: biases, privacy, engagement-based optimization
- Regulatory efforts for Recommender Systems.
- Selected research results related to bias in RS and privacy-aware RS from our group.

Lecture Objectives

By the end of this lecture, you should be able to:

- Explain why RecSys create ethical risks beyond prediction error.
- Identify common sources of bias in RecSys data, models, and interfaces.
- Explain human-centered evaluation.
- Describe privacy-preserving approaches and their trade-offs.
- Relate fairness, accessibility, and regulation to RecSys design.

Why Ethics in Recommender Systems?

Ethics: principles used to evaluate whether actions, institutions, and technologies are justifiable in relation to human interests, rights, duties, and social consequences.

Recommender systems: they do not just predict preferences; they shape exposure, behavior, and opportunity.

- **Allocation:** which items, creators, products, or viewpoints receive visibility?
- **Autonomy:** can users understand, control, and disengage from recommendations?
- **Fairness:** are benefits and burdens distributed equitably across groups?
- **Privacy:** what behavioral data are collected, inferred, and reused?

The Recommender Loop: A Sociotechnical System

The Feedback Loop:

- Recommendations influence user behavior (personalized lists of items).
- User interactions are collected as new data (clicks, purchases, likes).
- New data trains the algorithm, reinforcing existing patterns.

Ethical implication:

- Feedback loops can amplify popularity, stereotypes, polarization, and compulsive engagement.
- What the system optimizes becomes what the platform reinforces.

Whose objective does the system optimize?

- **User:** relevant, useful, diverse, controllable recommendations.
- **Platform:** engagement, retention, revenue, growth.
- **Providers/creators:** fair exposure and discoverability.
- **Society:** pluralism, safety, accessibility, non-discrimination.

Challenge: optimizing one stakeholder objective can harm another.

Recommendation as Optimization

Most RecSys can be abstracted as ranking items by a learned utility score:

$$\hat{y}_{ui} = f_\theta(u, i, c)$$

$$L_u = \mathrm{TopK}_{i \in I} \ \hat{y}_{ui}$$

- u: user, i: item, c: context.
- $\hat{y}_{ui}$: predicted relevance, click probability, rating, watch time, or utility.
- Ethical question: what is this score actually optimizing?

Multi-Objective Recommender Systems

A recommender rarely optimizes only relevance:

$$\max_{\pi} \ \lambda_1 \mathrm{Rel}(\pi) + \lambda_2 \mathrm{Div}(\pi) + \lambda_3 \mathrm{Fair}(\pi) + \lambda_4 \mathrm{Priv}(\pi) + \lambda_5 \mathrm{WellBeing}(\pi)$$

- π: ranking policy.
- λ_j: design choice, business choice, or policy constraint.
- Ethical conflicts often appear as trade-offs between objectives.

How We Evaluate Also Shapes What We See

- **Offline evaluation:** uses historical logs; efficient, but inherits logging bias.
- **Online A/B testing:** captures real behavior; often optimizes short-term engagement.
- **User studies:** captures experience, control, trust, regret, and perceived harm.
- **Audits:** test subgroup disparities, exposure patterns, and compliance risks.

Ethical point: harms that are not measured are often treated as if they do not exist.

Evaluating Recommender Systems: Beyond Accuracy

- **Fairness:** Are outcomes equitable for different user/item groups?
- **Diversity:** Is the range of recommended items wide enough?
- **Novelty:** Is the user seeing items they wouldn't have found otherwise?
- **User agency/control:** Can users stop, steer, or contest recommendations?
- **Well-being:** Does optimization create regret, compulsive use, or emotional harm?

Beyond Accuracy: Diversity and Novelty

Intra-list diversity:

$$\mathrm{ILD}(L_u) = \frac{2}{K(K-1)} \sum_{i<j,\, i,j \in L_u} d(i,j)$$

Novelty:

$$\mathrm{Novelty}(L_u) = \frac{1}{K} \sum_{i \in L_u} -\log_2 p(i)$$

- $d(i,j)$: dissimilarity between two recommended items.
- $p(i)$: item popularity or exposure probability.
- Novelty penalizes repeatedly recommending already-popular items.

Case Study: The Engagement Trap

Core problem:

- Many recommender systems optimize for engagement through continuous, personalized content streams.
- Popular example: short-form video platforms
- Recommendations may be highly relevant, but still difficult to disengage from.
- This creates an **Engagement Trap** [2]: relevant content + high disengagement effort.

Ethical issue: engagement optimization can disproportionately burden users with difficulties in self-regulation (e.g., ADHD).

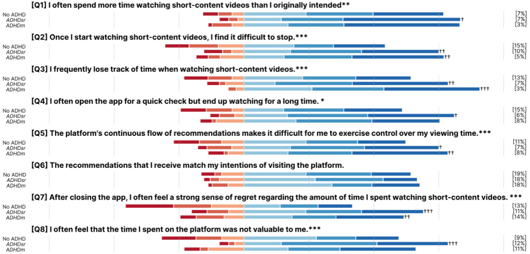
The Engagement Trap: Empirical Evidence

Study design:

- Stratified quantitative survey, $N = 302$.
- Groups: medically diagnosed ADHD, self-reported ADHD, and no ADHD.
- Context: short-form video platforms such as YouTube Shorts, Instagram Reels, and TikTok.

Findings

- Recommendation relevance did **not** differ significantly across groups.
- Users with ADHD reported significantly higher time blindness, disengagement difficulty, regret, guilt, frustration, and negative mood after use.



A Paradigm Change in RS Research

System-centered:

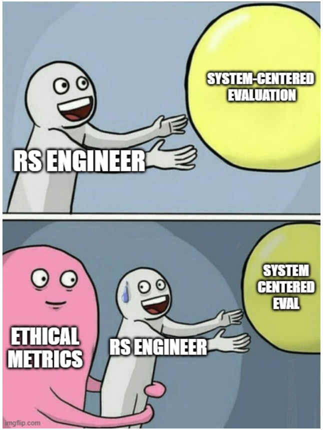
- Focus on **algorithmic performance.**
- Data-driven.
- Technical perspective.

$\Longrightarrow$

Human-centered:

- Focus on **user** and **societal benefits.**
- **Multidisciplinary** approaches (e.g., sociology, law, psychology).
- **Trustworthiness:** Ethical RS design.

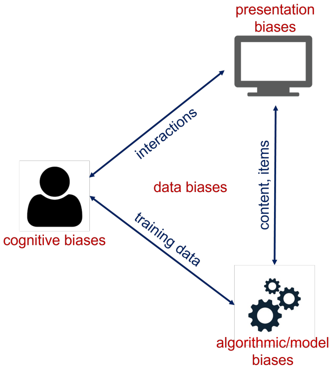
A Paradigm Change in RS Research



The Problem of Bias

Types of Bias:

- **Data Bias:** Reflects societal inequalities in training data.
- **Algorithmic Bias:** Reinforces existing stereotypes or patterns.
- **Presentation Bias:** User perception shaped by ranking and design.



Source: Schedl et al. [4]

Types of Bias in Recommender Systems

- **Selection bias:** training data reflects only observed interactions, not all preferences.
- **Popularity bias:** popular items receive more exposure and become even more popular.
- **Exposure bias:** users can only interact with items they were shown.
- **Presentation bias:** rank position affects clicks and perceived relevance.
- **Stereotyping bias:** user groups receive narrowed or essentialized recommendations.
- **Feedback-loop bias:** model outputs become future training data.

Example Bias Metrics: Disparity and Popularity Lift

Group performance gap:

$$\mathrm{RecGap}(G_a, G_b) = \mathrm{Metric}(G_a) - \mathrm{Metric}(G_b)$$

Popularity lift:

$$\mathrm{PopLift} = \frac{\frac{1}{|U|K} \sum_u \sum_{i \in L_u^K} \mathrm{pop}(i)}{\frac{1}{|I|} \sum_{i \in I} \mathrm{pop}(i)}$$

- RecGap measures unequal recommendation quality across groups.
- Popularity lift measures over-representation of popular items.

Case Study: Accessibility Bias in Recommender Systems

Fairness (well-studied)

- What gets recommended
- To whom & how often
- e.g., statistical parity, bias

Accessibility & Inclusion (underexplored)

- Can users perceive, understand, and act on recommended items?
- Barriers: language/reading level, assistive tech compatibility, geography

From *"what is recommended"* to ***"can users actually use it?"***

Accessibility and Inclusion in Recommender Systems

Image credit: *https://mobisoftinfotech.com/resources/blog/ui-ux-design/accessibility-in-ux-design*

Example: Accessibility in POI Recommender Systems [1]

Problem. POI RecSys shape how people discover places, yet they rarely model accessibility. Missing labels behave like *hidden barriers* for wheelchair users, especially outside dense urban areas.

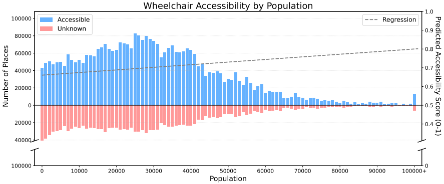
Research Questions

- How does the availability of wheelchair-accessibility metadata vary across urban vs. rural contexts?
- How do these disparities propagate into model bias in POI-RS?

Accessibility Metadata and POI Recommender Systems [1]

Urban-Rural Coverage Disparities

- Pos. association between county population and likelihood a POI is labeled accessible.
- Rural counties show markedly higher rates of **unknown** accessibility.
- For wheelchair users, unknown often functions as effectively inaccessible.



Takeaway ⇒ Sparse accessibility labels in rural settings are not only an equity issue in maps data, but a *systematic driver* of model bias in POI-RS.

Another Ethical Issue in Recommender Systems: Privacy

- Behavioral data can reveal sensitive traits, interests, and vulnerabilities.
- Privacy-preserving methods: data minimization, federated learning, anonymization, and differential privacy.

Differential privacy:

$$\Pr[M(D) \in S] \leq e^{\varepsilon} \Pr[M(D') \in S]$$

For neighboring datasets D and D' differing in one user's data, the mechanism M should produce nearly indistinguishable outputs. Smaller ε means stronger privacy.

Case Study: ReuseKNN

Problem: UserKNN recommends items using the ratings of a target user's nearest neighbors:

$$\hat{r}_{ui} = \frac{\sum_{v \in N_k(u)} sim(u,v)\, r_{vi}}{\sum_{v \in N_k(u)} |sim(u,v)|}$$

- Neighbor ratings can leak information about other users.
- Standard differential privacy protects ratings by adding noise, but often reduces recommendation accuracy.
- **ReuseKNN** [3] identifies small, highly reusable neighborhoods so that only frequently reused neighbors require DP protection.

ReuseKNN: Selecting Reusable Neighbors

Idea: ReuseKNN modifies neighborhood selection by considering both user similarity and the privacy benefit of reusing neighbors.

For a target user u and item i, the k neighbors are selected as:

$$N_{u,i}^{k} = \arg\max_{c \in U_i}^{k} \left[\text{ranking}(sim(u,c)) + \text{ranking}(reusability(c \mid u)) \right]$$

- U_i: users who rated item i.
- $sim(u,c)$: similarity between target user u and candidate neighbor c.
- $reusability(c \mid u)$: how useful candidate c is under the chosen ReuseKNN reuse strategy.

ReuseKNN: Reuse Strategies

ReuseKNN adds a reusability term to neighbor selection.

Expect: global reuse

$$\mathrm{Expect}(c) = \sum_{i \in I_c} \frac{|U_i|}{|U|}$$

Uses candidates who rated broadly useful, popular items.

Gain: user-specific reuse

$$\mathrm{Gain}(c \mid u) = \frac{|I_c \cap I_u|}{|I_u|}$$

Uses candidates whose ratings overlap with the target user.

Accountability in Recommender Systems

- The ability to attribute decisions and outcomes to specific actors (designers, operators, data providers).
- Mechanism to redress when harm occurs due to system failure or bias.

Elements of Accountability:

- **Auditability:** Systems must be designed for external review and verification.
- **Documentation:** Clear records of system design choices, data used, and performance evaluations.
- **Governing Body:** Having an internal or external body responsible for ethical oversight.

Types of Audits:

- **External Audits:** Independent third parties review the system, data, and algorithms.
- **Internal Audits:** Continuous monitoring and review by the companys own ethics or audit teams.
- **Data Audits:** Analyzing the composition and potential biases within the training/operational data.

Trustworthy AI

Do companies care about ethical guidelines?

TU Graz

Regulation as a Driver for Trustworthy AI

Self-regulation (ethical guidelines by companies) has proven insufficient due to conflicting business incentives (relevance vs. fairness).

The Need for Regulation:

- *Standardization*: Provides a common set of enforceable rules (e.g., for safety, data quality).
- *Mandatory Requirements*: Ensures organizations adopt fairness, privacy, and transparency measures regardless of cost/convenience.
- *Public Trust*: Restores confidence in the deployment of AI systems.

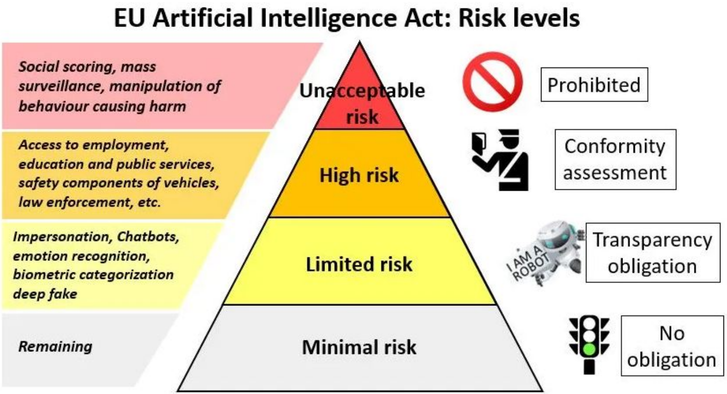
Trustworthy AI: Regulatory Efforts

EU regulatory frameworks

- **GDPR:** data protection, consent, profiling, automated decision-making rights.
- **Digital Services Act:** transparency obligations for online platforms and recommender systems.
- **EU AI Act:** risk-based obligations for AI systems, especially high-risk use cases.
- **European Accessibility Act:** accessibility requirements for many digital products and services.

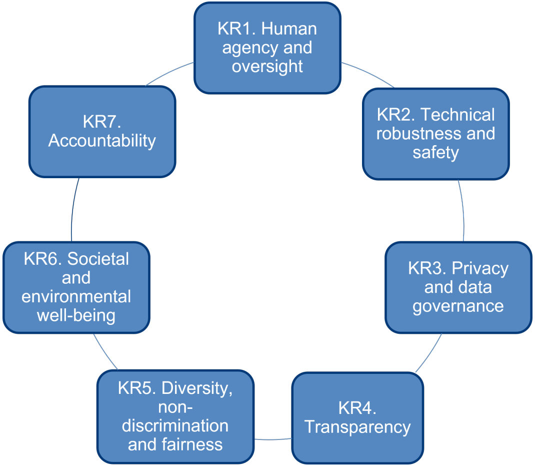
RecSys implication: regulation increasingly targets not only data use, but also transparency, user choice, auditability, and accessibility.

Trustworthy AI: Regulatory Efforts



https://d18x2uyjeekruj.cloudfront.net/wp-content/uploads/2023/05/ai.jpg

Key Requirements for Trustworthy AI



Overview:

- Seven key requirements for Trustworthy AI.
- Cover fairness, safety, privacy, and transparency.
- Essential for non-discriminatory and ethical RS.

Implications

Does regulation hinder innovation, or does it push for more responsible, and ultimately, better systems?

Summary: Principles for Ethical RecSys Design

- **Fairness:** Minimize bias and ensure equitable access.
- **Privacy:** Use privacy-preserving techniques.
- **Accountability:** Enable audits and evaluations.
- **Human Agency:** Design systems that allow users to have meaningful control over their data and recommendations.

Takeaways

- RecSys ethics is not only about "bad" recommendations; it is about optimization objectives, feedback loops, and power.
- Accuracy and engagement can conflict with fairness, autonomy, accessibility, privacy, and well-being.
- Ethical evaluation requires subgroup analysis, user studies, audits, and multi-stakeholder metrics.
- Trustworthy RecSys design requires technical, legal, and human-centered interventions.

Discussion question: What should a platform be allowed to optimize?

Questions?

References for Challenges i

[1] Gregor Mayr, Markus Reiter-Haas, and Elisabeth Lex. **"Mind the Gap: Urban-Rural Disparities in Wheelchair Accessibility for POI Recommendations".** In: *RecSoGood workshop at ACM RecSys 2025*. 2025.

[2] Vedad Misirlic, Gregor Mayr, and Elisabeth Lex. **"Quantifying the Engagement Trap: Impact of Short-form Video Recommender Systems on Users with ADHD".** In: *ASSETS '26: The 28th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM. 2026.

[3] Peter Müllner et al. **"ReuseKNN: Neighborhood reuse for differentially private KNN-based recommendations".** In: *ACM Transactions on Intelligent Systems and Technology* 14.5 (2023), pp. 1–29.

References for Challenges ii

[4] Markus Schedl, Vito Walter Anelli, and Elisabeth Lex. ***Technical and Regulatory Perspectives on Information Retrieval and Recommender Systems: Fairness, Transparency, and Privacy.*** Springer, 2025.