

Gamers help highlight disparities in algorithm data

September 29 2023, by Louis DiPietro



Credit: Unsplash/CC0 Public Domain

Is The Witcher immersive? Is The Sims a role-playing game?

Gamers from around the world may have differing opinions, but this

diversity of thought makes for better algorithms that help audiences everywhere pick the right games, according to new research from Cornell, Xbox and Microsoft Research.

With the help of more than 5,000 [gamers](#), researchers show that predictive models, fed on [massive datasets](#) labeled by gamers from different countries, offer better personalized gaming recommendations than those labeled by gamers from a single country.

The team's findings and corresponding guidelines have broad application beyond gaming for researchers and practitioners who seek more globally applicable data labeling and, in turn, more accurate predictive artificial intelligence (AI) models.

"We show that, in fact, you can do just as well, if not better, by diversifying the underlying data that goes into predictive models," said Allison Koenecke, assistant professor of information science in the Cornell Ann S. Bowers College of Computing and Information Science.

Koenecke is the senior author of "[Auditing Cross-Cultural Consistency of Human-Annotated Labels for Recommendation Systems](#)," which was presented at the [Association for Computing Machinery Fairness, Accountability, and Transparency \(ACM FAccT\) conference](#), in June.

Massive datasets inform the [predictive models](#) behind recommendation systems. The model's accuracy depends on its underlying data, especially the proper labeling of each individual piece within that massive trove. Researchers and practitioners are increasingly turning to crowdsourced workers to do this labeling for them, but [crowdsourced workforces tend to be homogenous](#).

During this data-labeling phase, cultural bias can creep in and, ultimately, skew a predictive model intended to serve global audiences,

Koenecke said.

"For the datasets used in algorithmic processes, someone still has to come up with either some rules or just some general idea of what it means for a data point to be labeled in some way," Koenecke said.

"That's where this human aspect comes in, because humans do have to be the decision makers at some point in this process."

The team surveyed 5,174 Xbox gamers from around the world to help label gaming titles. They were asked to apply labels like "cozy," "fantasy," or "pacifist" to games they had played, and to consider different factors, such as whether a title is low or high complexity, or the difficulty of the game controls.

Some game labels—like "zen," which is used to describe peaceful, calming games—were applied consistently across countries; others, like whether a game is "replayable," were applied inconsistently. To explain these inconsistencies, the team used [computational methods](#) to find that both [cultural differences](#) among gamers and translational and linguistic quirks of certain labels contributed to labeling differences across countries.

The researchers then built two models that could predict how gamers from each country would label a certain game—one was fed [survey data](#) from globally representative gamers, and the second used survey data from only U.S. gamers. They found that the model trained on labels from diverse global populations improved predictions by 8% for gamers everywhere when compared to the other model trained on labels from just American gamers.

"We see improvement for everyone—even for gamers from the U.S.—when the [training data](#) is shifted from being entirely U.S.-centric to being more globally representative," Koenecke said.

In addition to their findings, researchers crafted a framework to guide fellow researchers and practitioners on ways to audit underlying data labels to check for global inclusivity.

"Companies tend to use homogeneous data labelers to do their data labeling, and if you're trying to build a global product, you'll run into issues," Koenecke said. "With our framework, any academic researcher or practitioner could audit their own underlying data to see if they might be running into issues of representation via their data labels or choices."

More information: Rock Yuren Pang et al, Auditing Cross-Cultural Consistency of Human-Annotated Labels for Recommendation Systems, *2023 ACM Conference on Fairness, Accountability, and Transparency* (2023). [DOI: 10.1145/3593013.3594098](https://doi.org/10.1145/3593013.3594098)

Provided by Cornell University

Citation: Gamers help highlight disparities in algorithm data (2023, September 29) retrieved 25 February 2024 from <https://techxplore.com/news/2023-09-gamers-highlight-disparities-algorithm.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.