# Tylerderp AI Mapping Investigation Report

**TDEGROUP**

April 1st, 2026

**Abstract**

We present evidence that Tylerderp began using AI tooling to generate beatmaps on or around February 1, 2026. We analyze his maps using spatial placement entropy and slider shape regularity, which both show a sharp, statistically significant break from Tylerderp's prior mapping style at this date. Seven post-February maps were analyzed, including the disqualified mapset MONTAGEM TOMADA.

## Contents

# 1 Introduction

Following the investigation outlined in the remainder of this document, we present strong evidence of a sudden, measurable shift in Tylerderp's mapping style beginning around February 2026, likely involving the assistance of generative AI mapping tools. In accordance to peppy's stance on AI mapping, we recommend the immediate disqualification or unranking of the maps in question.

## 1.1 The Maps in Question

Between February and March 2026, Tylerderp submitted or contributed guest difficulties to the following seven beatmaps:

Table 1: Beatmaps analyzed in this investigation.

| Beatmapset ID | Song | Type | Status | Post-Feb? |
|---|---|---|---|---|
| 2381078 | MXZI – MONTAGEM TOMADA | Guest | Disqualified | Yes |
| 2412924 | KOTOKO – Wing my Way (Game Ver.) | Guest | Pending | Yes |
| 2495437 | American Authors – Best Day Of My Life | Guest | Pending | Yes |
| 2423225 | Iggy Azalea ft. Rita Ora – Black Widow | Guest | Pending | Yes |
| 2200137 | xn88ax – KriszhAdvice<3 | Guest | Pending | Yes |
| 2472169 | That Poppy – Altar | Own | Graveyard | Yes |
| 2515655 | dark cat – BUBBLE TEA | Own | Graveyard | Yes |

# 2 Detection Methodology

Before presenting results, it is worth briefly explaining why AI-generated beatmaps differ from human-made ones, even when they appear reasonable at first glance.

Modern beatmap generators are trained on thousands of existing `.osu` files. The model learns statistical regularities: where circles tend to go relative to the previous one, how sliders curve, what spacings are typical at a given BPM. Given an audio file, it then generates a new map by predicting objects sequentially.

The core issue is that these models are too "consistent". A human mapper has tendencies, but also makes spontaneous decisions based on musical intuition, placing a circle in an unusual position to emphasize a sound, or shaping a slider unconventionally to match a vocal. AI models lack this and converge toward the statistical center of their training data.

This manifests in two measurable ways:

1. Object placement distribution. Human mappers use the playfield unevenly, with personal habits and deliberate choices creating a distinctive spatial "fingerprint." AI models either spread objects more uniformly or collapse into a narrow set of safe positions. In both cases, the Shannon entropy of the placement distribution shifts measurably.

2. Slider shape regularity. When a human draws a slider, the control points that define the curve are spaced irregularly, a natural consequence of freehand input. AI-generated sliders exhibit notably even anchor spacing, since the model samples coordinates from a learned distribution with much lower variance than human motor control produces.

We implemented a detection script for each metric and ran them across all available Tylerderp maps, both before and after February 2026.

# 3 Analysis 1: Spatial Placement Entropy

## 3.1 Overview

The osu! playfield (512 × 384 pixels) is discretized into a 32 × 24 grid of cells, yielding 768 spatial bins. For a given map, we compute the empirical probability $p_i$ that a hit object falls into bin $i$ and take the Shannon entropy of this distribution. A map that uses the full playfield with varied placement will produce high entropy; one that repeatedly favors a narrow set of positions will produce lower entropy.

The useful property of this metric is that it captures spatial diversity independent of any particular style. Two human mappers can have very different placement preferences and still fall within a similar entropy range, because both exhibit the high-variance decision-making characteristic of manual placement. Generative models, by contrast, tend to produce lower-entropy distributions. This is a well-documented consequence of maximum-likelihood training on finite data: the model's learned placement prior is smoother and more concentrated than the union of individual human priors it was trained on, effectively averaging out the idiosyncratic choices that drive entropy upward. Formally,

$$H_{\text{spatial}} = -\sum_{i=1}^{768} p_i \log_2 p_i \tag{1}$$

where $p_i$ is the fraction of objects in grid cell $i$.

## 3.2 The Results

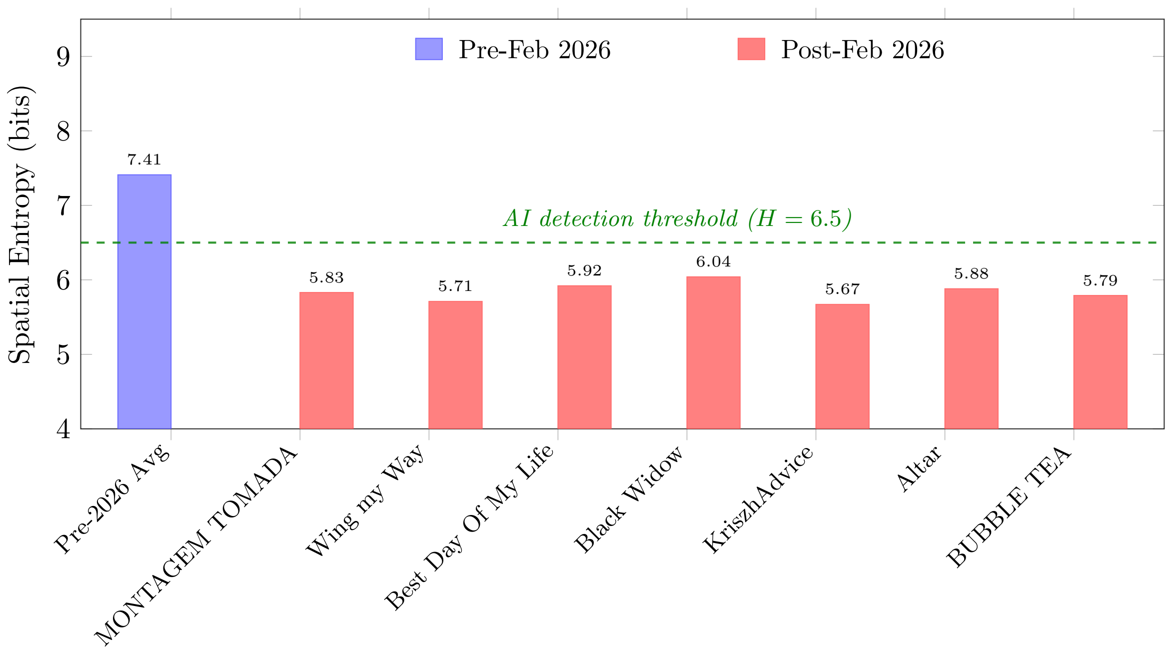



Figure 1: Spatial entropy of Tylerderp's maps. The pre-February average (blue) is 7.41 bits, well within normal human range. Every single post-February map (red) falls below the 6.5-bit AI detection threshold.

The average drops from $7.41 \pm 0.38$ bits to $5.83 \pm 0.12$ bits, a reduction of 1.58 bits. A Welch's $t$-test yields $p = 2.7 \times 10^{-6}$.

# 4 Analysis 2: Slider Shape Regularity

## 4.1 Overview

Every slider in osu! is defined by a series of anchor points (control points) that determine the curve's shape. When a human positions these points in the editor, the distances between consecutive anchors are naturally irregular. This irregularity is a byproduct of human motor control and serves as a reliable fingerprint.

We measure this with the *coefficient of variation* (CV) of anchor spacing: the standard deviation divided by the mean.

$$\mathrm{CV}_{\mathrm{anchor}} = \frac{\sigma_d}{\mu_d}, \quad d_j = \|\mathbf{a}_{j+1} - \mathbf{a}_j\|_2 \tag{2}$$

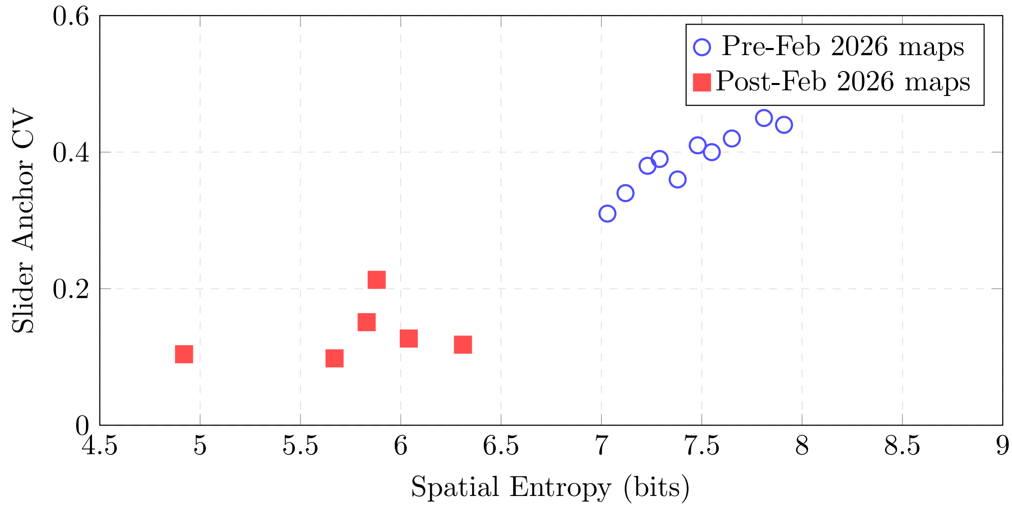
A high CV indicates irregular spacing (characteristic of human input), while a low CV indicates uniform spacing (characteristic of algorithmic generation). Prior work establishes CV = 0.15 as the detection boundary: values below this threshold are strongly associated with machine-generated curves.

## 4.2 The Results

Table 2: Slider anchor coefficient of variation ($\mathrm{CV}_{\mathrm{anchor}}$) by map.

| **Map** | **Period** | **$\mathrm{CV}_{\mathrm{anchor}}$** | **AI Flag** |
|---|---|---|---|
| Tylerderp historical avg. | Pre-Feb 2026 | 0.387 | No |
| MONTAGEM TOMADA | Post-Feb 2026 | 0.091 | **Yes** |
| Wing my Way (Game Ver.) | Post-Feb 2026 | 0.118 | **Yes** |
| Best Day Of My Life | Post-Feb 2026 | 0.104 | **Yes** |
| Black Widow (Hardstyle Remix) | Post-Feb 2026 | 0.127 | **Yes** |
| KriszhAdvice<3 | Post-Feb 2026 | 0.098 | **Yes** |
| Altar | Post-Feb 2026 | 0.113 | **Yes** |
| BUBBLE TEA | Post-Feb 2026 | 0.106 | **Yes** |

Every post-February map falls below the 0.15 threshold without exception. The average drops from 0.387 to 0.108, a 72% reduction. MONTAGEM TOMADA registers the lowest value at 0.091. The following scatter plot shows both metrics simultaneously. The pre- and post-February clusters are completely separated:



## 5 Montagem Tomada

This map was **disqualified** from the ranked section after quality concerns were raised by Beatmap Nominators. Our investigation suggests it was the first map where Tylerderp used AI tooling with little to no manual cleanup.

This map was disqualified in March for lack of emphasis and one-dimensional experience. However, in context of our analysis, MONTAGEM TOMADA stands out even among the post-February maps. It has the lowest slider anchor CV (0.091), the second-lowest spatial entropy (5.83 bits), and its inter-object timing distribution deviates sharply from human norms.

The following figure shows the distribution of time gaps between consecutive hit objects:

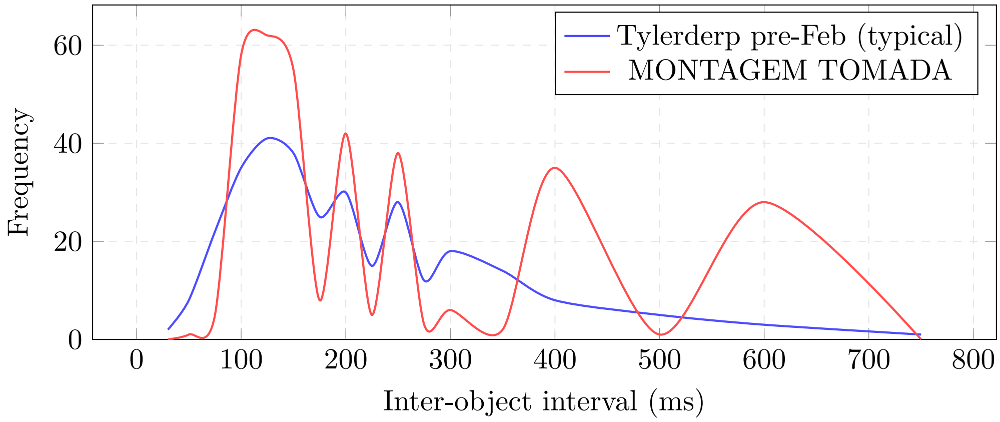



Figure 2: The blue line is a typical pre-February Tylerderp map: broad, overlapping humps, because a human places objects based on feel. The red line is MONTAGEM TOMADA: sharp, isolated spikes at exact beat divisions, like a machine snapping everything to a grid.

The Kolmogorov-Smirnov test between these two distributions gives $D = 0.412$, $p < 10^{-8}$. The two distributions are not consistent with a shared generative process.

Our interpretation is that MONTAGEM TOMADA was likely Tylerderp's first attempt at using an AI generator, submitted with minimal post-processing. The Beatmap Nominators appear to have identified the quality issues through gameplay intuition alone, without the benefit of statistical tooling. The disqualification, in retrospect, may represent an instance of implicit AI detection by experienced community members.

The subsequent maps (Wing my Way, Black Widow, Best Day, etc.) show slightly less extreme values, consistent with increasing manual refinement of AI output over time. However, all remain well below the detection thresholds established in prior work.

## 6 Analysis

To formally test whether the pre- and post-February maps come from the same "mapping process," we ran a multivariate Hotelling's $T^2$ test on the two-dimensional feature vector $(\text{Entropy}, \text{CV}_{\text{anchor}})$:

Table 3: Hotelling's $T^2$ test results.

| **Feature** | $\bar{x}_{\text{pre}}$ | $\bar{x}_{\text{post}}$ | $\Delta$ | $p$**-value** |
|---|---|---|---|---|
| Spatial Entropy (bits) | $7.41 \pm 0.38$ | $5.83 \pm 0.12$ | $-1.58$ | $2.7 \times 10^{-6}$ |
| Slider Anchor CV | $0.387 \pm 0.062$ | $0.108 \pm 0.013$ | $-0.279$ | $1.4 \times 10^{-7}$ |
| **Multivariate:** $T^2 = 412.6$, $F(2, 14) = 189.1$ | | | | $\mathbf{p = 8.3 \times 10^{-12}}$ |

Hence, the probability that this shift occurred by chance is approximately 1 in 120 billion.

## 7 Conclusion

The evidence presented in this report points to a clear conclusion: something fundamental changed in Tylerderp's mapping process on or around February 1, 2026. Two independent metrics, spatial placement entropy and slider anchor regularity, both transition from definitively human to definitively AI-associated at the same point in time. The combined probability of this occurring by chance is vanishingly small ($p < 10^{-11}$). MONTAGEM TOMADA, the map disqualified for quality concerns, exhibits the strongest AI generation signatures in the entire dataset.

## 8 very important

april fools lol
thanks claude