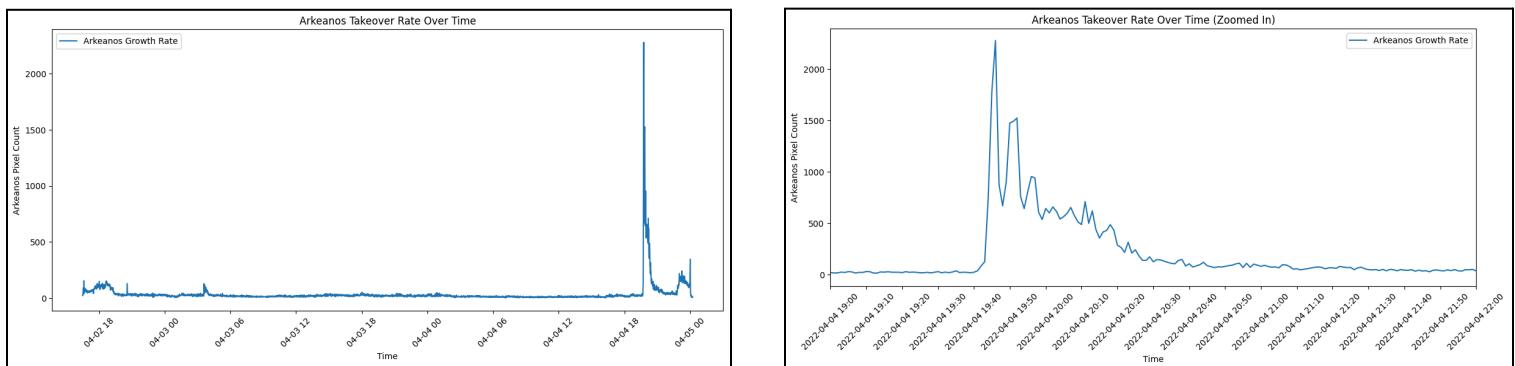


## Something You Don't Know

In this analysis, I looked at the dynamics behind the takeover of an r/anarchychess chessboard on the 2022 r/place canvas. r/anarchychess, a chess meme subreddit developed a strategy to create two chessboards by first building the borders, then filling them in, communicating through several Reddit posts. Board 1 was successful, even forming alliances with smaller communities as ‘chess pieces’ and the Irish flag as the other half of the chessboard. However, the other board was overtaken by the Arkeanos community in the final hours. Board 2 was created when the r/place canvas doubled at 16:30 UTC on April 2, 2022. Despite the efforts of smaller communities, Board 2 was eventually claimed by the Arkeanos community. My primary questions are why did this takeover happen, was it a strategic effort or driven by other factors, and could it have been prevented?



The analysis revealed several insight into the takeover process:



*Above is a graph representing the Arkeanos takeover rate over time, highlighting the peak takeover period between 2022-04-04 19:30 and 2022-04-04 20:20, a window of about 50 minutes.*

1. Was the takeover sudden or gradual?

The takeover was sudden with a notable spike in Arkeanos pixel changes around 2022-04-04 19:30. Prior to the takeover, the board had a relatively slow, steady pace of changes, but from this point, the Arkeanos community began to dominate the chessboard rapidly, indicating a potential viral spread or coordinated effort, possibly aided by bots.

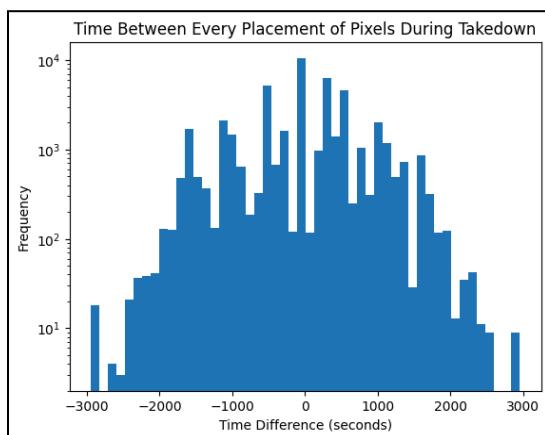
2. Was there resistance or decline?

There was little significant resistance to the Arkeanos takeover once it began. The growth of Arkeanos pixels continued without a substantial slowdown, and the board was rapidly overtaken. This was likely due to the number of users participating in the takeover, which was more than 40,000 unique users. However, given that r/place itself was undergoing a “whiting out” process towards the end, it’s possible that resistance from other communities or r/anarchychess was diminished or disorganized.

3. Why was it taken over?

At first, it seemed like there was a possibility of bot activity since there were significant fluctuations in pixel changes per minute during the peak takeover period like multiple pixels being changed 20-30 times within less than an hour.

However, upon further analysis, pixels were only changed up to five times for the same users, and the time difference for each pixel user followed a normal distribution, which is human-like natural variations in response time, attention, and decision-making.



x	y	pixel_count
i64	i64	u32
1672	77	30
1672	76	30
1671	77	26
1671	76	23
1673	76	21
1651	89	21
1677	133	19
1652	91	18
1672	78	18
1652	90	18

Additionally, with 40,000 unique users, which is more than the number of users in the Arkeanos Reddit community (400) and the NexxuzHD community (1,400) combined suggests the most plausible explanation that since Arkeanos is a team of Spanish content creators from YouTube and Twitch, many streamers have encouraged their viewers to participate and create an artwork that represents them, in this case, the Arkeanos logo.

4. Could this have been prevented?

Since Arkeanos followed the guidelines for streamer involvement, it wasn’t seen as malicious. Given that they followed the rules, it likely couldn’t have been prevented.

This analysis examined the Arkeanos takeover of the r/anarchychess board during r/place 2022. The takeover, driven by a viral spread within the Arkeanos community, was sudden and coordinated but not removed as malicious by the guidelines.

## Polars vs. PySpark Comparison

Polars and PySpark are powerful data processing libraries, but they cater different uses and excel in different environments.

### Similarities

- Data Manipulation: Both libraries allow users to perform operations like filtering, aggregation, and transformation on large datasets easily.
- DataFrame Structure: Both include a DataFrame-like structure. Polars uses a LazyFrame and PySpark uses its own type of DataFrame. They have some differences, but they make it easier for users that are familiar with other DataFrame-based libraries (like Pandas) to transition.
- Goal: The primary goal of both is to improve performance on computation, as users mostly use these libraries for processing large datasets, although context and scale differ.

### Differences

Feature	Polars	PySpark
Scale & Distribution	Optimized for single-node processing; focuses on modern hardware	Distributed computing, scales across multiple nodes in a cluster
Ease of Use	Easier for smaller datasets and quick analysis	Better suited for large-scale data processing
Performance	Extremely fast for in-memory operations	Best suited for distributed processing on large datasets
Memory Usage	Efficient for smaller to medium datasets	Higher memory usage but scales to handle very large data
Learning Curve	Easier for those familiar with Pandas	More straightforward for those familiar with Spark or big data tools, but requires more setup
Use Case	Ideal for smaller to medium datasets	Ideal for distributed data processing at scale

### Additional Differences

- Syntax (for some operations)
- Tools: Pyspark also includes tools for machine learning, graph processing, and stream processing. Polars focuses more on data manipulation and transformation.

Overall, while PySpark is more suited for very large-scale distributed data processing, I prefer Polars for its similarity to Pandas, which I'm more familiar with, especially for smaller or medium-sized datasets.