Proceedings Paper:
Pedrassoli Chitayat, Alan (2025) AI vs. the Algorithm: Measuring Success on Twitch. In: 2025 IEEE Conference on Games (CoG). IEEE Conference on Games, 26-29 Aug 2025, Técnico Innovation Center, Instituto Superior Técnico. IEEE Conference on Computational Intelligence and Games. , PRT.

https://doi.org/10.1109/CoG64752.2025.11114409

AI vs. the Algorithm: Measuring Success on Twitch

Alan Pedrassoli Chitayat
University of York
York, UK
alan.pchitayat@york.ac.uk

***Abstract*—Games livestreaming has become an invaluable tool for game studios, supporting game discoverability, community building, and community management. Understanding the different forms of success on livestreaming platforms such as Twitch, along with the relevant metrics and target benchmarks, is crucial for maximising engagement. Similarly, gaining insight into how short-term success influences long-term performance can empower studios to strategically plan, design, and implement future content or game releases. However, the diverse ways in which games perform on Twitch present challenges for detailed analysis. To address this, this paper applies unsupervised machine learning techniques to identify and present seven archetypes of success, enabling studios to gain deeper insights into their game's performance.**

***Index Terms*—games livestream, clustering, archetype analysis, metrics, twitch**

I. Introduction

Games livestreaming is one of the fastest-growing streaming markets and a major sector of the gaming industry [1, 2, 3].

As part of its widespread popularity, games livestreaming has become a crucial tool for game studios, enhancing game discoverability, fostering community building, and providing an interactive environment for incorporating player feedback into future development [4, 5].

The academic literature has examined several aspects of what makes livestreaming a compelling media format. However, most studies have primarily focused on audience experience and the challenges faced by streamers [6, 7, 8]. In contrast, understanding the different ways in which games perform within the vast livestreaming ecosystem remains a complex task.

Streamers often adapt their content to maintain audience engagement, cycling through different games and modifying their presentation style to attract viewers [9]. Game design elements have also been shown to influence audience interest, particularly when integrated with livestreaming platforms to enhance interactivity [7, 5].

The wide array of available metrics for tracking game and stream performance — such as those provided by platforms like *TwitchTracker*[1] and *TwitchMetrics*[2] — further complicates the interpretation of success. It is not always clear which metrics are the most relevant for a given game title or how multiple metrics interact, making in-depth analysis less accessible to studios. This challenge is particularly significant for smaller studios, which may lack the resources to conduct detailed analyses while simultaneously maintaining their game and engaging with their community.

This paper examines nearly nine years of data on the most successful games on Twitch to classify distinct archetypes of success. By understanding how games perform over time, this classification provides insights into how short-term performance influences long-term outcomes.

The contribution of this paper is twofold.

Firstly, it introduces a set of archetypes that define the different ways games succeed on Twitch. These archetypes offer game studios a structured tool for understanding their game's performance on one of the most popular livestreaming platforms, using commonly available metrics. Additionally, as a result of the archetype design, each classification provides a set of target values for different types of success. This allows studios to identify which areas of livestreaming performance need improvement to transition between archetypes. The archetype model, along with the necessary steps for its application, is made freely available[3] to enable further study and to allow studios to classify their own games without restrictions.

Secondly, the paper analyses how games perform over extended periods, providing a deeper understanding of how transitions between archetypes unfold over several years. By identifying common trends in the data, this analysis empowers studios to better interpret their game's trajectory and the underlying patterns that shape its success on Twitch.

II. Literature Review

Games livestreaming has grown to become a major sector of the gaming industry [3], with some researchers referring to livestreaming platforms, such as Twitch, as "the future of gaming" as early as 2015 [10]. This rapid growth has contributed to increasing academic interest in the domain.

One example of this interest is the work by Herrewijn and Charleer [4], which describes an iterative design process for enhancing viewer customisation of their livestream experience. Through a series of questionnaires, the authors identified the

This research was supported by the UKRI Arts and Humanities Research Council funded CoSTAR Live Lab project (AH/Y001079/1).

[1] https://twitchtracker.com/

[2] https://www.twitchmetrics.net/

[3] https://github.com/ChitaAPC/ArchetypalSuccessOnTwitch

importance of enabling viewers to discover and navigate different game titles on Twitch. This study focused on designing interfaces that improve the viewing experience [4].

In a related study, Herrewijn and Charleer [9] explored the intersection between viewer and streamer needs, emphasising the importance of streamers customising their experience. The authors highlighted that streamers must make informed decisions about which games to stream, a finding that aligns with their previous viewer-focused study [4]. These studies exemplify academic efforts to improve both the viewer and streamer experience through better design.

From a game development perspective, some academic work has examined specific game design aspects that enhance livestream suitability. For example, Glickman et al. [5] identified key design challenges for livestream-integrated games, particularly regarding audience participation. This study explored different ways in which games can be designed to interact with livestreaming, outlining usability concerns related to audience influence on gameplay. The authors noted that such interactive features can be received positively or negatively depending on streamer community size.

Similarly, Lessel et al. [7] investigated viewer perceptions and preferences in games livestreaming, identifying that interactive features - especially those embedded within the game itself - enhance the viewing experience, provided they do not overly interfere with streamer performance. Importantly, this applied to both active viewers, who engage with the stream via live chat or social media, and passive viewers, who prefer a more observational experience.

In addition to interactivity, Lessel et al. [7] found that game-specific overlays providing real-time explanations and additional information were among the most well-received features. Such overlays have been explored in esports environments, either integrated directly into the stream [11, 12], presented as third-party interactive dashboards [13, 14], or developed as companion apps for viewers to use alongside streams [15].

Notably, many of these dashboard systems utilise artificial intelligence (AI) techniques to enhance storytelling [11, 12, 15], such as win prediction models [16]. For example, Demediuk et al. [17] employed unsupervised learning techniques, including K-means or Mean Shift clustering, to automatically detect different player roles in esports teams. Similarly, Pedrassoli Chitayat et al. [18] used clustering techniques to analyse character representation in a constantly evolving esports environment. By clustering several years of patch-note data, the study provided a structured representation of in-game characters, enabling machine learning (ML) models to maintain performance across future patches.

In these cases, AI and ML have been leveraged to enhance dataset interpretation, facilitating deeper insights into game environments. A particularly relevant approach for this purpose is archetype analysis [19, 20, 21], including archetypal-based clustering [22]. Archetype analysis [23] focuses on identifying the most extreme yet representative data points. K-Maxoids clustering [22], an archetypal variation of K-Means, is an example of this approach, where centroids are selected based on maximising their distance from other nodes in the cluster. As a result, cluster centroids represent extreme cases, making them particularly relevant for analysis.

For instance, Demediuk et al. [19] used archetypes to detect player roles in live esports matches. By comparing in-game values to cluster centroids, the authors developed a performance index to assess individual players in a live environment. Similarly, archetypal analysis has been used to detect anomalies in live environments, such as identifying key moments in esports matches for enhanced storytelling [20] or recognising anomalous player behaviours, including cheating, in live service games [21].

Overall, much of the existing literature has focused on the interaction between streamer and viewer, either from a sociological and interactive media perspective [4, 9] or from a game development standpoint [5, 7]. Additionally, researchers have highlighted the significant role AI plays in games livestreaming, whether for data-driven storytelling [15, 20] or for analytical purposes [21].

Despite these advancements, no existing studies could be found that explore the performance of games themselves within the livestreaming domain. This paper addresses this gap by employing archetypal clustering techniques (K-Maxoids [22]) to examine how different games perform and succeed on Twitch.

III. Methodology

A. Clustering

As discussed in Section I, this study aims to classify different types of success on Twitch. To achieve this, archetypal clustering - specifically K-Maxoids [22] - was employed.

Clustering techniques have been widely used in the literature to extract meaningful insights from large, complex datasets [17]. They have also been shown to improve the effectiveness of data analysis [18], enabling models to identify and model patterns that persist over time.

Beyond classification, this study aims to provide developers and studios with actionable insights into how different games succeed on Twitch. Archetypal analysis is particularly relevant for this purpose, as it identifies extreme yet representative cases. By using this approach, the study can define meaningful performance benchmarks for different game categories, similar to its use in esports performance indexing [19]. These archetypes can then be used as an analysis tool for studios and researches in order to understand complex patterns and the interaction between game design, marketing strategies and game performance on livestream platforms.

B. Dataset

This study utilises the Kaggle dataset “Top games on Twitch 2016 - 2023” [24], specifically the “twitch game data” CSV file. The dataset contains Twitch performance metrics for the top 200 ranked categories each month. At the time of writing, the dataset (version 17) includes monthly rankings from January 2016 to September 2024 (105 months).

TABLE I
DATASET FEATURE DESCRIPTIONS

Feature Name	Description
Rank	Monthly ranking of the category, from 1 (highest) to 200 (lowest in dataset)
Game	Unique identifier for the Twitch category (typically the game title)
Month	Calendar month corresponding to the entry (1 to 12)
Year	Calendar year corresponding to the entry (2016 to 2024)
Hours_watched	Total hours watched across all streams in the category for the month
Hours_streamed	Total hours streamed in the category for the month
Peak_viewers	Maximum number of concurrent viewers in the category for the month
Peak_channels	Maximum number of concurrent streams in the category for the month
Streamers	Total number of unique streamers broadcasting in the category for the month
Avg_viewers	Average number of viewers across the month
Avg_channels	Average number of concurrent streams across the month
Avg_viewer_ratio	Average viewer-to-stream ratio (Hours_watched / Hours_streamed) for the month

While most categories correspond to specific games, some non-game categories, such as “Arts & Crafts” or “Music & DJ,” are also present. Each monthly entry includes 12 features (Table I), providing detailed performance metrics for each ranked category. In total, the dataset contains *21,000* entries, representing monthly performance across the top 200 categories for each of the 105 months.

C. Data Preprocessing

To apply the K-Maxoids algorithm, the dataset was preprocessed to ensure consistency and suitability for clustering.

Since the dataset only includes the top 200 ranked games per month, individual game data appears sporadically. Some games, such as *League of Legends*, have been ranked consistently throughout the dataset’s entire timespan. Others, like *Diablo III*, appear intermittently (e.g., 76 times despite its 2012 release). Additionally, newer games (e.g., *Overwatch 2*, released in October 2022) do not have data for earlier years.

To address this variability, a *sliding window approach* with a 12-month period was applied, as depicted in Figure 1. Each game was represented by consecutive 12-month segments, starting from its first available month. For example, *League of Legends* resulted in 9 entries, each covering January to December for different years (2016–2024), as depicted by the 9 boxes in the figure.

For games with missing months, such as *Diablo III*, absent values were set to zero. For example, in its fourth entry (2021), all values for June, September, October and November were zeroed out, as the game did not rank in the top 200 during those months. This can be observed by the empty (white) months for that game in Figure 1. Similarly, *Overwatch 2* first ranked in April 2022 during its beta phase, initiating a sliding window from April 2022 to March 2023. Furthermore, any entries with a period starting after October 2023 contain missing values, since the dataset only covers up to September 2024.

1) Handling Unknown Values: All missing values were replaced with zeros, with known values enteres as is, with the exception of *Rank*. Since higher values indicate larger quantities (e.g., “*Streamers*” = 200 means 200 unique channels streamed the game), most features naturally accommodate zero values. However, *Rank* follows an inverse scale, where 200 represents the lowest rank. To correct for this, rankings were transformed using:

$$\text{Adjusted Rank} = 201 - \text{Original Rank}$$

This ensures that higher values consistently indicate higher performance across all features.

2) Data Encoding: Each encoded entry consists of a game category, followed by 12 values (one per month) for each feature in Table I (excluding “Month” and “Year”). Notably, this encoding removes explicit temporal information, meaning the dataset represents 12-month sequences without specifying start or end dates. This approach introduces a limitation: if a game appears only twice (e.g., in January 2020 and January 2021), it will have two entries dominated by zeros, a similar behaviour can be observed in the last two entries for Diablo III depicted in Figure 1.

3) Outlier Removal: Due to extreme variance in feature values, certain entries disproportionately influenced clustering. This was particularly problematic for K-Maxoids, which selects maximally distant centroids. To mitigate this, outliers were removed before normalisation and training using a quartile-based filtering approach. Only entries within the interquartile range (IQR) (i.e. second and third quartiles) were retained.

After filtering, the dataset was reduced from *21,000* to *10,640* entries. Following the encoding described in Section III-C2, this resulted in *1,476* encoded entries, which were then used for training and testing of the K-Maxoids model.

4) Data Normalisation: Finally, the dataset was standardised using a *z-score normalisation* approach:

$$X' = \frac{X - \mu}{\sigma}$$

where μ is the mean and σ is the standard deviation for each feature. This transformation ensures all features have a mean of 0 and unit variance, improving clustering performance.

IV. Results

A. Model Training and Selection

To train the K-Maxoids model, the encoded dataset was split into training and testing sets using a *60/40%* split. The

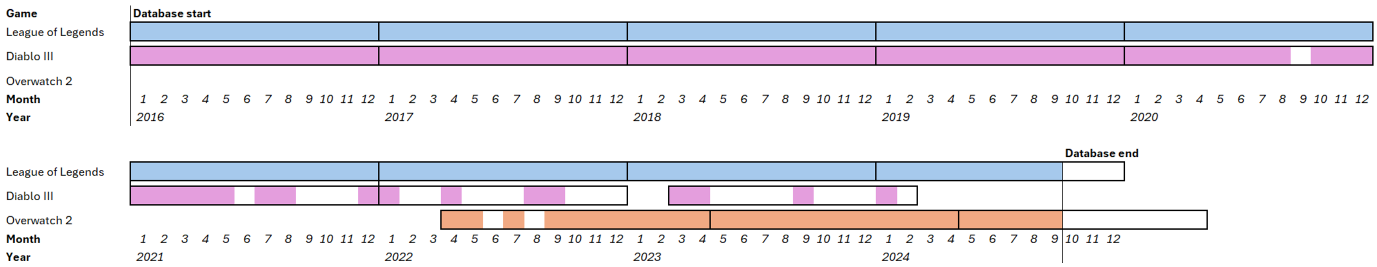

Fig. 1. Depiction of how data is available for certain games in the dataset. Sliding window periods have been highlighted with a box.

model training process followed the methodology outlined in the literature [22].

Given the time-dependent nature of the encoded dataset (Section III-C), Dynamic Time Warping (DTW) was enabled. This ensures better alignment of time-series features, improving the clustering of games with similar performance trajectories.

Model training was performed with the following hyperparameters:

- Convergence tolerance: 0.0001
- Maximum iterations: 100
- Cluster range: $K = 3$ to $K = 10$

Additionally, to aid in replicability, a random state seed of 11 was introduced both in the data split and the training process. Figure 2 presents the silhouette scores for both training and testing datasets across different values of K.

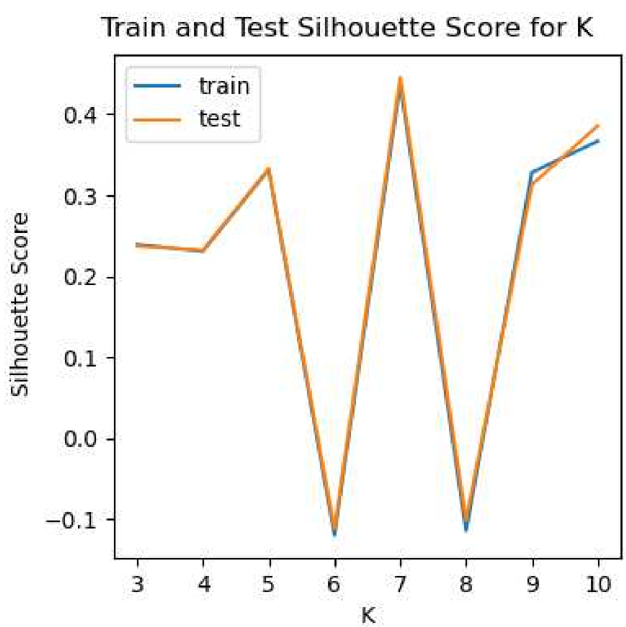

Fig. 2. Silhouette Scores for different values of K.

As depicted in Figure 2, $K = 7$ was selected as the final model, as it produced the highest silhouette score. Notably, a potential positive trend can be observed between $K = 9$ and $K = 10$. However this was not explored further because of two main factors. Firstly, a divination in training and test results can be observed from this stage, suggesting that models with higher K values may not generalise well to unseen data. Secondly, as discussed in Section V, the clusters are designed to provide game studios with understandable and actionable insights into the data, to be used as an analysis tool in a similar way as design personas describe potential users for User-Centred Design methodologies. To enable this, smaller values of K are favourable, as large numbers of archetypes may limit the understandability and application of the model.

B. Cluster Analysis

After determining $K = 7$ as the optimal number of clusters, each cluster was examined based on:

1) Centroid characteristics, capturing representative performance trends.
2) Overall game performance over the nine-year period, evaluating the long-term success patterns.

This analysis is further discussed in Section V.

To ensure robustness, the previously removed outliers (as described in Section III-C3) were reintroduced. These entries were encoded and normalised using the same mean (μ) and standard deviation (σ) values derived from the original dataset to maintain consistency in results. Table II presents the labels as well as the distribution of entries across the 7 identified clusters.

TABLE II
CLUSTER DISTRIBUTION OF THE TOTAL DATA

Cluster Label	Number of Entries
Momentum Builders	317
Supernova	569
Esports Giants	751
Replay Royalty	41
Short-lived Fame	2,411
Roller Coaster	91
Wildcard	134

The distribution of games across clusters provides insight into distinct success patterns on Twitch, which will be further explored in the following section.

V. DISCUSSION

To interpret the seven clusters, the centroids and the most representative entries - defined as those with minimal distance from the centroid - were analysed. The labels for each archetype were chosen to clearly communicate the trends observed, serving a similar role to user personas often used during the design process for a software or application. Once the clusters were characterised, this study also examined how games evolve over time, allowing for the observation of long-term trends beyond the initial classifications.

This section first details the seven clusters, outlining their common characteristics and providing meaningful descriptions. Following this, three games are analysed as case-studies, demonstrating how individual games transition between archetypes over time, beyond the 12-month period used for classification.

A. Cluster Descriptions

Each of the seven clusters is defined based on common trends observed in the most representative entries. To aid comprehension, the three most prominent features of each archetype are highlighted.

1) Momentum Builders: Games in this cluster exhibit a steady rise in popularity over a 12-month period, typically ending the year at a higher rank than where they started. These games experience a significant increase in audience demand, as indicated by a near doubling of hours watched, while the number of hours streamed remains relatively stable.

Key Characteristics (12-Month Comparison):

- Monthly Hours Streamed: Remains approximately consistent.
- Monthly Hours Watched: Roughly doubles.
- Monthly Rank: Improves by approximately 40 positions.

2) Supernova: Games in this cluster experience intense but short-lived success, attracting a large audience while maintaining a relatively low number of streamers. This results in an exceptionally high viewer ratio, as significantly more people watch the game than play it. Their popularity is concentrated in bursts lasting a few months before a sharp decline.

Key Characteristics (Measured Over a 4-Month Period):

- Monthly Average Viewer Ratio: 80+
- Monthly Hours Watched: 150,000+
- Monthly Rank: Within the top 200

3) Esports Giants: This cluster contains dominant and long-standing esports titles that have secured their place in the streaming ecosystem. These games have the highest hours streamed in the dataset, supported by a vast player base ranging from streamers with small communities to high-profile professionals. While their viewer ratio is lower than other clusters, their sheer scale enable continued visibility and engagement.

This clustered was named as most successful esport titles are consistently classified within it. However, it is important to note that some non-esport titles can also be observed in this archetype as discussed in the case-studies in this section.

Key Characteristics (Measured Over a 12-Month Period, with Monthly Targets):

- Monthly Hours Streamed: 70,000+
- Monthly Peak Channels: 300+
- Monthly Hours Watched: 120,000+

4) Replay Royalty: This cluster consists of highly replayable games that maintain a steady presence on Twitch without reaching massive viewer peaks. Many of these are single-player titles (e.g., *Fallout*, *Skyrim*) that thrive on dedicated but smaller audiences rather than widespread streamer interest.

Key Characteristics (Measured Over a 12-Month Period, with Monthly Targets):

- Monthly Hours Streamed: 50,000+
- Monthly Average Channels Streaming the Game: 70+
- Monthly Average Viewers: 400+

5) Short-Lived Fame: The largest cluster, but also the most fleeting type of success. These games experience short bursts of popularity before fading, often appearing in the top 200 twice within a year and then disappearing. They may represent entry-level success - a brief moment in the spotlight before temporally vanishing.

Key Characteristics (Measured Over a 12-Month Period, Appearing at Least Twice):

- Monthly Rank: Within the top 200
- Monthly Average Viewer Ratio: 35+
- Monthly Peak Viewers: 20,000+ at least once

6) Roller Coasters: These titles are inconsistent performers, experiencing strong months followed by gaps where they do not place in the top 200. Their success is driven more by many active streamers than just raw viewer interest. They perform well but unpredictably, making them harder to sustain over time.

Key Characteristics (Measured Over a 4-Month Consecutive Period):

- Monthly Total Streamers: 6,000+
- Monthly Peak Channels Streaming the Game: 200+
- Monthly Hours Watched: 100,000+

7) Wildcards: A more extreme version of Roller Coasters, these games experience bigger spikes and deeper drops in popularity. They swing between high peak viewership and complete disappearances, making their success highly unpredictable. Unlike Roller coasters, Wildcards depend more on viewer hype than streamer activity, leading to volatile performance patterns.

Key Characteristics (Measured Over a 4-Month Consecutive Period):

- Monthly Hours Watched: 150,000+
- Monthly Peak Viewers: 60,000+ (at least twice)
- Monthly Peak Channels Streaming the Game: 200+

B. Archetypes Over Time

This section presents the performance of three distinct games as three case-studies. These games were chosen as meaningful illustrations of general trends observed in the data, particularly with how games tend to transition from one archetype to another. While general trends are identified, and possible causes are provided, further work (as discussed in Section VI) is required to more formally determine the root causes of these transitions. Therefore, while this section offers some interpretations, continued study is needed to fully understand the underlying causes. However, the interpretations provided in the case-studies serve as an illustration of how these archetypes can be used by studios to identify powerful trends on the data and guide continued work in their game development lifecycle.

1) Case-Study: No Man's Sky:

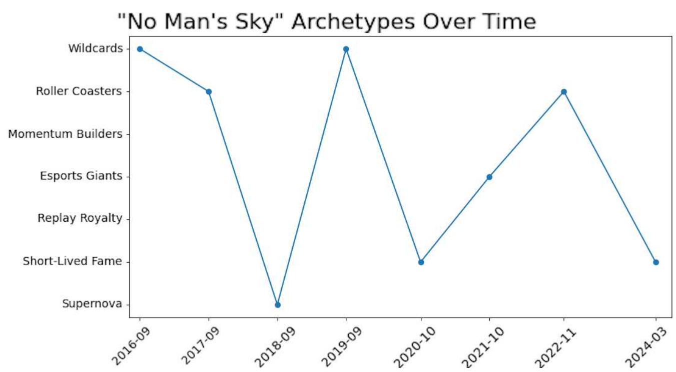

Fig. 3. No Man's Sky timeline based on the observed archetypes

No Man's Sky, a popular survival game, provides a clear example of how a title can oscillate between unpredictable archetypes over time. As shown in Figure 3, the game exhibits recurring shifts between the Roller Coaster and Wildcard archetypes, particularly during the periods beginning in September 2016, 2017, 2019, and November 2022. These fluctuations suggest that No Man's Sky undergoes bursts of renewed interest, likely driven by major updates and community engagement cycles.

A broader trend observed in the dataset is the transition from *Short-Lived Fame* to *Esports Giant*, which is evident in No Man's Sky's performance between October 2020 and 2021. Notably, this period coincided with the COVID-19 pandemic, which may have influenced the game's increased popularity. However, it also aligns with the release of three major content updates[4]:

- Origins (September 2020)
- Expeditions (March 2021)
- Frontiers (September 2021)

Each of these updates triggered a surge in streams, viewership, and overall Twitch ranking, as expected. However, in most cases, this interest was temporary, leading to a decline once players and streamers had explored the new content. Despite covering multiple significant updates, the period from October 2020 to September 2021 was ultimately categorized under the *Short-Lived Fame* archetype, highlighting how temporary surges do not always translate to sustained success.

By contrast, the Frontiers update in September 2021 had a more lasting impact. As shown in Figure 3, despite an expected drop in key metrics after the update's initial surge, the decline was not as steep as in previous cases. This sustained engagement allowed No Man's Sky to transition into the *Esports Giant* archetype.

This analysis demonstrates how archetypes can provide developers with valuable insights into a game's long-term performance. By understanding the patterns behind their game's popularity, studios can investigate the factors that contributed to periods of exceptional success - whether through community engagement, content design, or external events - and use these findings to refine future updates and marketing strategies.

2) Case-Study: Slay the Spire:

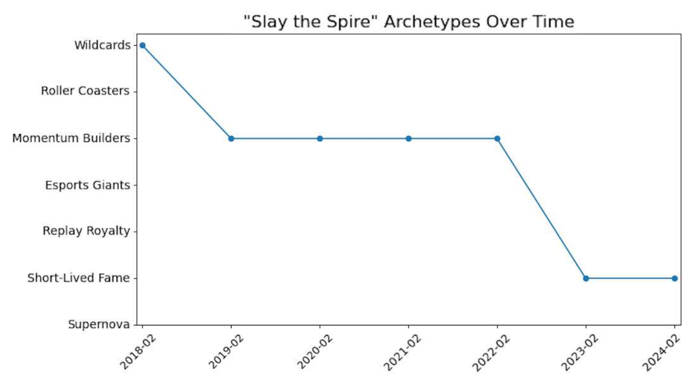

Fig. 4. Slay the Spire timeline based on the observed archetypes

Slay the Spire was selected as a case study because its developers have explicitly stated that they designed the game not only to be enjoyable to play but also engaging to watch [25]. Particular attention was given to UI and gameplay clarity to enhance the viewing experience. The impact of these design choices can be observed in the game's Twitch performance, as shown in Figure 4.

Unlike many other games that experience short bursts of success, Slay the Spire demonstrates a rare example of continuous, long-term growth on Twitch. Over a four-year period, the game maintained a steady increase in viewership, a pattern that repeatedly aligns with the *Momentum Builder* archetype. This sustained success suggests that game design choices made with streamability in mind can contribute to long-term engagement, rather than just short-term visibility.

This case-study highlights the importance of intentional design for livestreaming. Slay the Spire's success on Twitch is not merely a by-product of its gameplay mechanics but a reflection of deliberate efforts to enhance its watchability. This supports the broader argument that livestreaming can play a critical role in a game's overall success and longevity. Moreover, it reinforces the need for developers to analyse

[4]https://nomanssky.fandom.com/wiki/Patch_notes#Update_3.00_(Origins)

game performance data to refine and optimise their designs for streaming platforms.

3) Case-Study: Raft:

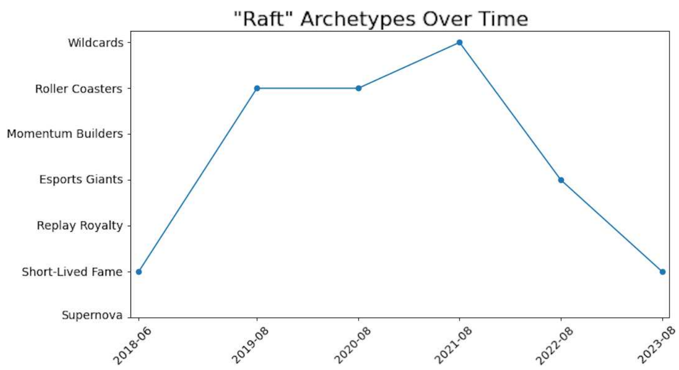

Fig. 5. Raft timeline based on the observed archetypes

Raft presents another example of how games can transition from inconsistent success to widespread popularity. As illustrated in Figure 5, Raft initially exhibited an unstable pattern, fluctuating between the *Roller Coaster* and *Wildcard* archetypes. This pattern reflects a common challenge faced by many games - gaining traction in bursts but struggling to maintain sustained engagement.

Over time, however, Raft's trajectory shifted. Eventually, the game reached a level of consistent popularity that allowed it to enter the *Esports Giant* archetype. This transition mirrors the pattern observed in No Man's Sky, further emphasizing how certain games evolve from unpredictable performance to more sustained success.

A key insight from this case study is that many of these transitions, like those shown in Figures 3 and 5, do not last beyond a one-year period. This highlights the importance of continued analysis and strategic intervention. Developers aiming to sustain long-term success must not only focus on achieving large audience spikes but also implement strategies to retain player interest beyond the initial surge. By monitoring a game's archetype evolution, studios can better understand when and why their audience engagement fluctuates, ultimately leading to more informed decisions about content updates, marketing efforts, and community engagement.

VI. Limitations & Future Work

The use of the archetypes described in this paper can provide game studios with valuable insights into their game performance on Twitch, as discussed in Section V.

However, it is important to emphasise that the work presented here should be seen as a tool for analysis rather than a definitive framework. The interpretations outlined in this paper describe possible ways in which studios might utilise the archetypes, but future studies are required to better understand the impact that studio strategies have on the streamability of games. For example, continued analysis of game performance across archetypes, particularly in as a collaboration with game studios, could provide insights into how gameplay changes influence a game's streamability and whether such changes can be predicted.

Additionally, one clear limitation of the archetype model presented is that it rely on a 12-month period to classify performance. Therefore, it cannot be used to quickly identify or predict the impact of new strategies; instead, it serves as a post-hoc analysis tool. Future research could focus on designing and training a predictive model capable of estimating a game's performance (i.e. the archetype it will belong to) within a shorter time frame.

One potential direction for future work is the development of a predictive system that uses the pre-classified games in this study as labelled data for training purposes. Such a model could attempt to classify a game into an archetype at the end of a 12-month period using only data from a shorter window, such as the most recent 2 to 4 months. This would enable studios to gain faster insights into the potential impact of changes in gameplay or marketing strategies, without needing to wait for a full year of data.

VII. Conclusion

This paper demonstrates how archetypal clustering can be used to identify different types of success for games on Twitch. By applying a 12-month sliding window to encode the dataset, a K-Maxoids model was trained that divides nine years of Twitch category data into seven distinct archetypes of success.

Through the analysis of long-term game performance via three case studies, this paper illustrates how the archetypes can serve as a meaningful analytical tool. Future research in this domain may further refine this approach, enhancing its ability to provide studios with deeper insights into game performance and streamability.

To support continued research and the practical application of these findings, the clustering model is made fully available, detailed in Section I. This allows both researchers and game studios to leverage the results presented in this paper, fostering further advancements in understanding game success on Twitch and designing games for streamability.

VIII. Acknowledgement

This research was supported by the UKRI Arts and Humanities Research Council funded CoSTAR Live Lab project (AH/Y001079/1).

References

[1] Newzoo, "Newzoo global esports market report 2019," 2019, [Online]. Available: https://newzoo.com/insights/trend-reports/newzoo-global-esports-market-report-2019-light-version/ (accessed: 2025-03-09).

[2] ——, "Newzoo's global esports & live streaming market report 2022 — free version," 2022, [Online]. Available: https://newzoo.com/resources/trend-reports/newzoo-global-esports-live-streaming-market-report-2022-free-versi (accessed: 2025-03-09).

[3] ——, "Newzoo's global games market report 2023 — may 2024 update," 2024, [Online]. Available: https://newzoo.com/resources/trend-reports/newzoo-global-games-market-report-2023-free-version (accessed: 2025-03-09).

[4] L. Herrewijn and S. Charleer, "Power to the spectator: Towards an enhanced video game stream discovery experience," in *International Conference on Computer-Human Interaction Research and Applications*. Springer, 2023, pp. 349–360.

[5] S. Glickman, N. McKenzie, J. Seering, R. Moeller, and J. Hammer, "Design challenges for livestreamed audience participation games," in *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play*, 2018, pp. 187–199.

[6] Z. Meng, Q. Gao, D. Guo, Y. Li, B. Li, H. Fei, S. Wu, F. Li, C. Teng, and D. Ji, "Mmlscu: A dataset for multimodal multi-domain live streaming comment understanding," in *Proceedings of the ACM on Web Conference 2024*, 2024, pp. 4395–4406.

[7] P. Lessel, M. Altmeyer, and A. Krüger, "Viewers' perception of elements used in game live-streams," in *Proceedings of the 22nd International Academic Mindtrek Conference*, 2018, pp. 59–68.

[8] T. Taylor, *Watch me play: Twitch and the rise of game live streaming*. Princeton University Press, 2018.

[9] L. Herrewijn and S. Charleer, "Spectator lf streamer: facilitating better stream discovery through spectator motivations and stream affordances," in *Companion Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, 2023, pp. 135–142.

[10] B. Burroughs and P. Rama, "The esports trojan horse: Twitch and streaming futures," *Journal For Virtual Worlds Research*, vol. 8, no. 2, 2015.

[11] F. Block, V. Hodge, S. Hobson, N. Sephton, S. Devlin, M. F. Ursu, A. Drachen, and P. I. Cowling, "Narrative bytes: Data-driven content production in esports," in *Proceedings of the 2018 ACM international conference on interactive experiences for TV and online video*, 2018, pp. 29–41.

[12] A. Pedrassoli Chitayat, A. Coates, F. Block, A. Drachen, J. A. Walker, J. Dean, M. McConachie, and P. York, "From passive viewer to active fan: Towards the design and large-scale evaluation of interactive audience experiences in esports and beyond," in *Proceedings of the 2024 ACM International Conference on Interactive Media Experiences*, 2024, pp. 94–107.

[13] S. Charleer, K. Gerling, F. Gutiérrez, H. Cauwenbergh, B. Luycx, and K. Verbert, "Real-time dashboards to support esports spectating," in *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play*, ser. CHI PLAY '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 59–71. [Online]. Available: https://doi.org/10.1145/3242671.3242680

[14] Z. Yang, Z. Pan, Y. Wang, D. Cai, S. Shi, S.-L. Huang, W. Bi, and X. Liu, "Interpretable real-time win prediction for honor of kings—a popular mobile moba esport," *IEEE Transactions on Games*, vol. 14, no. 4, pp. 589–597, 2022.

[15] A. V. Kokkinakis, S. Demediuk, I. Nölle, O. Olarewaju, S. Patra, J. Robertson, P. York, A. P. Pedrassoli Chitayat, A. Coates, D. Slawson *et al.*, "Dax: Data-driven audience experiences in esports," in *ACM International Conference on Interactive Media Experiences*, 2020, pp. 94–105.

[16] V. J. Hodge, S. Devlin, N. Sephton, F. Block, P. I. Cowling, and A. Drachen, "Win prediction in multiplayer esports: Live professional match prediction," *IEEE Transactions on Games*, vol. 13, no. 4, pp. 368–379, 2019.

[17] S. Demediuk, P. York, A. Drachen, J. A. Walker, and F. Block, "Role identification for accurate analysis in dota 2," in *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, vol. 15, 2019, pp. 130–138.

[18] A. Pedrassoli Chitayat, F. Block, J. Walker, and A. Drachen, "Beyond the meta: Leveraging game design parameters for patch-agnostic esport analitics," in *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, 2023, pp. 116–125.

[19] S. Demediuk, A. Kokkinakis, M. S. Patra, J. Robertson, B. Kirman, A. Coates, A. Chitayat, J. Hook, I. Nolle, O. Olarewaju *et al.*, "Performance index: A new way to compare players," in *2021 MIT Sloan Sports Analytics Conference. IEEE*, 2021.

[20] R. Sifa, A. Drachen, F. Block, S. Moon, A. Dubhashi, H. Xiao, Z. Li, D. Klabjan, and S. Demediuk, "Archetypal analysis based anomaly detection for improved storytelling in multiplayer online battle arena games," in *Proceedings of the 2021 Australasian Computer Science Week Multiconference*, 2021, pp. 1–8.

[21] X. Qian, R. Sifa, X. Liu, S. Ganguly, B. Yadamsuren, D. Klabjan, A. Drachen, and S. Demediuk, "Anomaly detection in player performances in multiplayer online battle arena games," in *Proceedings of the 2022 Australasian Computer Science Week*, ser. ACSW '22. New York, NY, USA: Association for Computing Machinery, 2022, p. 23–30. [Online]. Available: https://doi.org/10.1145/3511616.3513095

[22] C. Bauckhage and R. Sifa, "k-maxoids clustering." in *LWA*, 2015, pp. 133–144.

[23] A. Cutler and L. Breiman, "Archetypal analysis," *Technometrics*, vol. 36, no. 4, pp. 338–347, 1994.

[24] R. Kirshenboim, "Top games on twitch 2016 - 2023," Kaggle, V17, Acceess: https://www.kaggle.com/datasets/rankirsh/evolution-of-top-games-on-twitch, 2024.

[25] A. Wiltshire, "Why revealing all is the secret of Slay The Spire's success," [Online]. Available: https://www.rockpapershotgun.com/why-revealing-all-is-the-secret-of-slay-the-spires-success (accessed: 2025-03-09), 2018.