PlayMyData: a curated dataset of multi-platform video games

Andrea D'Angelo
andrea.dangelo6@graduate.univaq.it
Università degli studi dell'Aquila
L'Aquila, Italy

Claudio Di Sipio
claudio.disipio@univaq.it
Università degli studi dell'Aquila
L'Aquila, Italy

Cristiano Politowski
cristiano.politowski@umontreal.ca
DIRO, University of Montreal
Montreal, Canada

Riccardo Rubei
riccardo.rubei@univaq.it
Università degli studi dell'Aquila
L'Aquila, Italy

ABSTRACT

Being predominant in digital entertainment for decades, video games have been recognized as valuable software artifacts by the software engineering (SE) community just recently. Such an acknowledgment has unveiled several research opportunities, spanning from empirical studies to the application of AI techniques for classification tasks. In this respect, several curated game datasets have been disclosed for research purposes even though the collected data are insufficient to support the application of advanced models or to enable interdisciplinary studies. Moreover, the majority of those are limited to PC games, thus excluding notorious gaming platforms, e.g., PlayStation, Xbox, and Nintendo. In this paper, we propose PlayMyData, a curated dataset composed of 99,864 multi-platform games gathered by the IGDB website. By exploiting a dedicated API, we collect relevant metadata for each game, e.g., description, genre, rating, gameplay video URLs, and screenshots. Furthermore, we enrich PlayMyData with the timing needed to complete each game by mining the HLTB website. To the best of our knowledge, this is the most comprehensive dataset in the domain that can be used to support different automated tasks in SE. More importantly, PlayMyData can be used to foster cross-domain investigations built on top of the provided multimedia data.

KEYWORDS

video games, data mining, software engineering

ACM Reference Format:

Andrea D'Angelo, Claudio Di Sipio, Cristiano Politowski, and Riccardo Rubei. 2024. PlayMyData: a curated dataset of multi-platform video games. In *21st International Conference on Mining Software Repositories (MSR '24), April 15–16, 2024, Lisbon, Portugal.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3643991.3644869

1 INTRODUCTION

Over the last decades, digital entertainment has become a billion-dollar business and it is still growing [6]. In this respect, the videogame industry plays a crucial role by offering tons of new products every year. This expanding market segment has attracted the academic interest in different domains, e.g., behavior analysis [7, 16, 28], gender bias detection[5, 18] and addiction [8, 34]. Just recently, the software engineering (SE) community has started to consider games as knowledgeable software artifacts [23], thus opening several research opportunities, e.g., code smell detection [21], sentiment analysis [2, 33], and serious games [4]. Similar to traditional software artifacts, the main challenge remains the data gathering since most of the well-known digital stores and websites [1, 26, 32] collect unstructured data or don't support mining activities with a proper API, thus requiring extra development efforts.

To fill this gap, we propose PlayMyData, a curated videogame dataset of 99,864 games belonging to main gaming platforms (called *platforms* hereon) i.e., PlayStation, Xbox, Nintendo, and PC, stored on IGDB website [1]. The rationale behind this choice is *i)* it offers a dedicated API [12] to collect the needed data and *ii)* the mined data are reusable for non-commercial usage [2], thus fostering the reproducibility of the results. In the scope of the paper, we mined all relevant metadata to support automated text classification, e.g., descriptions, game genres, and ratings. In addition, we collect 43,812 video gameplay URLs and 443,630 screenshots that can be used to support computer vision tasks. We further enhance PlayMyData with game completion times, i.e., the number of hours to complete a game, stored on HLTB website[3] by using a community-based API [20]. To this end, we query HLTB by applying the Levensthein distance [22] to the game title, with the aim of reducing possible false positives. To the best of our knowledge, PlayMyData is the first multi-platform dataset that includes completion times. Even though the primary usage of PlayMyData is towards the support of automated approaches, we can envision different potential usages in the social domain by means of the collected multimedia data, i.e., screenshots and gameplay videos. The mined dataset is made publicly available at https://zenodo.org/records/10262075.

2 PLAYMYDATA COLLECTION

Figure 1 depicts the data collection process of the two selected websites, i.e., IGDB and HLTB. Concerning the former, we collect relevant metadata to our purpose on IGDB including the screenshots and the URLs gameplay videos. Afterwards, we searched the titles of the retrieved games on the HLTB platform and we matched them using the Levenstein similarity function.

[1]https://www.igdb.com/
[2]https://www.twitch.tv/p/it-it/legal/developer-agreement/
[3]howlongtobeat.com

MSR '24, April 15–16, 2024, Lisbon, Portugal

ACM ISBN 979-8-4007-0587-8/24/04.
https://doi.org/10.1145/3643991.3644869

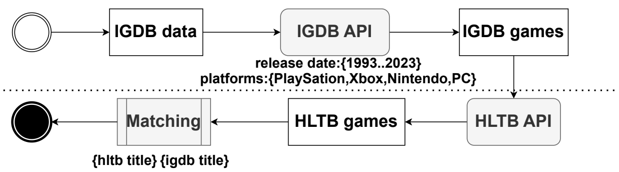

Figure 1: The PlayMyData collection process.

2.1 PlayMyData data model

Figure 2 depicts the PlayMyData data model. The main concept is the `Game` class that represents a game, with a unique integer id and a name, i.e., the title of the game. Since we mined games from two different storage, the `IGDBGame` and `HLTBGame` classes represent the collected data from IGDB and HLTB respectively. An `IGDBGame` contains a `summary` of the game and a more succinct description, i.e., the `storyline`. Furthermore, the average rating assigned by the users is represented by the attribute `rating`. A game can belong to 23 different `genres` available on IGDB[4]. Similarly, the same game has been released on different `platforms`. We define the `PlatformFamily` enumeration that describes the considered platforms i.e., PlayStation, Xbox, Nintendo, and PC. Roughly speaking, we consider all the platform versions for a specific family. For instance, PlayStation literal indicates all the PlayStation version, e.g., PSone, PS2. Each `IGDBGame` can possibly have a set of gameplay video and screenshots represented by the `Video` and `Screenshot` classes.

Similarly, the `HLTBGame` class represents specific information concerning the gameplay time expressed in hours for each game. In particular, HLTB stores the time needed to complete the `main` story, `extra` that includes the side-quests times, and the time to collect all the game achievements, i.e., `completionist` attribute. In addition, we collect the number of users that submit their completion time, i.e., `people_polled` attribute. In particular, this attribute measures the popularity and engagement levels of each game within the gaming community.

2.2 IGDB mining

To collect data from IGDB, we rely on the dedicated service [12] that facilitates the data-gathering phase through an authorized API. Mining all the games stored on IGDB is out of the scope of this paper. Thus, we focus on games released from 1^{st} January 1993 up to 30^{th} November 2023 belonging to the best-selling platforms[5], i.e., PlayStation, Xbox, Nintendo, and PC. To this end, we exploit the *first release* field available on IGDB. For each game, we fetch the metadata represented in the data model (cf. Section 2.1) using the query shown in Listing 1:

```
"fields [list fields]; where platforms= [platform ids] and
      first_release_date >= min date and  first_release_date < max
      date; sort rating desc; limit 500;"
```

Listing 1: The IGDB query to collect metadata.

[4]The reader can refer to the complete list of genres in the provided archive
[5]https://www.techradar.com/news/best-consoles

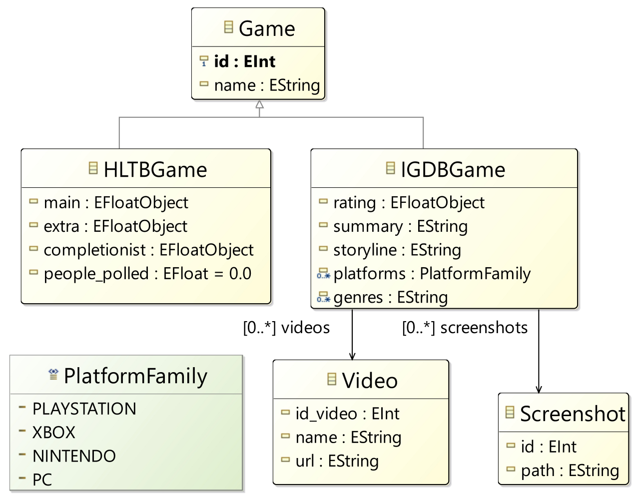

Figure 2: PlayMyData data model.

The *list fields* and *platforms ids* are the attributes of an `IGDBGame` as described in Section 2.1 and the list of unique IDs for each platform respectively. Meanwhile, *min date* and *max date* represent the time interval for each considered year. To handle the API rate limit, we set this interval equal to one month to fetch the maximum number of games for each month, by setting the *limit* equal to 500. The retrieved games are eventually ordered by the average user ratings, from the most rated to the least one.

In addition, we downloaded the screenshots and the gameplay video metadata for each game using two different sets of queries available here [25] that are similar to the one shown in Listing 1. We downloaded the screenshot files for all games since are in the *thumbnails* size, i.e., the entire size of the dataset is still manageable. Contrariwise, we opt to fetch only the YouTube URLs of the gameplay videos. Even though PlayMyData does not include the corresponding `.MP4` files, we provide a dedicated functionality that makes use of the PyTube library [6]. We made available an explanatory video in the supporting GitHub repository [25].

For each video, we store the name as appears on IGDB. This information can be further exploited to categorize the videos into trailers and actual gameplay released by the software house.

Overall, we collected 99,864 games, 443,630 screenshots, and 43,812 video URLs. Metadata related to games and videos are stored in CSV format while we group the screenshots using the genre of the corresponding game.

2.3 HLTB mining

The second phase involves the data collection from HLTB. In the scope of this work, we customize the community-built API [20] to retrieve the data related to the game completion as discussed in Section 2.1.

The conceived components execute a POST request on the HLTB website using the title of the game as shown in Listing 2.

[6]https://pytube.io/en/latest/

```
{"searchType": "games", "searchTerms": ["game_title"], "
  searchPage": 1, "size": 20}
```

Listing 2: The explanatory HLTB query.

In particular, the *game_title* parameter is the name attribute of an IGDB game as described in Figure 2 while the *search page* is set to 1 since we limit ourselves to the first result page. The *size* parameter represents the number of possible matches for that tile. In the scope of our analysis, we select the first one from this list. This component eventually retrieves a set of possible candidates that need to be mapped to IGDB games collected in the previous phase.

2.4 Matching IGDB and HLTB games

To match the collected data gathered from HLTB, we exploit the common element between the two platforms, i.e., the game title. However, this information may not perfectly match those stored on HLTB or may not be present at all, resulting in the retrieval of an entirely different game.

To overcome these issues, when merging the data from the two different sources, we compute the Levenshtein distance [22], a widely-used technique to measure the edit distance between two strings according to the standard formula:

$$L_{s_1,s_2}(i,j) = \begin{cases} \max(i,j) & \text{if } \min(i,j)=0, \\ L_{s_1,s_2}(i-1,j-1) & \text{if } s_1[i]=s_2[j], \\ \min \begin{cases} L_{s_1,s_2}(i-1,j)+1 \\ L_{s_1,s_2}(i,j-1)+1 \\ L_{s_1,s_2}(i-1,j-1)+1 \end{cases} & \text{otherwise.} \end{cases} \tag{1}$$

where s_1 and s_2 are the IGDB and HLTB game titles respectively, and i and j are their indices, starting from their length. In the scope of our work, we empirically set the similarity threshold equal to 3. On one hand, a higher threshold might lead to the inclusion of titles that are too different, potentially merging unrelated games. On the other hand, a lower threshold could be overly restrictive, excluding valid matches where titles vary slightly between the two databases. Roughly speaking, a Levenshtein distance of 3 allows for minor typographical differences or variations in game titles, thus maintaining an adequate level of accuracy during the matching phase. In total, we merged completion times data for 35,815 different games from HLTB.

3 PLAYMYDATA OVERVIEW

This section presents an overview of the collected data by providing basic descriptive statistics. In addition, we analyze the collected games over time by considering their completion time.

Descriptive statistic: PlayMyData contains 99,864 unique games over on a total number of 266,072 currently stored on IGDB[7]. By carefully inspecting the gathered data, we counted missing values for the *storyline*, *genres*, and *ratings* metadata. In particular, the storyline description is missing for only 445 entries while a relevant number of games have no rating, i.e., 32,859 games. A possible explanation is that only popular games are often rated by the users. This is confirmed by the HLTB since the *people_polled* value is equal to 0 for 62,450 games, meaning that any HLTB user didn't submit its time completion for these games.

We further analyzed the data checking the top five genres for each game considered platform as shown in Figure 3. Games tagged as *Shooter* are more common on Xbox and PlayStation as depicted in Figure 3a and Figure 3b. On the other hand, Nintendo games are mostly in the *Platform* genre, followed by *Puzzle* games. *Indie* games only appear on PC, with more than 8,200 titles, thus confirming the recent trend that the releases are "heavily skewed by small indie projects" [8]. Overall, the most popular genres are *Adventure*, *Shooter*, *RPG*, and *Simulator*, with more than 10K games each.

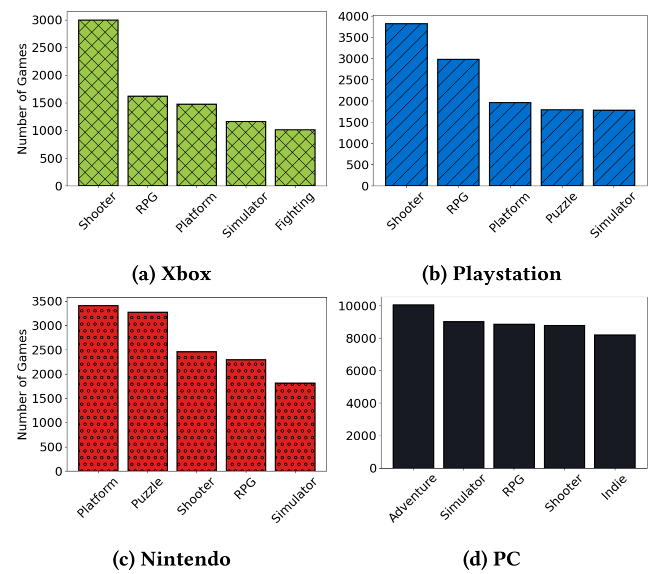

Figure 3: Most spread genres grouped by platform family.

Game completion over time: We explored the HLTB data and checked how the completion time has evolved over the years. Figure 4 depicts the number of games released per year and the corresponding completion time considering the three different attributes, i.e., main, extra, and completionist described in Section 2.1. As expected, the number of released games per year has grown during the considered time window, i.e., from less than 1,000 games released in 1993 to more than 8,000 in 2022 on average.

Even though the completion time follows a similar trend on average, the number of hours collected from HLTB doesn't grow at the same scale. The conducted analysis shows that completing a game nowadays requires less time compared to the middle 2000s, but still more than in the early 1990s. Even though recent findings show that the games are becoming longer "to beat" [9], this trend doesn't affect all the genres. In addition, the analysis conducted on our dataset reveals a peak in 2013, in particular for collecting all the

[7]https://www.igdb.com/about

[8]https://tinyurl.com/2bw7zw4t

[9]https://tinyurl.com/4xxk7dty

game achievements, i.e., the `completionist` attribute reached 80 hours on average. Investigating the possible reasons behind such a phenomenon is beyond the scope of this paper and is left as a possible future work.

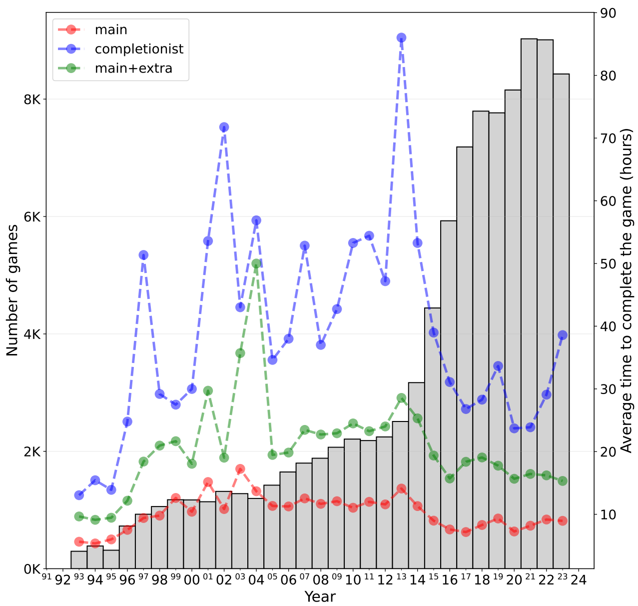

Figure 4: Game complention time collected from HLTB

4 RESEARCH OPPORTUNITIES

Text classification: The existing literature exploits textual data to automatically categorize games [3, 11, 30] or provide a personalized list of games [17, 24]. Compared to existing approaches, PlayMyData collects a large set of multi-platform games and the corresponding relevant data to facilitate the application of advanced techniques, e.g., pre-trained models. Furthermore, we augment the IGDB data with the completion time for each game collected from the HLTB storage, thus enabling the possibility of including such data to enhance automatic classification or developing recommender systems for games.

Image classification tasks: In the realm of image classification tasks, games remain an underexplored domain despite the widespread popularity of such models. Notably, in [13], images were leveraged to feed a multi-modal model designed for genre classification in video games. PlayMyData provides a vast repository of thumbnail screenshots, all standardized to the same size of 90x90 pixels. This consistency in image dimensions adheres to best practices established in the field [15, 31]. The screenshots provided by PlayMyData serve as a helpful resource to build models able to classify a game by genre or other attributes from still images.

Analyzing gameplay videos: Gameplay and video contribution have been employed to detect errors or anomalies [9] and for classification purposes [35]. Recently, gameplay videos have been employed to derive user behavior in gamification [10]. Even though PlayMyData collects only the YouTube URL videos, we offer an easy way to download and increase the collection of those artifacts.

5 THREATS TO VALIDITY

In this section, we briefly discuss the limitations of our paper. Concerning the *internal validity*, one possible threat is the selected time window,i.e., we consider only games released from 1993 to 2023. To mitigate any bias possible, we considered the most spread platforms,i.e., PlayStation, Xbox, Nintendo, and PC, thus building a representative enough dataset composed of 37% of the games stored on IGDB. In addition, we provide only the gameplay video URLs instead of the actual .MP4 file. To handle this, we provide a dedicated function to download the video from YouTube, thus allowing interested researchers to get all the stored videos. Threats to *external validity* are related to missing game information of PlayMyData, i.e., videos, ratings, and time completion. In particular, we may miss relevant data from HLTB since the stored games are different from IGDB. We opt for the Levensthein distance to match the game titles and minimize the number of missing entries.

6 RELATED WORKS

Jiang and Zheng [13] proposed a multi-modal classifier based on IGDB games. The authors used the cover image textual data to classify 50,000 PC games using their genres. Politowski et al. [27] manually analyzed postmortems of released games and created a dataset with 200 postmortems from 1998 to 2018, extracting 1,035 SE-related problems with traceability links. Melhart et al. [19] provided a dataset for affect modeling named AGAIN. The dataset consists of 37 hours of video footage from 1,116 gameplay sessions played by 124 participants. ViGGO [14] is a corpus for data-to-text Natural Language Generation that includes 100 games and about 7000 pairs of structured meaning representations. SIMS4ACTION [29] dataset collects more than 10 hours of gameplay of The Sims 4 game to simulate Activities of Daily Living (ADL).

Compared to the abovementioned approaches, PlayMyData collects a larger number of games, i.e., 99,864, considering additional gaming platforms apart from PC, i.e., PlayStation, Xbox, and Nintendo. Furthermore, we consider the completion time gathered from HLTB, screenshots, and gameplay video URLs.

7 CONCLUSION AND FUTURE WORKS

Intending to support the intersection between SE and entertainment, we present PlayMyData, a well-structured dataset composed of multi-platform games belonging to the main gaming platforms, i.e., Xbox, PlayStation, Nintendo, and PC. First, we collected 99,864 games that date back to 30 years ago available on the IGDB website including their relevant metadata, 443,630 screenshots, and 43,812 URLs of gameplay videos. Second, completion time has been collected from HLTB, a dedicated website that discloses such data, by applying the Levensthein distance on titles to reduce any possible bias. Although some data are missing, PlayMyData can be used to support cross-domain investigations, moving from social science to the application of automated techniques. In future works, we plan to fill in the missing entries and further expand the offered metadata, e.g., including game reviews. In addition, the older platforms can be included in the analysis, e.g., Sega or Amiga, to investigate the evolution of games over time. Last but not least, we can exploit multimedia data, i.e., screenshots and videos, to feed multimodal approaches to understand their impact on the classification task.

REFERENCES

[1] 2023. Video Game Reviews, Articles, Trailers and more. https://www.metacritic.com/game/ [last accessed on 2023-12-06].

[2] Rian Ardianto, Tri Rivanie, Yuris Alkhalifi, Fitra Septia Nugraha, and Windu Gata. 2020. Sentiment analysis on E-sports for education curriculum using naive Bayes and support vector machine. *Jurnal Ilmu Komputer dan Informasi* 13, 2 (2020), 109–122.

[3] Paul Bertens, Anna Guitart, Pei Pei Chen, and Africa Perianez. 2018. A machine-learning item recommendation system for video games. In *2018 IEEE Conference on Computational Intelligence and Games (CIG)*. IEEE, 1–4.

[4] Antonio Bucchiarone, Kendra M. L. Cooper, Dayi Lin, Edward F. Melcer, and Kelvin Sung. 2023. Games and Software Engineering: Engineering Fun, Inspiration, and Motivation. *SIGSOFT Softw. Eng. Notes* 48, 1 (jan 2023), 85–89. https://doi.org/10.1145/3573074.3573096

[5] Dalila Forni. 2020. Horizon Zero Dawn: The educational influence of video games in counteracting gender stereotypes. *Transactions of the Digital Games Research Association* 5, 1 (2020).

[6] Erin Gibson, Mark D Griffiths, Filipa Calado, and Andrew Harris. 2022. The relationship between videogame micro-transactions and problem gaming and gambling: A systematic review. *Computers in Human Behavior* 131 (2022), 107219.

[7] Isabela Granic, Adam Lobel, and Rutger CME Engels. 2014. The benefits of playing video games. *American psychologist* 69, 1 (2014), 66.

[8] Mark D Griffiths and Alex Meredith. 2009. Videogame addiction and its treatment. *Journal of Contemporary Psychotherapy* 39 (2009), 247–253.

[9] Emanuela Guglielmi, Simone Scalabrino, Gabriele Bavota, and Rocco Oliveto. 2023. Using gameplay videos for detecting issues in video games. *Empirical Software Engineering* 28, 6 (2023), 136.

[10] Reza Hadi Mogavi, Chao Deng, Jennifer Hoffman, Ehsan-Ul Haq, Sujit Gujar, Antonio Bucchiarone, and Pan Hui. 2023. Your Favorite Gameplay Speaks Volumes About You: Predicting User Behavior and Hexad Type. In *HCI in Games (Lecture Notes in Computer Science)*, Xiaowen Fang (Ed.). Springer Nature Switzerland, Cham, 210–228. https://doi.org/10.1007/978-3-031-35979-8_17

[11] Ismo Horppu, Antti Nikander, Elif Buyukcan, Jere Mäkiniemi, Amin Sorkhei, and Frederick Ayala-Gómez. 2021. Automatic Classification of Games using Support Vector Machine. *CoRR* abs/2105.05674 (2021). arXiv:2105.05674 https://arxiv.org/abs/2105.05674

[12] igdb API. 2023. Getting Started – IGDB API docs. https://api-docs.igdb.com/#getting-started

[13] Yuhang Jiang and Lukun Zheng. 2020. Deep learning for video game genre classification. *arXiv:2011.12143 [cs]* (Nov. 2020). http://arxiv.org/abs/2011.12143 00001 arXiv: 2011.12143.

[14] Juraj Juraska, Kevin K Bowden, and Marilyn Walker. 2019. ViGGO: A video game corpus for data-to-text generation in open-domain conversation. *arXiv preprint arXiv:1910.12129* (2019).

[15] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. 2017. ImageNet Classification with Deep Convolutional Neural Networks. *Commun. ACM* 60, 6 (may 2017), 84–90. https://doi.org/10.1145/3065386

[16] Simone Kühn, Dimitrij Tycho Kugler, Katharina Schmalen, Markus Weichenberger, Charlotte Witt, and Jürgen Gallinat. 2019. Does playing violent video games cause aggression? A longitudinal intervention study. *Molecular psychiatry* 24, 8 (2019), 1220–1234.

[17] Tianrui Liu. 2022. RecommenderPlus: New Content-based User-centered Game Recommendation System. In *2022 3rd International Conference on Computer Vision, Image and Deep Learning & International Conference on Computer Engineering and Applications (CVIDL & ICCEA)*. 767–770. https://doi.org/10.1109/CVIDLICCEA56201.2022.9825356

[18] Daniel Madden, Yuxuan Liu, Haowei Yu, Mustafa Feyyaz Sonbudak, Giovanni M Troiano, and Casper Harteveld. 2021. “Why are you playing games? You are a girl!”: Exploring gender biases in Esports. In *Proceedings of the 2021 CHI conference on human factors in computing systems*. 1–15.

[19] David Melhart, Antonios Liapis, and Georgios N Yannakakis. 2022. The arousal video game annotation (AGAIN) dataset. *IEEE Transactions on Affective Computing* 13, 4 (2022), 2171–2184.

[20] Michele. 2023. HowLongToBeat Python API. https://github.com/ScrappyCocco/HowLongToBeat-PythonAPI original-date: 2018-12-28T22:50:59Z.

[21] Vittoria Nardone, Biruk Muse, Mouna Abidi, Foutse Khomh, and Massimiliano Di Penta. 2023. Video Game Bad Smells: What They Are and How Developers Perceive Them. *ACM Trans. Softw. Eng. Methodol.* 32, 4, Article 88 (may 2023), 35 pages. https://doi.org/10.1145/3563214

[22] Gonzalo Navarro. 2001. A guided tour to approximate string matching. *Comput. Surveys* 33, 1 (2001), 31–88. https://doi.org/10.1145/375360.375365

[23] Luca Pascarella, Fabio Palomba, Massimiliano Di Penta, and Alberto Bacchelli. 2018. How Is Video Game Development Different from Software Development in Open Source?. In *2018 IEEE/ACM 15th International Conference on Mining Software Repositories (MSR)*. 392–402. ISSN: 2574-3864.

[24] Apurva Pathak, Kshitiz Gupta, and Julian McAuley. 2017. Generating and Personalizing Bundle Recommendations on Steam. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR ’17)*. Association for Computing Machinery, New York, NY, USA, 1073–1076. https://doi.org/10.1145/3077136.3080724

[25] PlayMyData. 2023. riccardoRubei/MSR2024-Data-Showcase: Repository for MSR2024 Data Showcase. https://github.com/riccardoRubei/MSR2024-Data-Showcase

[26] PlaystationStore. 2023. Latest | Official PlayStation™Store. https://store.playstation.com/ [last accessed on 2023-12-06].

[27] Cristiano Politowski, Fabio Petrillo, Gabriel Cavalheiro Ullmann, Josias de Andrade Werly, and Yann-Gaël Guéhéneuc. 2020. Dataset of video game development problems. In *Proceedings of the 17th International Conference on Mining Software Repositories*. 553–557.

[28] Muhannad Quwaider, Abdullah Alabed, and Rehab Duwairi. 2019. The impact of video games on the players behaviors: A survey. *Procedia Computer Science* 151 (2019), 575–582.

[29] Alina Roitberg, David Schneider, Aulia Djamal, Constantin Seibold, Simon Reiß, and Rainer Stiefelhagen. 2021. Let’s play for action: Recognizing activities of daily living by learning from life simulation video games. In *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 8563–8569.

[30] Riccardo Rubei and Claudio Di Sipio. 2021. AURYGA: A Recommender System for Game Tagging.. In *IIR*.

[31] Karen Simonyan and Andrew Zisserman. 2015. Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv:1409.1556 [cs.CV]

[32] Steam. 2023. Steam Store. https://store.steampowered.com/ [last accessed on 2023-12-06].

[33] Joseph J Thompson, Betty HM Leung, Mark R Blair, and Maite Taboada. 2017. Sentiment analysis of player chat messaging in the video game StarCraft 2: Extending a lexicon-based model. *Knowledge-Based Systems* 137 (2017), 149–162.

[34] Richard TA Wood. 2008. Problems with the concept of video game “addiction”: Some case study examples. *International journal of mental health and addiction* 6 (2008), 169–178.

[35] Saman Zadtootaghaj, Steven Schmidt, Nabajeet Barman, Sebastian Möller, and Maria G. Martini. 2018. A Classification of Video Games based on Game Characteristics linked to Video Coding Complexity. In *2018 16th Annual Workshop on Network and Systems Support for Games (NetGames)*. 1–6. https://doi.org/10.1109/NetGames.2018.8463434