

# Data-Driven Game Development: Ethical Considerations

Magy Seif El-Nasr  
Northeastern University  
magy@northeastern.edu

Erica Kleinman  
Northeastern University  
kleinman.e@husky.neu.edu

## ABSTRACT

In recent years, the games industry has made a major move towards data-driven development, using data analytics and player modeling to inform design decisions. Data-driven techniques are beneficial as they allow for the study of player behavior at scale, making them very applicable to modern digital game development. However, with this move towards data driven decision-making comes a number of ethical concerns. Previous work in player modeling [45] as well as work in the fields of AI and machine learning [9, 53] have demonstrated several ways in which algorithmic decision-making can be flawed due to data or algorithmic bias or lack of data from specific groups. Further, black box algorithms create a trust problem due to lack of interpretability and transparency of the results or models developed based on the data, requiring blind faith in the results. In this position paper, we discuss several factors affecting the use of game data in the development cycle. In addition to issues raised by previous work, we also raise issues with algorithms marginalizing certain player groups and flaws in the resulting models due to their inability to reason about situational factors affecting players' decisions. Further, we outline some work that seeks to address these problems and identify some open problems concerning ethics and game data science.

## CCS CONCEPTS

• Human-centered computing → Visualization.

## KEYWORDS

Game Design, Game Data, Player Modeling, Ethics, Human-in-the-Loop, Transparent Models

## ACM Reference Format:

Magy Seif El-Nasr and Erica Kleinman. 2018. Data-Driven Game Development: Ethical Considerations. In *The Fifteenth International Conference on the Foundations of Digital Games (FDG '20)*, September 15–18, 2020, Bugibba, Malta. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/1122445.1122456>

## 1 INTRODUCTION

Data Science is becoming a tool-set used in many applications, including recruitment, marketing, design, workflow analysis, and policy making — affecting people as consumers as well as producers of technology. Within game development, specifically, game data science and analytics have been used to tune and adapt design

[23, 32, 33], assess workflow [87], and recommend teams or items to players [15, 17]. With such an increase in the use of data within many of the processes affecting users, organizations, and developers, several researchers questioned the ethical implications of such practices. In her book, *Weapons of Math Destruction*, O'Neil discusses the issues stating: "consequently, racism is the most slovenly of predictive models. It is powered by haphazard data gathering and spurious correlations, reinforced by institutional inequities, and polluted by confirmation bias [53], [P. 23]."

O'Neil isn't the first to express concerns regarding the moral and ethical impact of Artificial Intelligence (AI) on our society. Bostrom and Yudkowsky [9] give a detailed overview of the types of ethical considerations that need to be taken into account as algorithms are increasingly charged with tasks that have legitimate social impact. They warn that lack of data, or incorrect selection of data, can inject bias into the output. Further, because the inner workings of many Machine Learning (ML) systems are often unknown to most, it can be difficult to know why an AI made a certain decision. This becomes increasingly problematic when those decisions have a negative social impact, such as denying mortgages to African American applicants [9], or misidentifying them as gorillas — a bias in the google image identification algorithm reported in 2015 [75].

These concerns also raise issues of users' trust or mistrust in the system. Studies show that the people who would, and do, interact with these algorithms in the wild rarely trust them [8, 26]. Binns et al. [8] conducted a study in which they found that many participants felt it was dehumanizing or impersonal to allow a machine to make decisions for humans. They also stated concerns over their inability to understand how the systems arrived at their answers, generalization, and statistical inference [8]. Eslami et al.'s [26] work examining the Yelp algorithm found that many users did not trust the opaque system and expressed that they felt it was discouraging or even impeding their freedom of speech, with many users demanding transparency. Users' concerns regarding transparency and explainability are also echoed by Siau and Wang [61] as critical factors affecting whether or not people can trust AI.

So what does that mean for games? As the game industry and academic research starts to streamline processes for the use of data in many aspects of game development, design, and research, they also will face similar issues and concerns regarding privacy, bias, unfairness, and marginalization of specific groups of players, which may raise issues of trust or distrust among the community. This may have more severe impact in applications of games in education, health, or others [12]. Mikkelsen et al. [45] discussed numerous ethical drawbacks to player modeling. For example, they raised the issue that algorithms designed to increase monetization and prevent churn do not consider the financial status of the player,

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](https://permissions.acm.org).  
FDG '20, September 15–18, 2020, Bugibba, Malta

© 2018 Association for Computing Machinery.  
ACM ISBN 978-1-4503-XXXX-X/18/06...\$15.00  
<https://doi.org/10.1145/1122445.1122456>

meaning they risk trapping the player in a predatory cycle, spending money they do not have. Automated banning systems can benefit the player base by punishing inappropriate behavior, but must be transparent in terms of what they target, so players can intentionally avoid them. Recommendation systems and adaptive systems may not properly accommodate all players equally, resulting in unpleasant gaming experiences for minorities or non-expert players. Recommendation systems may also potentially discriminate against players with less purchasing power. They also warn that using models during the balancing process can result in gameplay adjustments that do not apply to the entire player-base. In addition to all of these concerns, there is also the underlying issue of privacy [45].

Like O’Niel [53], Mikkelsen et al. [45] recommend the use of transparent models to address some of these issues. While transparency is important, there needs to be a concrete methodological approach that includes transparency as part of the process when addressing these issues. A fair body of work has been dedicated to creating algorithms and algorithmic methodologies that are considered more ethical or fair [5, 18, 68]. Zhu et al. [85] and Roselli et al. [58] attempted to resolve these ethical drawbacks by increasing human involvement in the modeling process in order to increase transparency, explainability, and auditability. Zhu et al. [85] developed a novel approach to algorithm design they call “Value-Sensitive Algorithm Design”. They sought to address two concerns 1) a lack of critical engagement with users and stakeholders, and 2) a reliance on historical judgements rather than insights into how the world can be improved. They incorporated stakeholder knowledge and feedback early in development in order to avoid bias or undermine stakeholder values in the algorithm’s design [85]. They use an iterative design process in which the algorithm is tuned based on feedback to remove unintended biases. Further, Roselli et al. [58] presented a series of steps that can be taken when training a model to reduce bias. They emphasized that substantial evidence should be present to back up a hypothesis and that training data should not be overly curated and guarded against manipulation, to avoid bias. They also pushed for a feedback loop in which the outcomes are checked against predictions and negative feedback loops are avoided by comparison with external evidence [58]. While this work presents a great stride towards a methodology that address such ethical issues, such processes have not been discussed within the context of the game data science pipeline.

These interpretability concerns have also given rise to the field of Explainable AI (XAI), which employs various methods and techniques to ensure that the outputs of an AI system are explainable and interpretable by humans [19, 35, 46, 81]. Work in XAI sought to normalize what “explainable” means [19] such that AI systems can be designed and implemented in a consistent manner. Further, work has sought to connect facets of “explainability” with facets of human reasoning to better guide XAI development [81]. However, there are concerns that much of the explainability guidelines are more focused on explaining the systems to the developers rather than to users [46]. Further, the work in this field is still in its beginning and more work needs to be done to make such approaches useful for the game data science pipeline.

While there has been a strong push for more ethical and explainable AI, little work has been done in the field of games. To

make matters more difficult, analyzing game data is an incredibly complicated process. Game data is often influenced by numerous factors, including, but not limited to, individual differences among players and the situational nature of gameplay. In terms of individual differences, there are various player differences that will naturally impact the way any given player approaches a game. These range from cultural differences [54, 70] to personality differences [11, 16] to skill based differences [16, 21]. At the same time, decision-making in games is heavily informed and influenced by the context in which it occurs [1]. This results in highly situational and individually varied data that cannot easily be aggregated or distributed for analysis without introducing some risk of bias or misinterpretation into the process.

In this position paper, we argue that issues impacting data-driven techniques and AI concerning fairness and ethics impact game data science as well. We expand on the concerns discussed by others echoing issues of individual differences and situational decision-making. We then build on previous works’ discussions of ethical concerns in game data by enumerating some new methods being developed and explored in games that can begin to address these concerns. We then provide takeaways that we believe can help inform more ethical data driven game design. We conclude by discussing the issues that need to be addressed by game data scientists, the open problems and their implications on our field. We hope this position paper can be a start to the discussion around game data science and ethics — as such discussion has already been underway in the fields of ML and AI.

## 2 ETHICAL CONCERNS WITH THE USE OF GAME DATA

Games, as is the case with other media, is a genre of entertainment that have long been concerned with ethics due to the potential for the medium to influence consumers [34, 34, 63, 73]. There have been many discussions regarding ethics in game development, ranging from game content and exposing players to violence [43, 66, 76] to the role that games can play as vehicles to inspire reflection and thought [41, 55, 62, 64]. We are particularly interested in the ethical use of data in the game development cycle, and thus we will focus the discussion here on this area.

### 2.1 Monetization and Its Impact on Players

In 2017, EA’s *Star Wars Battlefront 2* [24] announced a controversial lootbox system that had many claiming the game was “pay to win” [67]. The controversy directed attention to the monetization strategies of games and their tendency towards micro-transactions. Naturally, this gave rise to ethical concerns regarding the classification of such game elements as gambling [22]. Since then, a fair amount of work has examined the classification of lootboxes and the connection between them and gambling. A large scale study by Zendle and Cairns [84] found correlations between problematic gambling behaviors and the purchasing of lootboxes with real money. However, their study did not infer causation. In a more recent work Nielsen and Grabarczyk [52] developed a formal classification system for lootboxes (renamed “random reward mechanisms”) in which they break them down into categories, with only one type linked to gambling.

Much work still seeks to understand what motivates players to spend real money in otherwise free to play environments. Hamari et al. [31] performed a large scale study to address this issue. They identified several factors, including unobstructed play, social interaction, competition, and unlocking content. The use of monetization techniques have been criticized due to their use of player data to optimize on these factors. Mikkelsen et al. [45], for example, discussed how data can be used to determine the best way to encourage a player to make a purchase within a game. They then discussed that such data does not include any relevant information about the player, such as whether or not they are financially able to make such a purchase. The result is a predatory cycle, fueled by a desire for profit. Further, they illustrated how a player who makes financially responsible decisions, by not spending what they do not have, may be discriminated against by recommendation systems [45].

These examples make it clear that the ethical concerns surrounding game monetization are closely linked to the ethical concerns surrounding game data, as it is such data that is often used to design free-to-play monetization systems. In her GDC talk, Celia Hodent warned that there is a lack of understanding of how design decisions may be impacting users' purchasing habits. These "dark patterns", as she referred to them, include guilt tripping players into paying to make a character happy, or making it difficult to play for free [34].

Despite these concerns, a study by Alha et al. [2] found that most designers were only concerned about transactions being problematic in certain contexts, e.g., in the vicinity of children. They felt that free-to-play games with micro-transactions were no less ethical than a game you paid for upfront. However, issues of bias and recommendation systems' marginalization of particular players may be hidden in the black-box methods used.

## 2.2 Retention = Addiction

As the games industry focuses more and more on games that rely on subscriptions, repeated or long term play, and in-game purchases, research on the use of data began to focus on how to keep people in the game longer [6]. Various AI and player modeling techniques are used with the intent of achieving this goal. Xue et al. [82] developed a probabilistic model for dynamic difficulty adjustment to increase player engagement across levels. Hadiji et al. [30] developed a set of classifiers that can be used to predict churn in "freemium" (free to play, pay for bonuses) games. Additionally, Peterson et al. [56] used a mixed methods approach to evaluate the on-boarding phase of free-to-play mobile games as they relate to the churn rate of new players. These are only a few of many works that sought to maximize player retention. However, there is a fine line between keeping players engaged in a game and stimulating addiction.

Hodent addressed addiction in her ethics talk, illustrating that, while concerns were valid, and there are those who need help, there was little evidence that video game 'clinical' addiction actually occurred [34]. However, Hodent drew on the clinical definition of addiction in her discussion, and the exact term used to diagnose patients in a clinical setting may not apply to most cases discussed

for video games. There were nevertheless reported harmful behavioral signs among the game playing population, such as neglecting school, work, or family responsibilities [48, 80]. For example, a Korean couple allowed their child to starve to death while they played an online game [59]. Additionally, studies conducted by Grüsser et al. [29] showed that addictive gaming habits do not exist solely in correlation with monetization in games, and warned designers to be aware of how they entice players to continue playing, to avoid promoting damaging behavior. This is a particularly difficult task, as keeping players engaged is the overall goal of game design, and data is a helpful aid in achieving such a goal at scale. More work and discussion is needed on this dimension. On the flip side, there is also much data on players' game use and behavior which can be used to moderate play sessions.

## 2.3 Representation in Data

As Mikkelsen et al. [45] pointed out, it is difficult to create models that correctly account for everyone, and this becomes even more true when dealing with games that have large international communities and a broad user base. Previous work [54, 70] demonstrated that different location-based contexts will result in different gameplay motivations and behavior. This then raises a valid concern that aggregated data will not capture these differences. Further, data-sets that are drawn from a single region or skill level will result in a lack of representation in the data, and a model that is biased towards a particular group. This is an issue we will expand upon later in the paper.

## 2.4 Player Modeling

Despite the fact that player modeling has grown in popularity [28, 36, 65], the work on the ethical issues surrounding game AI and player modeling is sparse at best. Yannankis et al. [83] gave a detailed overview of the benefits and applications of player modeling. In discussing further work, they illustrated concerns that echo those regarding AI ethics, such as privacy, stereotypes, and censorship. Further, concerns regarding player's individual differences, and how models can adapt to these differences, are prominent. Hooshyar et al. [37] cite individual prediction as one of the challenges for player modeling in educational games, among others. Charles and Black [14] developed a framework for dynamic player modeling that attempted to address this by allowing the model to reclassify players as their performance drifted.

However, individual differences are only one of several ethical issues within the field of player modeling. Mikkelsen et al. [45] build on the work of ethics in AI to clearly articulate the ethical and moral shortcomings of player modeling as it relates to a variety of applications within games, including churn, recommendation systems, balancing, and matchmaking. In this paper, we echo some of these concerns and expand on them towards the development of methods that can account for these issues and biases.

## 3 OTHER ETHICAL ISSUES WITH DATA-DRIVEN GAME DEVELOPMENT

There are other concerns regarding data-driven game development that have not been discussed extensively by previous work, but have been apparent in studies regarding player behavior in games.



In this section we will discuss two issues we believe are important to the discussion here, particularly: Individual Differences and Situational Factors.

### 3.1 Individual Differences = Error or Outliers?

O’Neil [53] describes Simpson’s Paradox — a paradox that holds true with much of the work on game data. In her book, she states: “Simpson’s Paradox: when a whole body of data displays one trend, yet when broken into subgroups, the opposite trend comes into view for each of those subgroups [53][p. 136].”

Games are dynamic and complex environments that rarely force players to solve problems in a single way. And previous work has found that partitioning players differently can bring about different results [13]. There are usually numerous influences that impact the way one approaches gameplay ranging from culture [7] to gender [42, 60] to expertise in game types or gameplay [16, 21, 25]. The result is a myriad of vastly different play behaviors that players exhibit when playing a game — such variation results in data skews. For example, expert players will play more levels and play longer, minimizing the data collected from non-expert players or players who give up early [16]. Further, there is often a dominant strategy that will be employed by many players — especially players who are used to specific types of games or puzzles. Due to ease, efficiency, or popularity, these dominant strategies are likely to overwhelm player data, and overshadow any uncommon (or creative) strategies that may exist at the fringes of the set. Such skews and lack of data from particular groups is challenging for automated modeling techniques, such as predictive classification or clustering algorithms.

To illustrate the existence of such individual differences, Nguyen et al. [51] devised an inverse reinforcement learning technique to predict policies, i.e. learned problem solving patterns, from game data logs. They used a game called *Wuzzit Trouble* developed and published by BrainQuake Games [10]. In their visualization of players’ patterns, they showed how players varied in what rewards they were trying to maximize while playing the game. Further, they also showed that players varied in their goals and have, in fact, shifted strategies as they were playing due to lack of or changing opportunities and their own learning process. They concluded that inverse reinforcement learning techniques were not yet mature enough to address such variations in the data.

Further, in our own work, we explored variation in problem solving behaviors through data collected from an educational game [40]. In our analysis of problem solving behaviors, we visualized all patterns to see how they varied over time. The visualization shows similar results to Nguyen et al.’s work [51]. Similar to the previous example, the resulting visualizations indicated that there are several most-popular trajectories to solve the levels. However, there were also many diverse trajectories that existed in the data and were visualized, many of which belong to players who struggled with the debugging process. This is where the significance of individual differences really comes into play. Because this work deals with player performance in an educational game, and because the trajectories of those who were successful are the most common, aggregation or using vanilla ML techniques could marginalize players who need the most design adjustments or help.

Such skews in the data become problematic when dealing with massive multiplayer online games, where the common strategy is often colloquially referred to as the “meta”. The “meta” is followed most closely by the most dedicated players, who will also make up the majority of the data, simply due to their lengthy play time. Casual players, who are more likely to engage “non-meta” strategies are thus more likely to be drowned out in the data, and overlooked by analysis. This can result in design changes that confuse or frustrate a large portion of a game’s player-base, or, in some cases, provide unintentional advantages to those engaging deviant strategies that went unnoticed due to aggregation.

These examples clearly illustrate how varied player behavior can be, even in simple games. When gameplay data is analyzed by traditional machine learning or statistical techniques, uncommon strategies or behaviors are frequently overlooked. However, if grouped by personality or other grouping factors, some of these patterns can be picked up by a modeling algorithm, e.g., [11]. Either to help those who need it, maintain balance, or improve quality of life, it is critical that game data analysis is able to capture individual differences. However, the aggregation, feature extraction, or generation methods used by most state-of-the-art techniques result in marginalizing individual with different play patterns.

### 3.2 What about Situational Factors?

In addition to overshadowing individual differences, most analytics or data science techniques rarely preserve the context in which actions or decisions occurred. However, player behavior is made up of situated actions [69], and it is difficult if not impossible to truly comprehend the strategic behavior of players without knowing the context that drove them to action.

Previous work by Ahmad et al. [1] demonstrated the importance of spatio-temporal context to understanding player strategies in *Defense of the Ancients 2* (DotA2) [72]. They used a human-in-the-loop methodology to apply behavioral labels to log data displayed in a spatio-temporal visualization system. This method allowed them to see when and where player actions occurred, which facilitated the capture of the state of the game environment when a behavior was taken [1]. As a result, they were able to generate a more context-aware model of player behavior that allowed them to identify how strategies of winning and losing teams varied as a game progressed [1].

DotA2 players engage in dozens of behaviors over the course of a match. On a raw data level, what the analysts see are attacks targeting other entities (either other players or non-player characters), items being used, entities being killed, or players dying. But it is the when and where behind these actions that define what they are on a strategic level. If hero A kills hero B in lane, during the early game, surrounded by creeps (a type of non-player character), then it is simply a *one-on-one kill*. If hero A kills hero B by entering his lane from the side jungle, and hero A has an ally in lane who drew hero B’s attention to set them up for the attack, then it is a strategic behavior known as a *Gank*. Hero A can also kill hero B in the midst of a team fight, because hero B is the enemy team’s primary damage dealer and hero A wants to remove them from play. In all three cases, hero A killed hero B, and without the contextual details, it is unlikely that these kills will be classified correctly.

This may have implications on how the system recommends teams, items or guides the player through a game.

## 4 ADDRESSING ETHICAL CONCERNS

Data is a powerful tool for game development that allows designers to understand players and increase engagement and profit at a massive scale. As discussed above, there are several ethical issues discussed concerning the use of data, these can be summarized as follows:

- (1) *Retention strategies that cause addiction*: many of the data driven techniques meant to keep players invested can lead to uncontrolled, addictive gameplay habits.
- (2) *Monetization techniques that encourage irresponsible spending*: placing game content and bonuses behind paywalls often motivates players to spend money they may not have to continue progressing.
- (3) *Lack of data on specific groups*: there may be a lack of data on specific groups, such as non-expert players who may not produce enough data throughout gameplay to allow us to address this group or model their behaviors.
- (4) *Lack of models that capture individual differences*: even with some data, if we model based on all players' behaviors or with the big group rather than splitting the data into smaller subsets, we may marginalize specific groups of people.
- (5) *Lack of data capturing details of context*: most game data are often aggregated statistics that may not capture decisions made moment to moment or the context behind such decisions.
- (6) *Lack of transparency or interpretability of the models or results*: we often produce models using black-box methods that do not allow us to interpret the reasons behind the results produced.

These issues are relevant to data in general, and in the areas of machine learning and data science at large, there has been some work addressing problems 3, 4, and 6 above. Work addressing 3 has looked into selecting data subsets that are better representative of the sample to minimize bias [44]. Work addressing 4 has examined using methods that are adjusted and tuned to better capture differences between individuals [39, 49, 57, 74]. The field of Explainable AI seeks to address issue 6 through detailed study of how AI systems can be explained to humans [19, 46, 81]. However, some of these issues, especially 4 and 5, are of greater pressing concern to games due to their open, dynamic environments. Issue 6 is also of particular concern if one wishes to increase interpretability to ensure player data is understood correctly.

The first three issues, addiction, monetization, and marginalization, are, for the most part, not issues that can be addressed through novel methods for data analysis. The first two deal more with the ethical application of data, rather than its analysis, while the third relies on the ethical collection of data. In this section, we will focus specifically on the latter three issues: individual differences, context, and interpretability, which we believe can be addressed through methodological approaches.

### 4.1 Addressing Individual Differences

To address the issue of individual differences discussed above, Canossa et al. [13], in their work on *The Division* [13, 71], used a multi-step approach involving expert designers interpreting and grouping players. They first clustered player data and then presented the resulting clusters to the game designers allowing them to assign meaning to the clusters. If the clustering results were not meaningful, different clustering techniques were used. This enforced the idea of data used to guide and inform, but not make decisions. It also allowed designers to be the judge of figuring out the different groupings in their data rather than treating the data of all the players as one whole. Their second step was to use sequence analysis to analyze each clusters' sequence of actions [13]. While this is a good step towards addressing differences, the focus on clustering means there is still a risk of losing individuals who sit too far from the clusters' centers.

Another approach used and proposed by several researchers is the use of visualization to seek meaning in players' behaviors. In addition to its benefits to preserving context and providing interpretable models, it also helps with the preservation and analysis of individual differences. Several visualization systems have been proposed [27, 47, 78]. Most of these visualization systems used a game map to show how players moved and exhibited actions through their play. Visualization proved beneficial in preserving individual differences in game data [79].

Visualizing data makes it possible to easily see not only players who are similar to each other (who form the clusters), but those who are dramatically different from each other, and those who don't quite fit into any particular cluster. However, as more player data is displayed such individual traces become cluttered making it hard to decipher meaning. To remedy this, several researchers explored the visualization of clusters based on the sequences displayed. A good example of this is *Glyph* [50], where researchers showed players' actions using two synchronized windows: one showing the sequences of players' action and the other showing such sequences clustered by similarity. Using such a visualization was effective in allowing designers to see individual differences in how players solved problems or behaved through time, because the clusters clearly showed popular strategies or patterns and uncommon ones.

Additionally, Ahmad et al. [1] presented a mixed methods approach in which action sequences, a human-in-the-loop, and two visualization systems are used to preserve spatio-temporal context, and facilitate the analysis of individual differences. They use one visualization system, *Stratmapper*, to present the actions of players in the context of when and where in-game they occurred, preserving a greater amount of situational factors than traditional techniques. This allows humans to create behavioral labels based on domain knowledge, which allows the abstraction applied to the data to reflect how players think about gameplay [1]. The labeled data is then exported to a second visualization, *Glyph*, where individual action sequences and clusters of sequences can be seen simultaneously. This allows for the analysis of individual differences, as individual variations can be easily identified within the clusters, and their action sequences (and how they differed) can be viewed [1].

## 4.2 Addressing Context

While Canossa et al. [13] didn't address problem 5 (situational context) completely, their use of sequences preserves some aspect of the action and some context of what happened before the action was taken. They showed that this technique was superior to using aggregate data [20]. To address the further issue of context (problem 5), Aung et al. incorporated both temporal and spatial context while analyzing player telemetry data from *Just Cause 2* [3, 4]. They then used clustering approaches and allowed experts to establish meaning for the clusters by examining the trajectories of the players closest to the centroid of each cluster [3]. While this technique is effective at incorporating situational context into behavioral profiling, examining just the centroid of the clusters marginalizes individual differences.

Additionally, much of the works that leverage visualization towards observing individual differences also successfully preserved a greater deal of contextual information through the ability to observe data in the spatial context of the game map [1, 27, 47, 78, 79]. As an illustrative example, Ahmad et al.'s [1] work, discussed above, provides a spatio-temporal visualization that allows human analysts, and domain experts, to examine the progression of in-game events in the context of where they occurred on the map. The experts can then create data labels that reflect this spatio-temporal context, ensuring that the data visualized in *Glyph* will include key contextual components. This increases the chances that the sequences will be interpreted correctly by researchers [1].

## 4.3 Addressing Transparency and Interpretability

Beyond the scope of games, eXplainable AI (XAI) is a field dedicated to highlighting the dangers of opaque systems, and developing algorithms that are interpretable and explainable [19, 35]. They emphasize that there is value in ensuring interpretability on the user end as well as among the engineers [46]. There is not much work applying XAI to game data, thus it is still an open area of exploration to see how XAI practices could be applied to game data analytics to improve transparency and interpretability. However, Zhu et al. [86] present a detailed vision of how XAI could enhance AI related systems in the domain of games. Specifically, they give examples of how XAI can be applied to AI assisted co-creation systems, procedural content generation systems, and systems that can evaluate in-game AI agents [86].

Game data is complex and difficult to interpret, and while it is not practical to collect think aloud data in the wild, being able to expose models to players, and allowing them to comment on or correct those models, can increase their transparency (addressing problem 6 above). Such an approach can also increase their accuracy, as discussed by Mikkelsen et al. [45].

Work has shown the benefits to including the players' thoughts in the modeling process. Horn et al. [38] used a mixed-method approach to analyzing player's behavior and strategies in an educational game. They used play traces as well as think-aloud utterances in their analysis. Specifically, they analyzed player's progression by hand as well as used hierarchical clustering to cluster play traces. This resulted in both questions and conclusions regarding player's strategic thinking that they were then able to answer or

confirm by analyzing player session think-aloud data, and connecting what players said with the moves they made [38]. While this type of analysis is time consuming and qualitative, it does allow the analyst to be involved in the modeling process and addresses issues of interpretability.

The work of Nguyen et al. [50] and Ahmad et al. [1] discussed above also address transparency. Through their visualization systems, they facilitate human-in-the-loop modeling by facilitating human interpretation through visualization and relying on human reasoning to draw conclusions from the visualized data [1, 50]. A number of other works also draw on visualization techniques to increase the human interpretability of data [77–79].

All of the approaches discussed above have merit, and all successfully address at least some of the problems discussed above. However, none are able to address all concerns, and some concerns have yet to be sufficiently addressed. Further, while the techniques above discuss some early directions, we are still far from finding practical approaches that scale and address these ethical concerns. Therefore, further exploration is still needed in this area.

## 5 TAKEAWAYS

As we have illustrated in the previous section, existing work has explored new and innovative methodological approaches to data driven game design. Many of these approaches help address a number of the ethical concerns that we have outlined. In this section, we will summarize the contributions of previous work and provide takeaways that developers and researchers can use to guide their procedures. Further, we will summarize open problems as they relate to each ethical concern. A synopsis of these topics can be seen in Table 1.

Previous work has shown that there are concerns regarding retention and monetization. Many games leverage player data to increase playtime or spending, and there is concern that this practice can cause addiction or financially irresponsible decisions [34, 45]. Previous work has suggested that, in both cases, designers think critically about the potential harm that could be caused by using data towards such a goal. Further, with regards to monetization, it is encouraged that developers be upfront about paid content, and models that track spending, such that players can avoid games that may entice them into bad practices. However, with regards to both addiction and monetization concerns, these suggestions are somewhat at odds with the goal of turning a profit. Thus, more work is needed to determine the best practices to encourage safe gameplay without harming a game's success.

Previous work have shown that there is a lack of data pertaining to certain groups, which, as discussed above causes marginalization [53]. This is not necessarily a concern that can be addressed through new techniques. As such, little work has addressed this through methodological approaches. Instead, we believe that this could be addressed through regulations that govern the collection and organization of data. Although there is little exploration of the viability of solutions, previous works cite the concern, and could be used as a reference to guide policy [9, 45]. However, this is still an open problem especially since it is difficult to collect more data from non-expert players who don't usually have enough gameplay

Ethical concern	How it's been addressed	Open problems
Retention and Addiction	Previous work presents conflicting evidence regarding the extent to which video games can become addictive. However, as a precaution, it is suggested that designers evaluate how they use data to inform design and encourage players to continue playing with a critical eye.	More work is needed to determine whether or not video game retention techniques cause addictive habits, vs. external factors beyond the design of the game. Additional work is also needed to determine the best way to encourage gameplay while preventing damaging behavior.
Monetization	Previous work recommends being transparent with players regarding the presence of paid content or monetization models in games, such that players can choose to avoid such content if exposure would be financially unsafe.	Such an approach, while honest, is at odds with the goal of making a profit, and further work is needed to determine what regulation and presentation practices can best balance these two factors.
Lack of Data	Previous work recommends regulation or policy changes to better guide how data is collected and organized making sure that enough data is collected from under represented minorities when possible.	There is no clear way to ensure that enough data is collected from all groups, especially in the case of game-play data, where novice players may play significantly less than experts, and may not produce enough data at all.
Individual Differences	Previous work showed that individual differences in games is not just a demographic difference, but can be about different style of play, or connected to game expertise and personality. To address and capture individual differences, three methods are recommended: (1) Developing models on the different player behavioral clusters capturing variations in players' behaviors, (2) using sequential rather than aggregate data can capture granular individual differences, and (3) visualizations can be used to represent the data where individual differences are visible or highlighted.	While these methods have been more effective at capturing individual differences than traditional techniques, more studies are needed to examine the exact extent to which each technique is able to address concerns of marginalization.
Context	Previous work have shown importance of context in the analysis process. It is then recommended that analysts identify what contextual elements are important to the data and leverage data formats and presentation tools to encode that context into the data representation.	It is difficult to define what constitutes context, and how it can best be derived from game data. Including context also increases the complexity of already incredibly complex data.
Transparency and Interpretability	Previous work have shown the importance of transparency and interpretability and discussed three ways to accomplish this: (1) using mixed methods leveraging qualitative techniques to embed labels or knowledge into the modeling process, (2) using visualization to allow for better representation and transparency, and (3) using explainable AI to explain resulting models from black-box machine learning techniques.	More work is needed to determine the best practices for visualization and scalable mixed methods within the domain of game data. Additionally, there is space for work exploring the applicability of explainable AI techniques to the game data domain.

**Table 1: The six identified ethical issues with game data analytics, how they are addressed in the literature, and open problems.**

data due to being novice or non-expert. Further, collecting more data to capture behavioral variations in style or personality is hard.

Related to lack of data, previous methods face challenges accounting for individual differences as discussed above. This is one of the most pressing concerns for games, as game data is vast and

varied. Existing work has demonstrated the benefit of doing analysis on clusters of players separately rather than the whole population. Further, previous work have also shown that using sequential data [13, 50] over aggregates [20] can prevent the marginalization of behavioral variations that may not be common. Additionally, existing work leveraged visualization [1, 27, 47, 50, 78] as a way of



making all of the players' trajectories, no matter how clustered or disparate, visible to the analyst. While these methods have been shown to capture individual differences, we don't have scientific evidence or a study that discusses marginalization and behaviors variations and methods the best capture them. Thus, more work is still needed perhaps expanding on sequential data analysis, cluster analysis and visualizations with more studies targeting the issue and question of marginalization and individual differences.

Context-aware analytics, as discussed above, is another important issue especially for game data science. This is due to the importance of environmental factors and their impact on players' behavior. Again, sequential data has demonstrated its ability to capture temporal context [13] and can be combined with spatial data to also capture location based context [3]. Here too, visualizations are being leveraged to allow analysts to easily capture the spatio-temporal context of the data [1]. The primary takeaway, similar to the previous issue, is to consider what elements of the game environment are of primary concern to the data being examined (such as time and location) and leverage data formats (such as sequences) and visualizations (such as heat maps) as ways to encode such context into the data representation. However, there are still open problems to this issue, including what constitutes a virtual context? How can it be derived from game data? Additionally, including context means increasing an already high-dimensional data, and thus this adds the issue of dealing with high dimensional analysis space.

The lack of interpretability or transparency of modeling techniques such as black-box machine learning techniques is an important issue echoed above. Existing work in games demonstrated the benefit of mixed methods approaches in linking players' reasoning to their actions, ensuring that their data is interpreted correctly by researchers [38]. Similarly, visualizations have proven useful in increasing the transparency of data, such that it can be more easily interpreted and more accurately understood [1, 50, 77]. Beyond the domain of game data, explainable AI has dedicated much time and resources towards achieving the goals of transparency and interpretability [19, 46], and existing work has shown potential in applying XAI concepts to games [86]. However, it has yet to be seen how XAI can guide more ethical game data science.

Going forward, we encourage work in game data analytics to explore such mixed methods and explainable AI approaches to ensure more interpretable and transparent data. Additionally, we encourage the exploration of exposing the model to players, which could potentially increasing player trust in the collection of their data [45]. The open problems in this area is centered on the need for more work around the development of methods for visualization, mixed-methods approaches that scale as well as explainable AI techniques.

## 6 CONCLUSION

Game data science and Artificial Intelligence algorithms are powerful tools that can be used effectively within software development. However, they also have the capability of inflicting damage, when implemented improperly, trained on incomplete or biased data, or when decisions they make are obscured or hidden [9, 53]. Game research and production processes, where such data-driven work has

bloomed in the last few years [23], are not immune to these dangers. Game data is incredibly complicated. In addition to the information it contains, it is also informed by various external contexts that impact the way players engage with gameplay. In this paper we identified several data analysis problems, including lack of data, lack of models that capture individual differences and context, and lack of transparency as underlying issues causing several ethical concerns and potential dangers to players such as predatory monetization, marginalization, and misrepresentation [45].

We further outlined several promising directions to address these issues. It should be noted that this area has received much attention in AI and ML lately. However, much of the work within AI and ML either focused on algorithm design rather than other aspects of human engagement, didn't address the issues we discussed here with game data, or were not discussed or applied to the game data science. In game data science research, there are several approaches that showed great promise. These involved increasing human involvement in all data science and analysis stages, increasing the interpretability of the results, and the use of visualization to preserve context, capture individual differences, and increase transparency of models. However, such approaches are still exploratory and face challenges in terms of scale and practicality. More work is needed to engage the community, as well as ethicists and social scientists, in discussing various new avenues for methodologically tackling the ethical problems and their root causes discussed in this paper. Further, there are probably other causes and problems that we can discuss or uncover as we start to tap this important area of research.

## REFERENCES

- [1] Sabbir Ahmad, Andy Bryant, Erica Kleinman, Zhaoqing Teng, Truong-Huy D Nguyen, and Magy Seif El-Nasr. 2019. Modeling Individual and Team Behavior through Spatio-temporal Analysis. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*. 601–612.
- [2] Kati Alha, Elina Koskinen, Janne Paavilainen, Juho Hamari, and Jani Kinnunen. 2014. Free-to-play games: Professionals' perspectives. *Proceedings of nordic DiGRA 2014* (2014).
- [3] Myat Aung, Simon Demediuk, Yuan Sun, Ye Tu, Yu Ang, Siva Nekkanti, Shantanu Raghav, Diego Klabjan, Rafet Sifa, and Anders Drachen. 2019. The trails of Just Cause 2: spatio-temporal player profiling in open-world games. In *Proceedings of the 14th International Conference on the Foundations of Digital Games*. 1–11.
- [4] Avalanche Studios. 2010. Just Cause 2. [Digital Game].
- [5] Bettina Berendt. 2019. AI for the Common Good?! Pitfalls, challenges, and ethics pen-testing. *Paladyn, Journal of Behavioral Robotics* 10, 1 (2019), 44–65.
- [6] Kelly Bergstrom. 2019. Moving beyond churn: Barriers and constraints to playing a social network game. *Games and Culture* 14, 2 (2019), 170–189.
- [7] Mateusz Bialas, Shoshannah Tekofsky, and Pieter Spronck. 2014. Cultural influences on play style. In *2014 IEEE Conference on Computational Intelligence and Games*. IEEE, 1–7.
- [8] Reuben Binns, Max Van Kleek, Michael Veale, Ulrik Lyngs, Jun Zhao, and Nigel Shadbolt. 2018. 'It's Reducing a Human Being to a Percentage': Perceptions of Justice in Algorithmic Decisions. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 377.
- [9] Nick Bostrom and Eliezer Yudkowsky. 2014. The ethics of artificial intelligence. *The Cambridge handbook of artificial intelligence* 316 (2014), 334.
- [10] Brainquake Games. 2011. Wuzzit Trouble. [Digital Game].
- [11] Sara Bunian, Alessandro Canossa, Randy Colvin, and Magy Seif El-Nasr. 2017. Modeling individual differences in game behavior using HMM. In *Thirteenth Artificial Intelligence and Interactive Digital Entertainment Conference*.
- [12] Alessandro Canossa. 2014. Reporting from the snooping trenches: Changes in attitudes and perceptions towards behavior tracking in digital games. *Surveillance & Society* 12, 3 (2014), 433–436.
- [13] Alessandro Canossa, Sasha Makarovych, Julian Togelius, and Anders Drachenn. 2018. Like a DNA string: Sequence-based player profiling in Tom Clancy's



- the Division. In *Fourteenth Artificial Intelligence and Interactive Digital Entertainment Conference*.
- [14] Darryl Charles and Michaela Black. 2004. Dynamic player modeling: A framework for player-centered digital games. In *Proc. of the International Conference on Computer Games: Artificial Intelligence, Design and Education*. 29–35.
  - [15] Zhengxing Chen, Christopher Amato, Truong-Huy D Nguyen, Seth Cooper, Yizhou Sun, and Magy Seif El-Nasr. 2018. Q-deckrec: A fast deck recommendation system for collectible card games. In *2018 IEEE Conference on Computational Intelligence and Games (CIG)*. IEEE, 1–8.
  - [16] Zhengxing Chen, Magy Seif El Nasr, Alessandro Canossa, Jeremy Badler, Stefanie Tignor, and Randy Colvin. 2015. Modeling individual differences through frequent pattern mining on role-playing game actions. In *Eleventh Artificial Intelligence and Interactive Digital Entertainment Conference*.
  - [17] Zhengxing Chen, Truong-Huy D Nguyen, Yuyu Xu, Christopher Amato, Seth Cooper, Yizhou Sun, and Magy Seif El-Nasr. 2018. The art of drafting: a team-oriented hero recommendation system for multiplayer online battle arena games. In *Proceedings of the 12th ACM Conference on Recommender Systems*. ACM, 200–208.
  - [18] Sam Corbett-Davies and Sharad Goel. 2018. The measure and mismeasure of fairness: A critical review of fair machine learning. *arXiv preprint arXiv:1808.00023* (2018).
  - [19] Derek Doran, Sarah Schulz, and Tarek R Besold. 2017. What does explainable AI really mean? A new conceptualization of perspectives. *arXiv preprint arXiv:1710.00794* (2017).
  - [20] Anders Drachen, Rafet Sifa, Christian Bauckhage, and Christian Thureau. 2012. Guns, swords and data: Clustering of player behavior in computer games in the wild. In *2012 IEEE conference on Computational Intelligence and Games (CIG)*. IEEE, 163–170.
  - [21] Anders Drachen, Matthew Yancey, John Maguire, Derrek Chu, Iris Yuhui Wang, Tobias Mahlmann, Matthias Schubert, and Diego Klabajan. 2014. Skill-based differences in spatio-temporal team behaviour in defence of the ancients 2 (dota 2). In *2014 IEEE Games Media Entertainment*. IEEE, 1–8.
  - [22] Aaron Drummond and James D Sauer. 2018. Video game loot boxes are psychologically akin to gambling. *Nature human behaviour* 2, 8 (2018), 530–532.
  - [23] Magy Seif El-Nasr, Anders Drachen, and Alessandro Canossa. 2016. *Game analytics*. Springer.
  - [24] Electronic Arts. 2017. *Star Wars Battlefront 2*. [Digital Game].
  - [25] Mona Erfani, Magy Seif El-Nasr, David Milam, Bardia Aghabeigi, Beth Aileen Lameman, Bernhard E Riecke, Hamid Maygoli, and Sang Mah. 2010. The effect of age, gender, and previous gaming experience on game play performance. In *IIIP Human-Computer Interaction Symposium*. Springer, 293–296.
  - [26] Motahhare Eslami, Kristen Vaccaro, Min Kyung Lee, Amit Elazari Bar On, Eric Gilbert, and Karrie Karahalios. 2019. User Attitudes towards Algorithmic Opacity and Transparency in Online Reviewing Platforms. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–14.
  - [27] André R Gagné, Magy Seif El-Nasr, and Chris D Shaw. 2011. A Deeper Look at the use of Telemetry for Analysis of Player Behavior in RTS Games. In *International Conference on Entertainment Computing*. Springer, 247–257.
  - [28] Ben Geisler. 2004. Integrated machine learning for behavior modeling in video games. In *Challenges in game artificial intelligence: papers from the 2004 AAAI workshop*. AAAI Press, Menlo Park. 54–62.
  - [29] Sabine M Grüsser, Ralf Thalemann, and Mark D Griffiths. 2006. Excessive computer game playing: Evidence for addiction and aggression? *Cyberpsychology & behavior* 10, 2 (2006), 290–292.
  - [30] Fabian Hadiji, Rafet Sifa, Anders Drachen, Christian Thureau, Kristian Kersting, and Christian Bauckhage. 2014. Predicting player churn in the wild. In *2014 IEEE Conference on Computational Intelligence and Games*. IEEE, 1–8.
  - [31] Juho Hamari, Kati Alha, Simo Järvelä, J Matias Kivikangas, Jonna Koivisto, and Janne Paavilainen. 2017. Why do players buy in-game content? An empirical study on concrete purchase motivations. *Computers in Human Behavior* 68 (2017), 538–546.
  - [32] Brent Harrison. 2012. Creating Model-Based Adaptive Environments Using Game-Specific and Game-Dependent Analytics. In *Eighth Artificial Intelligence and Interactive Digital Entertainment Conference*.
  - [33] Jannicke Baalsrud Hauge, Riccardo Berta, Giusy Fiucci, Baltasar Fernández Manjón, Carmen Padrón-Nápoles, Wim Westra, and Rob Nadolski. 2014. Implications of learning analytics for serious game design. In *2014 IEEE 14th international conference on advanced learning technologies*. IEEE, 230–232.
  - [34] Celia Hodent. 2019. Cognitive Psychology Applied to User Experience in Video Games.
  - [35] Andreas Holzinger. 2018. From machine learning to explainable AI. In *2018 World Symposium on Digital Intelligence for Systems and Machines (DISA)*. IEEE, 55–66.
  - [36] Danial Hooshyar, Moslem Yousefi, and Heuiseok Lim. 2018. Data-driven approaches to game player modeling: a systematic literature review. *ACM Computing Surveys (CSUR)* 50, 6 (2018), 90.
  - [37] Danial Hooshyar, Moslem Yousefi, and Heuiseok Lim. 2019. A systematic review of data-driven approaches in player modeling of educational games. *Artificial Intelligence Review* 52, 3 (2019), 1997–2017.
  - [38] Britton Horn, Amy K Hoover, Jackie Barnes, Yetunde Folajimi, Gillian Smith, and Casper Hartevelde. 2016. Opening the black box of play: Strategy analysis of an educational game. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play*. 142–153.
  - [39] Xiao Huang, Sharon Oviatt, and Rebecca Lunsford. 2006. Combining user modeling and machine learning to predict users' multimodal integration patterns. In *International Workshop on Machine Learning for Multimodal Interaction*. Springer, 50–62.
  - [40] Chaima Jemmali, Erica Kleinman, Sara Bunian, Mia Victoria Almeda, Elizabeth Rowe, and Magy Seif El-Nasr. 2020. MAADS: Mixed-Methods Approach for the Analysis of Debugging Sequences of Beginner Programmers. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*. 86–92.
  - [41] Kristine Jørgensen. 2016. The positive discomfort of spec ops: The line. *Game studies* 16, 2 (2016).
  - [42] Rosa Mikeal Martey, Jennifer Stromer-Galley, Jaime Banks, Jingsi Wu, and Mia Consalvo. 2014. The strategic female: gender-switching and player behavior in online games. *Information, Communication & Society* 17, 3 (2014), 286–300.
  - [43] Matt McCormick. 2001. Is it wrong to play violent video games? *Ethics and Information Technology* 3, 4 (2001), 277–287.
  - [44] M McKenzie and SC Wong. 2015. Subset selection of training data for machine learning: a situational awareness system case study. In *Next-Generation Robotics II; and Machine Intelligence and Bio-inspired Computation: Theory and Applications IX*, Vol. 9494. International Society for Optics and Photonics, 94940U.
  - [45] Benedikte Mikkelsen, Christoffer Holmgård, and Julian Togelius. 2017. Ethical considerations for player modeling. In *Workshops at the Thirty-First AAAI Conference on Artificial Intelligence*.
  - [46] Tim Miller, Piers Howe, and Liz Sonenberg. 2017. Explainable AI: Beware of inmates running the asylum or: How I learnt to stop worrying and love the social and behavioural sciences. *arXiv preprint arXiv:1712.00547* (2017).
  - [47] Dinara Moura, Magy Seif el Nasr, and Christopher D Shaw. 2011. Visualizing and understanding players' behavior in video games: discovering patterns and supporting aggregation and comparison. In *ACM SIGGRAPH 2011 game papers*. 1–6.
  - [48] Peter Mozelius, Thomas Westin, Mats Wiklund, and Lena Norberg. 2016. Gaming habits, study habits and compulsive gaming among digital gaming natives. In *The 10th European Conference on Games Based Learning (ECGBL)*, Paisley, United Kingdom, 6th to 7th October 2016. Academic Conferences Publishing.
  - [49] Daniel J Navarro, Thomas L Griffiths, Mark Steyvers, and Michael D Lee. 2006. Modeling individual differences using Dirichlet processes. *Journal of mathematical Psychology* 50, 2 (2006), 101–122.
  - [50] Truong-Huy D Nguyen, Magy Seif El-Nasr, and Alessandro Canossa. 2015. Glyph: Visualization Tool for Understanding Problem Solving Strategies in Puzzle Games.. In *FDG*.
  - [51] Truong-Huy D Nguyen, Shree Subramanian, Magy Seif El-Nasr, and Alessandro Canossa. 2014. Strategy Detection in Wuzzit: A Decision Theoretic Approach. In *International Conference on Learning Science Workshop on Learning Analytics for Learning and Becoming a Practice*.
  - [52] Rune Kristian Lundedal Nielsen and Pawel Grabarczyk. 2019. Are Loot Boxes Gambling? Random reward mechanisms in video games. *Transactions of the Digital Games Research Association* 4, 3 (2019).
  - [53] Cathy O'neil. 2016. *Weapons of math destruction: How big data increases inequality and threatens democracy*. Broadway Books.
  - [54] Petr Parshakov and Marina Zavertieva. 2015. Success in eSports: Does country matter? Available at SSRN 2662343 (2015).
  - [55] Matthew Thomas Payne. 2014. War bytes: the critique of militainment in Spec Ops: The Line. *Critical Studies in Media Communication* 31, 4 (2014), 265–282.
  - [56] Falko Weigert Petersen, Line Ebdrup Thomsen, Pejman Mirza-Babaei, and Anders Drachen. 2017. Evaluating the onboarding phase of free-to-play mobile games: A mixed-method approach. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*. 377–388.
  - [57] Margaret M Recker and Peter Pirolli. 1995. Modeling individual differences in students' learning strategies. *The Journal of the Learning Sciences* 4, 1 (1995), 1–38.
  - [58] Drew Roselli, Jeanna Matthews, and Nisha Talagala. 2019. Managing Bias in AI. In *Companion Proceedings of The 2019 World Wide Web Conference*. ACM, 539–544.
  - [59] Andrew Salmon. 2010. Jail for couple whose baby died while they raised online child. *Prius Online* (2010).
  - [60] Cuihua Shen, Rabindra Ratan, Y Dora Cai, and Alex Leavitt. 2016. Do men advance faster than women? Debunking the gender performance gap in two massively multiplayer online games. *Journal of Computer-Mediated Communication* 21, 4 (2016), 312–329.
  - [61] Keng Siau and Weiyu Wang. 2018. Building trust in artificial intelligence, machine learning, and robotics. *Cutter Business Technology Journal* 31, 2 (2018), 47–53.

- [62] Miguel Sicart. 2009. The banality of simulated evil: designing ethical gameplay. *Ethics and information technology* 11, 3 (2009), 191–202.
- [63] Miguel Sicart. 2010. Values between systems: Designing ethical gameplay. In *Ethics and game design: Teaching values through play*. IGI global, 1–15.
- [64] Miguel Sicart. 2010. Wicked games: on the design of ethical gameplay. *DESIRE* 10 (2010), 101–111.
- [65] Adam M Smith, Chris Lewis, Kenneth Hullet, Gillian Smith, and Anne Sullivan. 2011. An inclusive view of player modeling. In *Proceedings of the 6th International Conference on Foundations of Digital Games*. ACM, 301–303.
- [66] Stacy L Smith, Ken Lachlan, and Ron Tamborini. 2003. Popular video games: Quantifying the presentation of violence and its context. *Journal of Broadcasting & Electronic Media* 47, 1 (2003), 58–76.
- [67] GameSpot Staff. 2017. Star Wars Battlefront 2's Loot Box Controversy Explained. *Gamespot*. <https://www.gamespot.com/articles/star-wars-battlefront-2s-loot-box-controversy-expl/1100-6455155> (2017).
- [68] Bernd Carsten Stahl and David Wright. 2018. Ethics and Privacy in AI and Big Data: Implementing Responsible Research and Innovation. *IEEE Security & Privacy* 16, 3 (2018), 26–33.
- [69] Lucy A Suchman. 1987. *Plans and situated actions: The problem of human-machine communication*. Cambridge university press.
- [70] Yaoyao Sun. 2017. Motivation to play esports: case of League of Legends. (2017).
- [71] Ubisoft. 2016. Tom Clancey's The Division. [Digital Game].
- [72] Valve Corporation. 2013. Defense of the Ancients 2. [Digital Game].
- [73] MJJM Versteeg. 2013. *Ethics & Gamification design: a moral framework for taking responsibility*. Master's thesis.
- [74] Richard J Viken, Teresa A Treat, Robert M Nosofsky, Richard M McFall, and Thomas J Palmeri. 2002. Modeling individual differences in perceptual and attentional processes related to bulimic symptoms. *Journal of Abnormal Psychology* 111, 4 (2002), 598.
- [75] James Vincent. 2018. Google's fixed racist algorithm by removing gorillas from its image-labeling tech. *The Verge* 12 (2018).
- [76] David I Waddington. 2007. Locating the wrongness in ultra-violent video games. *Ethics and Information Technology* 9, 2 (2007), 121–128.
- [77] Günter Wallner, Nour Halabi, and Pejman Mirza-Babaei. 2019. Aggregated visualization of playtesting data. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–12.
- [78] Günter Wallner and Simone Kriglstein. 2012. A spatiotemporal visualization approach for the analysis of gameplay data. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 1115–1124.
- [79] Günter Wallner and Simone Kriglstein. 2013. Visualization-based analysis of gameplay data—a review of literature. *Entertainment Computing* 4, 3 (2013), 143–155.
- [80] Chin-Sheng Wan and Wen-Bin Chiou. 2006. Psychological motives and online games addiction: A test of flow theory and humanistic needs theory for Taiwanese adolescents. *CyberPsychology & Behavior* 9, 3 (2006), 317–324.
- [81] Danding Wang, Qian Yang, Ashraf Abdul, and Brian Y Lim. 2019. Designing theory-driven user-centric explainable AI. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–15.
- [82] Su Xue, Meng Wu, John Kolen, Navid Aghdaie, and Kazi A Zaman. 2017. Dynamic difficulty adjustment for maximized engagement in digital games. In *Proceedings of the 26th International Conference on World Wide Web Companion*. 465–471.
- [83] Georgios N Yannakakis, Pieter Spronck, Daniele Loiacono, and Elisabeth André. 2013. Player modeling. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.
- [84] David Zendle and Paul Cairns. 2018. Video game loot boxes are linked to problem gambling: Results of a large-scale survey. *PloS one* 13, 11 (2018), e0206767.
- [85] Haiyi Zhu, Bowen Yu, Aaron Halfaker, and Loren Terveen. 2018. Value-sensitive algorithm design: method, case study, and lessons. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 194.
- [86] Jichen Zhu, Antonios Liapis, Sebastian Risi, Rafael Bidarra, and G Michael Youngblood. 2018. Explainable AI for designers: A human-centered perspective on mixed-initiative co-creation. In *2018 IEEE Conference on Computational Intelligence and Games (CIG)*. IEEE, 1–8.
- [87] Georg Zoeller. 2013. Game development telemetry in production. In *Game analytics*. Springer, 111–135.

This figure "sample-franklin.png" is available in "png" format from:

<http://arxiv.org/ps/2006.10808v1>