"Hey Players, there is a problem...": On Attribute Inference Attacks against Videogamers

Linus Eisele, Giovanni Apruzzese
Liechtenstein Business School, University of Liechtenstein
{name.surname}@uni.li

***Abstract*—We focus on a subtle privacy issue that affects (potentially hundreds of) millions of videogamers: attribute inference attacks (AIA). Through AIA, evildoers can infer gamers' *private* attributes (e.g., age, gender, occupation) by leveraging in-game statistics that are *publicly available*. Despite some research efforts revealing the practicality of AIA in DOTA2, such a threat has been mostly ignored by the overarching gaming community. This is a problem: AIA can only be mitigated through the cooperation of the entire gaming community—and this cooperation can only begin if all stakeholders *acknowledge* the threat of AIA.**

We seek to promote such a positive change by raising the gaming ecosystem's awareness about AIA. First, we provide evidence that AIA have truly been overlooked in the gaming domain. Then, we scrutinize the gaming landscape, pinpointing (i) the games that are more prone to AIA, and (ii) the respective communities that facilitate the enactment of AIA. Finally, through an (ethical) user survey (n=516) resembling a fundamental data-collection step of AIA, we (iii) proactively assess the threat of AIA. We advocate gamers and developers to reflect upon our findings—which we disseminate in an educational campaign: the subtle threat of AIA cannot be countered solely by researchers.

I. INTRODUCTION

Video games represent the world's leading entertainment industry [1], totaling over $250 billion in 2023 [2]. These numbers are driven by the immense popularity of videogames across all age groups, genders, cultures, and income levels [1]: today, over 40% of the World's population play videogames [3]. Online multi-player videogames, in particular, are preferred [4] to single-player videogames—predominantly due to their intrinsic trait of enabling social interactivity [5–7].

Unfortunately, players of such videogames are exposed to various privacy threats [1]. The gaming ecosystem generates large amounts of data which may "leak" information about the players themselves. For instance, in 2014 [8], it was *suggested* that—by using data accessible only to game-developers—it could be possible to infer certain personal attributes of a given player (e.g., their gender, or their age). Such a possibility was confirmed by numerous studies (e.g., [9]), showing correlations between players' *(a)* in-game activities and their *(b)* off-game personal attributes—which could be used by developers to improve their products [10]. While arguably raising some ethical concerns, such "profiling" activities are not necessarily malicious. Yet, in 2023, a new study **proved** that such correlations between in- and off-game data can be maliciously exploited to launch a subtle form of privacy violation: "attribute inference attacks," or AIA [7].

979-8-3503-5067-8/24/$31.00 ©2024 IEEE

The fundamental (and dangerous) aspect of AIA is that they can be carried out by leveraging *publicly available data*. Indeed, in AIA (see Fig. 1), the attacker relies on in-game statistics of players retrieved by the so-called "tracking websites" (e.g., stratz.com), which are openly accessible. Such statistics (e.g., win/loss ratio, favourite weapon/hero) are used to infer a given player's personal attributes—which are *private information*. Through a user study of 484 DOTA2 players, Tricomi et al. [7] demonstrated that it is possible to, e.g., identify underage players in a population with ≈100% precision.

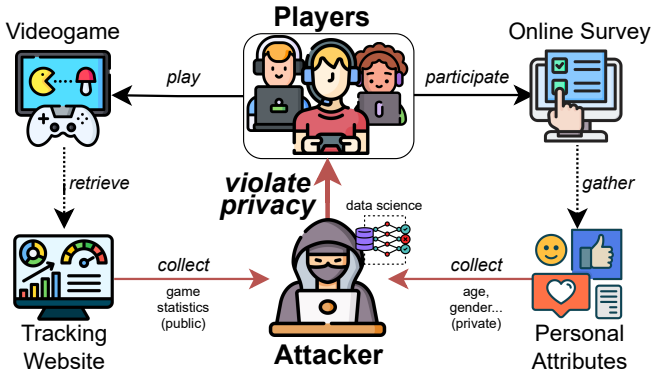

Fig. 1: **Attribute Inference Attacks in Videogames** – The in-game statistics of players are retrieved by "tracking websites", and are *publicly* accessible. An attacker can collect such data, and then distribute "online surveys" used to collect *private* personal attributes (e.g., age, gender, occupation...) of participating players. Finally, by exploiting data science techniques, it is possible to infer information about *other* players—violating their privacy.

According to Tricomi et al. [7], the only way to mitigate AIA (without irreparably disrupting the gaming ecosystem) is through a joint effort between *players* and *developers*: the former should be more "mindful of privacy," whereas the latter should provide more "privacy options" that the players can use to protect themselves. There is, however, a problem: the findings by Tricomi et al. [7] received **little visibility**. As of March 2024, the paper [7] received only 5 citations (according to Google Scholar) – despite having first appeared online in Dec. 2022. Moreover, we reached out to the authors of this publication, and they revealed that they informed VALVE of the threat of AIA in Nov. 2022, but received no response.

In this vision paper, we reinforce the message broadcast by Tricomi et al. [7]. We believe AIA are a threat that should not be underestimated in the gaming landscape. To "open the eyes" of the entire gaming community (and hence inspire a privacy-friendly change against AIA), we shed further light on this subtle threat—elucidating why gamers should care.

CONTRIBUTIONS. We seek to raise awareness of AIA in videogames. After summarizing the quintessential properties of AIA and providing factual evidence that such a privacy threat has been overlooked in game-related research (§II), we:

- pinpoint the *games "theoretically" prone to AIA* (§III). We systematically analyse the gaming landscape and **identify 11 titles** which bear high risk of being targeted by AIA.
- scrutinize which *gaming communities "practically" enable AIA* (§IV); With a structured approach, we **find 23 communities** facilitating the data-collection phase to setup AIA;
- through a **user survey (n=516)**, we gauge the extent to which *players would "unconsciously contribute"* to AIA (§V).

We also elucidate some concealed "pros-and-cons," which we use to define desiderata (§VI) for future work, and disseminate our findings to educate gaming communities about AIA.

II. BACKGROUND AND MOTIVATION

To appreciate our contributions, we describe the major characteristics of AIA (§II-A), and highlight the research gap (§II-B).

A. Gentle introduction to AIA in the Videogame Ecosystem

To elucidate why AIA are a concerning issue, we emphasize the necessary steps to enact AIA, and highlight the ways in which the gaming community can be affected by AIA.

*1) **Requirements**:* AIA are possible thanks to the capabilities of machine learning models to find hidden patterns in data after undergoing a training phase [11, 12]. In the scenario envisioned by Tricomi et al. [7], a model is trained to associate *(a)* a player's in-game statistics with *(b)* their personal attributes—the latter being the ground truth used to guide the learning phase of the model. Hence, to enact AIA an attacker needs to collect a certain amount of such "associations" which will enable the model to infer the personal attributes of other players (i.e., those not included in the training dataset).

*2) **Collection**:* There are many ways an attacker can use to gather the associations required to develop an AIA-ready model [7]. However, real attackers operate with a cost/benefit mindset [13], favoring tactics that are cheap to stage. Unfortunately, the ecosystem of multiplayer videogames makes it simple for evildoers to prepare an AIA: *tracking websites are a reservoir of in-game statistics*, and players publicly share their profiles [6]; whereas *personal attributes are obtainable by exploiting the social nature of videogamers*. In particular, *online surveys* are a convenient means of doing so: thousands of players provided their "personal attributes" in two DOTA2 surveys shared on reddit [14]; Tricomi et al. [7] distributed their questionnaire on other social networks, collecting hundreds of responses with minimal effort. It is even possible to "recruit" participants by offering a small compensation as an incentive (e.g., [9, 15, 16]). Of course—contrarily to all these research works—a real attacker would do so "unethically."

*3) **Consequences**:* With an AIA-ready model, the attacker can hence exploit its predictive capabilities to infer the personal attributes of "unknown" players. Such a weapon can be used in three exemplary ways [7]: *(i)* Given the in-game statistics of a player, infer its personal attributes. For instance, a mischievous player may want to see if their opponents can be verbally harassed during a match [17, 18]. *(ii)* Given the in-game statistics of multiple players, find associations "in bulk." For instance, an attacker can scrape the public profiles of many players, infer those who are more willing to purchase in-game content, and sell such information to advertising companies for targeted ads [19]. *(iii)* Identify specific individuals among a set of players. This is a variant of the previous way: an attacker may want to pinpoint the underage players among all the users of a given tracking website—and then bully them [20].[1]

⚠ **A subtle threat.** The gaming ecosystem is exposed to AIA due to the intrinsic "social" nature of its players. However, the elusive side of AIA is that even those players who are aware of privacy issues can fall victim to AIA.[a]

[a] If "unaware" players participate in (unethical) surveys, the attacker can use such data for AIA and infer the personal attributes of "aware" players.

B. AIA and Privacy in Gaming Research (Related Work)

To further motivate the need for this vision paper, we show that the themes of AIA and *privacy* in general tend to be overlooked in related literature—starting from IEEE CoG.

*1) **IEEE CoG**:* This venue changed its name from "CIG" to the current one in 2019. Hence, we take all papers that have appeared in the proceedings of IEEE CoG from 2019–2023, obtaining 682 papers. Next, we perform a keyword search for the terms "AIA", "attribute inference attack" and "privacy". As expected, we find no match for the first two—despite the term having gained visibility in the security community since 2016 [12], but in the social network context. For "privacy", we have 20 papers (3%) that mention the term at least once: a quick investigation reveals that for 10 of these papers the term appears only in references/appendices, whereas in 6 papers it is mentioned only for "data collection procedures" (e.g., to indicate that user studies have been done respectfully of participants' privacy); in 3 cases it is mentioned out of context.[2] The only relevant match (out of 682) is [22], which hints at potential privacy issues in the ethical statement.[3]

*2) **Google Scholar**:* We refine our analysis by querying Google Scholar (in Dec. 2023) with the term "attribute inference attack" and "game". We specifically look for papers that deal with this issue in the gaming ecosystem (some works [23] mention "game theory", which is outside our scope). We could not find any match besides the work by Tricomi et al. [7]. We then expanded our search by looking for game-related studies that look for associations between players' in-/off-game data, scrutinizing whether such papers hint at potential "attacks" or "privacy violations" that can be exploited by leveraging such associations with publicly available resources. We could not find any match: all such papers do not stress the privacy implications (towards players) of their own findings.[4] Some

[1] For low-level technical details about machine learning for AIA, see [7, 11].

[2] Intriguingly, it occurs 33 times in [21] because it considers "how *bathrooms* in games convey the sense of 'privacy' expected by real-world ones."

[3] Even [9], despite collecting 2.5k responses pertaining to LoL, has no occurrences of either "privacy" or "ethic" (and, of course, "AIA").

[4] E.g., there are 6.7k players considered in [24], over 4k in [25], and 1k in [26], but the term "privacy" never occurs in either (and [26] is from PETS).

papers carry out privacy-centered user studies (e.g., [27]), but do not focus on in-/off-game associations. Others raise the attention [1] on the data collection policies in the gaming ecosystem, but without carrying out any original investigation.

PROBLEM STATEMENT. AIA are overlooked by game-related research.[a] We challenge the status quo, and: examine the current gaming landscape, revealing *where* and *why* AIA can be staged; gauge the players' *awareness and perception of AIA*; and outline *what can be done* by all stakeholders to collectively mitigate the threat of AIA in videogames.

[a] AIA are known in the security/privacy domain, but we want to make the *gaming* community aware of this threat—and CoG is the best venue for this.

III. VIDEOGAMES PRONE TO AIA (IN THEORY)

As our first major scientific contribution, we pinpoint the games that "theoretically" enable the enactment of AIA, and explain how to do so systematically (§III-A).[5] We will use these findings (§III-B) as a basis for the rest of our research.

A. Criteria (how to determine if a game is prone to AIA?)

To reach our first goal, we examine the landscape of multi-player videogames under the lens of an attacker willing to carry out AIA. Specifically, we ask ourselves: "what games present the characteristics that would make an AIA *economically attractive* and *practically viable*?" We answer this question by reflecting on our previous explanation (§II-A2). Hence, we derive an **original list of assessment criteria**, centered on the attacker's cost/benefit mindset [13], that must be scrutinized to determine whether a videogame is "AIA-prone." We express our list through four high-level questions:

- "*Is the game popular?*" Setting up an AIA requires expertise and a resource investment (for data collection, filtering, and model training), hence games with a small playerbase may not be attractive for real attackers.
- "*Do tracking websites exist for the game?*" Tracking websites are essential to ensure that the AIA is feasible (collecting in-game statistics from the game itself is doable, but much less practical [7]).
- "*Does the playerbase contribute to online surveys?*" If true, then it would be a signal that harvesting the ground truth (via "unethical" surveys) will yield practical results (i.e., many and heterogeneous responses) for the attacker.
- "*Have correlations between players' in-/off-game data been found for this game?*" If a prior study proved the existence of such correlations (e.g., [9]), then an attacker can use it as a scaffold for preparing the AIA.[6]

In summary,[7] a videogame for which the answer to all the abovementioned questions is "yes" is an AIA-prone title.

[5] Tricomi et al. [7] listed some games wherein AIA could be conceived, but such a list included only eSport (which are a subset of multiplayer games), and was not derived with a systematic approach (which we adopt and propose).

[6] AIA require the existence of correlations that allow the model to associate public with private data [7]. The attacker can find the correlations autonomously (as done in [7]) but this increases the cost of the campaign.

[7] We present high-level criteria. However, depending on the attacker's objective, there may be additional fine-grained ones, e.g., "*do underage people play the game?*" or "*is this game popular among people of a certain gender?*"

⚠ **An upsetting scenario.** Our criteria rely on the assumption that an attacker develops an AIA-ready model *from scratch*. However, it is entirely possible to "share" an AIA-ready model, which can be used at no cost by any entity with malicious intentions. We hope this is not already happening.[a]

[a] Albeit darkweb marketplaces do deal with similar "merchandise" [28].

B. Findings (what games are prone to AIA in 2024?)

Our investigation follows the criteria presented in §III-A.

Method. We begin by looking for those (multi-player) games having a large playerbase. Hence, we rely on popular websites (e.g., activeplayer, steamcharts, playercounter) to derive a list of the 20 games having the largest number of concurrent players. Then, for each game, we

- search for a *tracking website*, i.e., a platform which must be *(i)* publicly available, and which *(ii)* allows to retrieve extensive in-game statistics of a large subset of the game's population, which could *(iii)* potentially be useful for AIA (for instance, we exclude "achievement trackers");
- search for *prior surveys*, through Google queries with the title and two terms among "player,survey,results,reddit". We consider only surveys with >200 responses, and disseminated by individuals unrelated to the game-devs;
- search for *literature-found correlations* between in-/off-game data, through queries on Google Scholar with the title of the game and any combination of "correlation,analysis,profiling,inference,privacy,personality,prediction". We used the snowball method to further investigate the references of relevant papers.

We perform these operations in Oct. 2023, and we repeat them another time in Feb. 2024 for validation purposes.[8] The results of our analysis are shown in Table I (described in the caption).

Ethical Statement: We acknowledge that the following information may be helpful for evildoers. However, we favor *ethical disclosure*, respecting the right to be informed of gamers [29, 30].

Results. We found tracking websites for 15 games, online surveys for 17, and also that 12 have correlations discussed in related literature. Given our criteria (§III-A), we consider those games (which are "popular" by definition) having all three of these elements (11 in total) to be AIA-prone, and are marked with a red cell in Table I. In contrast, games for which we could not find a tracking website (which is crucial) are marked with a light-blue cell, and we consider them to have a low likelihood to be involved in AIA (e.g., for Minecraft we could not find a tracking website that could enable AIA). As a side note, all papers discussing these correlations *never* mention the word "privacy" (aside from Tricomi et al. [7]).

⚠ **Attackers appreciate these efforts.** Works that announce the existence of correlations between in-/off-game data make the attackers' job easier (cf. §III-A). The attacker can also use results from prior surveys for validation purposes after (unethically) collecting their own dataset (as done in [7]).[a]

[a] Prior surveys may even induce attackers to steal such data for AIA [31].

[8] The operations are done by two authors, who resolved issues via discussions. For some games, we consider also their previous installments, since they tend to be similar (e.g., for BF2042, we consider [24] which is on BF3).

TABLE I: **Games prone to AIA –** We check the 20 games having the highest number of concurrent players. We report the number of active and concurrent players (over the last 30 days), the link to an exemplary tracking website, the number of participants of a representative survey (and its link), and a paper which showed correlations/predictions between the in-/off-game data. Red (blue) cells denote games being prone (not prone) to AIA (according to §III-A); games in boldface are those considered in our following analyses.

Game	Popularity Active – Concurr	Tracking Website?	Prior Survey?	Correl. Found?
LoL	142M – 900K	📈	3.7k	[9]
WoW	32M – 250K	📈	500	[32]
CSGO	31M – 900K	📈	13k	[33]
Fortnite	237M – 1.15M	📈	1k	[34]
PUBG	320M – 200K	📈	4.4k	[35]
OW2	25M – 350K	📈	3.2k	[33]
Valorant	24M – 600K	📈	1.4k	[36]
CoD:WZ	71M – 300K	📈	751	[37]
RS:S	10M – 120K	📈	4.8k	[33]
Destiny2	14M – 50K	📈	450	[38]
DOTA2	14M – 430K	📈	7.3k	[7]
Apex	52M – 250K	📈	296	
RktLg	85M – 220K	📈	6k	
GTA:O	24M – 110K	📈	1.9k	
BF2042	300K – 15K	📈		[24]
FIFA	5M – 50K			
Minecraft	169M – 900K		4k	
Roblox	213M – 1.5M		1.5k	
HeartStone	6M – 370K		21k	
Wildlands	370K – 5K			

In the next section, we will examine some communities of all games in boldface in Table I, i.e., those for which we found an "AIA-compliant" tracking website—but we exclude DOTA2 since it has been extensively covered by Tricomi et al. [7].

IV. COMMUNITIES THAT FACILITATE AIA (IN PRACTICE)

Here, we further scrutinize the ecosystem of 14 games by examining their *communities*, i.e., (online) platforms wherein users interested a common subject (ideally a game) tend to meet and socially interact with each other. Such a contribution serves to identify which ecosystems should be prioritised in the fight against AIA—by gamers, developers, and researchers.

A. Reflection (in what way would a community facilitate AIA?)

Even if a game has a tracking website, the attacker must still collect the data (and, particularly, ground truth for personal attributes) necessary to train an AIA-ready model (§II-A).

Context. We put ourselves in the attacker's shoes once more. Hence, after having identified 14 games (§III-B) wherein AIA can be staged, we ask ourselves "what would the attacker do now?" The answer is straightforward: they would investigate the communities of such games, and devise ways to deceive users to "release" their in-/off-game–data associations—and the best way to do so is via online surveys [7]. However, while *making* a survey is trivial, *ensuring* that such a survey achieves the intended (malicious) purpose is not.[9]

Challenges. Online social networks, such as reddit or forums, can have millions of users. To prevent "spam" which would annoy the community, the administrators typically enforce content moderation policies. For instance, new accounts may be prevented from posting in some sections of a given board; or certain types of content (e.g., links) may have to be approved before being publicly displayed. Therefore, an attacker attempting an AIA must deal with such obstacles, and hence must review the guidelines of each community to understand *what can be done*. Such a "review" is exactly what we will do in this analysis—but we will do so ethically.

Structured Approach. At a high-level, a community that "facilitates AIA" is one that makes it simple for an attacker to collect ground truth via online surveys. We scrutinize such "simplicity" through three steps. Given a game, we first search the Web for relevant communities, excluding those that are clearly inactive or have a small population, and derive a list of candidate communities. For each of these, we **(1)** Review its content moderation policies. Four cases can occur: *(1a)* if surveys are explicitly prohibited, we remove this community from our list; *(1b)* if surveys are explicitly allowed, we include this community among those that facilitate AIA; *(1c)* if there is no mention of "surveys" OR *(1d)* if surveys require mod-approval, we investigate it further. Specifically, we **(2)** Contact the community managers, asking for their permission to post a public survey (if no clear contact is provided, we go directly to step-3). Three cases can occur: *(2a)* if we are denied permission, we remove this community from our list; *(2b)* if we are given permission, we include this community among those that facilitate AIA; *(2c)* if we get no response after 14 days, we post the survey (unless *(1d)* is true). Then, we **(3)** Wait and see: *(3a)* if the post is deleted, we remove the community from our list; whereas *(3b)* if no action is taken by the mods, we include the community among those that facilitate AIA.

Ethical Statement: We never post a survey if it would violate the policies of a community. In messaging the community managers, we informed them of our research and of the goal of the survey.

From an attacker's perspective, achieving *(1b)* is better than *(2b)*, which is better than *(3b)*, since reaching *(3b)* takes longer time. N.B.: We will discuss the survey in the next section (§V).

B. Analysis (what communities make AIA easier to stage?)

Let us explain how we applied our structured approach for our low-level analysis, schematically depicted in Fig. 2.

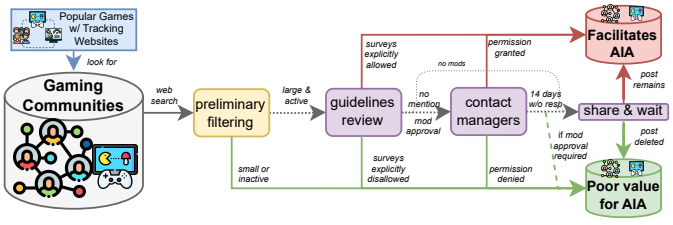

Fig. 2: **Workflow for the Community Analysis** – We scrutinize which gaming communities make it simpler for attackers to stage AIA – under the assumption that the attacker gathers the ground truth via online surveys. After identifying communities with a large userbase, we review the guidelines; if necessary, we contact the administrators to get their permission: if we receive no response, we post a survey in a given community, and see what happens.

Method. For every game among the 14 we consider, we look in well-known forums, on reddit, and also in popular Twitch streams: we set ourselves the target of finding 50 "large and active" communities (we cannot search the whole Web!).

[9] An attacker can certainly attempt a brute-force search, and create (or purchase) a plethora of fake accounts/bots which continuously post the link to the survey on various social networks—potentially promising a reward to increase the response rate. While costly, this can be an effective strategy which we will not pursue (due to its triviality, and also for being unethical).

Importantly, we focus on international communities, i.e., we do not consider communities of a specific language/region (albeit this would very well be within the reach of an attacker). For each potential community, we do a preliminary (qualitative) analysis to determine if it matches our criteria (e.g., we exclude /r/Gamingthoughts because it has barely 120 redditors). We eventually found 50 communities (among which we have 70 streamers, which we count as a single "Twitch" community); most of these communities are from Reddit. After reading their guidelines (if available), 3 explicitly allowed surveys, whereas 12 did not (e.g., Steam), 18 did not mention any rule related to surveys (or did not have rules to begin with), and 17 required mod approval. We then contacted the community managers (we could not find any specific contact for 5 communities). After writing 129 messages, we received permission by 12 communities, whereas 10 did not give us permission (notably, Twitch streamers were not collaborative), and we did not receive any response from 7; among these, 3 (e.g., r/privacy) required approval by mods, but since we received no response we did not proceed onwards. We then posted our survey for the 10 remaining communities: in 2 cases, the post was deleted, whereas in 8 cases no action was taken. (N.B.: we ultimately posted the survey also in the 15 other communities for which we obtained permission. We will discuss the survey in §V)

Strict rules / admins. Some communities have rules that explicitly prohibit sharing public surveys. We even contacted the respective community administrators, and they made it clear that any link to a survey would be deleted. This is a *positive finding* in light of countering AIA.[a]

[a]Even though this may prevent some (ethical) investigations, such strict policies also prevent (unethical) distribution of surveys that facilitate AIA.

Findings. We report in Table II the list of 23 communities that reached step *(1b)* (✓), *(2b)* (👍) or *(3b)* (🕒) of our approach during our analysis (and wherein we posted surveys). Hence, such communities can be considered as "facilitating AIA," since an attacker would be able to exploit their "openness" to distribute (unethical) surveys. A 📝 denotes those communities for which it was explicitly required to obtain the mod approval to post surveys, whereas a ? is when there was no mentioning of surveys in the guidelines. Communities in red make it straightforward to share surveys. Communities with ? can be considered equivalent to red ones, since the attacker can post the survey and it would not be taken down. The "harder" communities to exploit (at least from a theoretical perspective) are those with a 📝, since it is required to get permission from the mods—but even in this case, it only takes a short message exchange (which we did ethically, but an attacker can leverage social engineering techniques [39] to trick the community manager). Overall, the "red" communities count over 1.5M of users (according to their statistics).

V. Assessing and Educating Players about AIA

As our final contribution, we now focus on the *user survey*, which we only mentioned in the previous section. The purpose is twofold: *assessing* the risk of AIA in our considered games and communities, and *educating* players on privacy threats.

TABLE II: **Communities that Facilitate AIA –** We report the 23 communities (having players of the 14 games boldfaced in Table I) that, after our analysis (Fig. 2) can be classified as facilitating AIA—due to either allowing surveys by default (red), or after messaging the mod (blue), or which did not remove our survey after we posted it (gray cells do not have specific mods to contact; yellow cells have mods, but did not respond to our request after 14 days).

Community Large and Active	Survey Allowed?	Admin Response?	Msgs Sent
truegaming	✓		
SampleSize	✓		
JoyFreak	✓		
Rainbow6	📝	👍	1
wow	📝	👍	1
leagueoflegends	📝	👍	6
Overwatch	📝	👍	4
VALORANT	📝	👍	3
youtubegaming	📝	👍	1
GameTheorists	?	🕒	1
videogames	?	🕒	1
consoles	?	🕒	1
AskGames	?	👍	1
MMORPG	?	👍	2
playstation	?	🕒	2
ubisoft	?	👍	1
Instant Gaming	?	👍	1
RocketLeague	?	👍	3
gamers	?	👍	1
PC Gamer	?	–	
COD Forums	?	–	
Valorant Forums	?	–	
GTA Forums	?	–	

A. Survey Design

Our questionnaire is publicly observable in our repository [40].

Overview. Our survey is meant to represent the actual data-collection phase from the perspective of a real attacker who wants to setup an AIA—which takes place after the previous "intelligence gathering" phases (which we simulated in §III and §IV). Hence, we will distribute our survey in the communities in Table II. Our survey has a similar structure to the one by Tricomi et al. [7] with two major differences. **(1)** Our survey is shorter. Specifically, we will ask less questions about personality and the game than [7]. This is because *our intent is not to enact an AIA*: we simply want to assess whether the participants to our survey "are willing" to release their personal attributes. **(2)** Our survey assumes a "multi-game" setting. The survey in [7] focused only on DOTA2, whereas ours focuses on the 14 games in boldface in Table I.

Ethical Statement: Our institutions are aware of our research. We informed the participants that: *(i)* our study was for research; *(ii)* the questionnaire was anonymous; *(iii)* no data would be released. Participation was voluntary and we did not offer any compensation.

Organization. The questionnaire has five sections:

1) *Demographics.* We ask exactly the same questions as [7].
2) *Personality.* We ask six personality-related questions.
3) *Gaming.* We first inquire for the most played game (among the 14 we consider; plus an "other" category). Then, depending on the choice, we ask three questions; one for the gamertag (necessary for AIA), one for vali-

dation purposes (e.g., "what is your most played hero?", which we can verify with the gamertag), and one for generic knowledge about the game (an attention check).

4) *Privacy,* We ask eight privacy-related questions, and inquire for concerns about privacy issues in videogames.
5) *Extensions.* We ask three questions, inquiring if the participant regularly plays any other game among our considered 14, and which communities they follow.

We distribute our questionnaire for every community (Table II) in Dec.2023/Jan.2024, and collect responses over 3 weeks.[10]

Ethical Statement: We acknowledge that our "shorter" questionnaire (w.r.t. [7]) could be seen as a limitation from the perspective of a "real AIA", but *we do not want to pursue this objective*. Our primary goal is gauging how users of our considered communities respond to surveys—which could be unethically used for AIA. *We do not want to assist attackers* by providing "novel" information about potential correlations obtainable through our survey, which is also why we *(i)* do not set any target number of responses and *(ii)* will not provide details on demographics or personality.

B. Risk Assessment (are these surveys useful to an attacker?)

Here, we focus on three aspects that are most relevant for the sake of our paper—gauging how useful our surveys are from the perspective of an attacker who wants to carry out an AIA.

Which responses are valid? By aggregating the results of all our surveys, we obtain 651 responses. Of these, we remove 29 because they specified an "other" game which did not exist or failed the attention check. We then analyse the remaining 622 responses, scrutinizing which ones are "useful" for an attacker. Specifically, we focus on those answers that *provided a "valid" gamertag*. We find that 106 (16%) answers included an incorrect gamertag, or one which did not match the validation. Intriguingly, we find that in many instances the string provided in the gamertag was *criticizing our survey*. For instance, some participants wrote "NotGivingThatInformation" or "invasive, not answering". We find this intriguing: when we posted our surveys, we clearly specified that the questionnaire required users to provide their gamertag—hence, users were aware of this request.[11] We believe that these "skeptical users" are a *positive result* from the perspective of AIA, since it shows that not all users "blindly" trust requests to fill online questionnaires. However, the remaining 516 (79%) responses can be used for AIA: we will now analyse these.

Mindful players. Some participants of our survey refused to provide their in-game handle. We find it *positive*: there is no true reason for providing such information (which we use for validation), which is vital for attackers to setup AIA.[a]

[a]The username is necessary to associate the ground truth (i.e., personal attributes) to the in-game statistics retrievable through tracking websites [7].

What games do our participants play? We analyse the 516 valid responses, investigating the extent to which our surveys enabled to collect data of our considered 14 games (having a tracking website). We report the results of this analysis in Table III, showing the top10 "popular" games among our participants. Specifically, the first row shows the number of participants which marked a game as their "primary" game, and the second row denotes those who specified the game as "another game that they play regularly" (at the end of the survey). We can see that our questionnaires enabled to solicit (valid) responses from ≈100 players for 7 of our games (of which 5 are AIA-prone), despite our limited efforts in promoting our questionnaire. Importantly: *AIA do not need to consider only "professional" players* (as hinted in [7]). Hence, even if a player does not have one of our games as their primary choice, they would still participate in (unethical) deceitful online questionnaire that could fuel an AIA.

TABLE III: **Most popular games (top10)** – We aggregate the "primary" with "other" games (among the 14 considered) *often played* by participants. The color refers to Table I. Seven games are often played by ≥95 participants.

Game	OW2	LoL	CSGO	WoW	Apex	RS:S	GTA:O	Virnt	Frtnt	Dstny2
Primary	35	57	13	46	9	48	16	28	9	11
Other	113	83	110	75	89	49	79	58	76	59
Total	**148**	**140**	**123**	**121**	**98**	**97**	**95**	86	85	70

What communities are responsive? Lastly, we examine the communities that solicited the most responses to our surveys (recall that we distributed one questionnaire to each community in Table II; we ensure there are no duplicate answers). We report the top-10 most "responsive" communities in Table IV. We see that these communities are all from reddit. Intriguingly, we received no response at all from three communities entailing forums/boards (JoyFreak, COD Forums, Valorant Forums). More generally, we believe users of such communities to be unlikely to "contribute" to AIA. We conclude by inspecting the answers to "which gaming communities do you follow?". The top-3 most common responses are: 390 (76%) "Reddit", 209 (41%) "Discord", 90 (17%) "Steam". Reddit being first is expected (most of our respondents *are* from reddit!). The popularity of Discord (which we did not investigate thoroughly) makes such a channel also viable for AIA. Steam being third is encouraging: the current guidelines of Steam prohibit surveys. However, such guidelines may not be strictly followed by its community: we did find some surveys on the Steam's Community Hub (e.g., [41–44]).

TABLE IV: **Most responsive communities (top10)** – Communities (cf. Table II) from which we received the most (valid) responses to our survey.

Commu.	/r/truegaming	/r/MMORPG	/r/Rainbow6	/r/SampleSize	/r/wow/	/r/leagueoflegends	/r/Overwatch	/r/videogames	/r/VALORANT	/r/GameTheorists
Absolute	207	56	42	35	30	22	21	19	18	18
Relative	40.12%	10.85%	8.14%	6.78%	5.81%	4.26%	4.07%	3.68%	3.49%	3.49%

C. Awareness (is data-privacy in our participants' mind?)

We now focus on the questions in the fourth section of our survey, which inquire the participants' opinion on privacy-related issues in videogames—educating them in the process.

What do you know? The first question asks participants to rate their "knowledge about data collection and privacy issues in videogames"; the answer is in a [1–6] Likert scale (1: novice, 6: expert). Across 516 (valid) responses, the average value is 3.17 (std=1.3), which is below the middle point

[10]We carried out pilot tests with colleagues for feedback (avg length=10m).
[11]Users of 7 communities even voiced such a concern in the thread.

of 3.5 (confirmed with a t-test, $p \ll 0.05$): this denotes that our participants are not very informed about privacy in general.

Have you ever worried? Next, we consider the responses to three (binary) questions: "do you know that your gaming data are being collected by other entities?" and "do you know that it is possible to predict your personal attributes from your gaming data?" (i.e., AIA), and "have you ever worried about your anonymity in games being compromised?". We visualize the responses to these questions with the 3D-plot in Fig. 3, showing the inner relationships between the responses to each questions. At a high-level, 456 (88%) know that their data are being collected, but 122 (24%) *do not know* about AIA, and 293 (57%) *have never worried* about their anonymity being compromised. We find it instructive to further analyse these results: among those (394, 76%) who "know about AIA", 172 have worried about their anonymity, but 222 have not. Given that the overall level of knowledge is below average, this result indicates that even if players "know" (or "suspect") that their personal attributes can be predicted, they may overlook *what can be predicted* (and Tricomi et al. [7] showed that certain attributes are easy to infer, such as age and occupation).

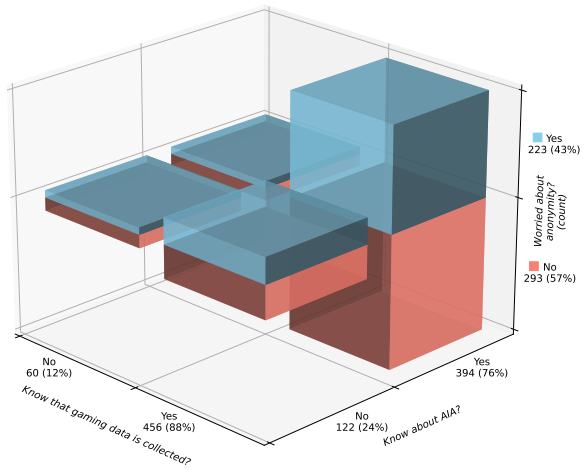

Fig. 3: **Distribution of the responses to three privacy questions** – Each axis denotes a (binary) question; we report the overall number of responses on the axes. The vertical axis shows the "count" of the responses to the question "have you ever worried about your anonymity being compromised?"

Do you share your data? Lastly, we inquire about data-sharing. First, we ask "did you explicitly choose to share your statistics?"; four answers were possible. The results are enlightening: 198 (38%) answered "No (there is no option or I do not know about such an option)"; 160 (31%) answered "Yes"; 95 (18%) answered "I am not sure"; and only 63 (12%) answered "No, I explicitly choose not to". These results underscore the *lack of transparent opt-out options*. As for the last question, "Generally, would you choose to publicly share your personal data (age, gender, marital and economic status etc.)?", 383 (74%) answered "No" and 133 (26%) answered "Yes". We find this turnout alarming: first, because *these participants did provide such data* in our survey; second, because such people clearly do not want such information to be known to others, *but they can be targeted by AIA*—which, as shown in [7], can reveal such personal attributes.

VI. Discussion, Disclaimers, and the Way Forward

We now coalesce our contributions, and derive an agenda that can be used to counter the problem of AIA in videogames.

Our Findings revealed a snapshot of the current gaming landscape from the viewpoint of privacy. Our results (in §III) serve as a guide for game-studies. A researcher can focus on games which bear low-risk (e.g., Hearthstone, or [45]) of AIA, and finding correlations there would not help any (potential) attacker. A researcher can also focus on AIA-prone games: in which case, caution should be taken, so that participants (and downstream research) are adequately informed of the risks. Although we do not claim generalizability of our results, our survey evidences that our participants are not very knowledgeable about privacy-related issues (§V-C). Plus, our structured approach (§IV-B) shows how to ethically study AIA.

Our Goal is to inspire a *reflective exercise* by the entire gaming landscape: players, developers and researchers. Importantly: this is *not a finger-pointing attempt*. The authors of prior work which discovered relationships between in-/off-game data (e.g. [9] in 2019) were acting in good faith, and *we are glad* that their contributions advanced our body of knowledge. Similarly, players/communities that are open to participating in research studies are *beneficial for science*. Indeed, before the publication of [7] (Apr. 2023) there was (relatively) little reason to be worried about AIA in videogames. Hence, it was acceptable to avoid stressing privacy-related issues in this context. However, as we demonstrated, AIA are a concrete risk which affects many games (§III) and communities (§IV) and which have been overlooked by abundant related work (§II-B). The situation is aggravated by *(i)* real attackers who are much more persistent than us, and can incentivize participation in surveys with economical rewards; *(ii)* the possibility of sharing/renting AIA-ready models; and *(iii)* the fact that many players are not adequately informed (§V-C). In light of the above, we believe that the current gaming landscape necessitates an overhaul from the perspective of data-privacy.

Our Vision. We want to ensure that AIA in videogames are properly accounted for. We hence *extend* the recommendations by Tricomi et al. [7], who only mentioned "players" and "developers", by adding "researchers" to the army that can fight off AIA. We advocate for the following (Fig. 4):

- Researchers should *promote a privacy-oriented mindset* when conducting their studies, educating players and informing developers of their findings.
- Developers should *listen to researchers*, relaying and amplifying their discoveries to the players, and proactively *introduce more privacy-related options* in their products (games, but also tracking websites).
- Players should *pay attention* to the evolution of privacy-related issues, voice their skepticism/concerns when they suspect potential privacy threats, and demand more privacy-preserving options and educational initiatives.

Finally, *security researchers can help*, too—but they may not be aware of problems in gaming.[a] We encourage more cooperation between gaming and security research.

[a] During some discussions on the Distinguished Paper Award winner [46] at USENIX SEC23 (a top-venue in security), some security researchers commented that "I did not know *aimbots* were a security problem!"

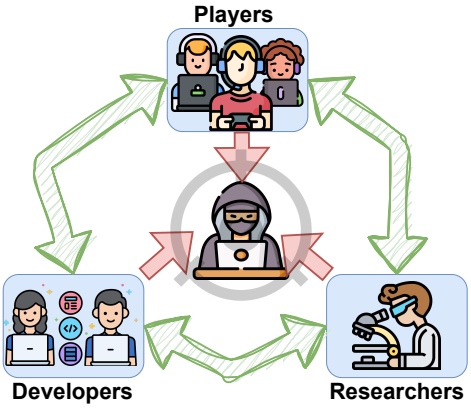

Fig. 4: **Our Vision to counter AIA in Videogames** – AIA can only be mitigated with a collective effort. *Researchers* should inform players and developers of novel privacy threats. *Developers* should account for the findings of research, and devise privacy-friendly initiatives. *Players* should be aware of privacy risks, and engage in social activities with a privacy-oriented mindset.

CONCLUSIONS. This paper is a call to action. We have provided further evidence of the threats of AIA in videogames. AIA are a subtle (and overlooked) threat that can target even privacy-aware players, as well as minorities (e.g., children). Every stakeholder (gamers, developers, researchers) should be more mindful of privacy, and collectively contribute to raising the overall awareness of privacy-related issues in videogames. We hope our paper inspires such a change.[a]

[a]We have created a repository [40] including: our questionnaire (§V-A); a document explaining our literature (§II-B) and community (§IV-B) analyses; additional evidence (mails/slides/links) to support some claims.

Educational Campaign. We disseminated our findings and takeaways among the communities wherein we shared our survey, thereby alerting (and preparing) them about AIA.

REFERENCES

[1] J. L. Kröger, P. Raschke, J. Percy Campbell, and S. Ullrich, "Surveilling the gamers: Privacy impacts of the video game industry," *Entertainment Computing*, 2023.
[2] https://statista.com/outlook/dmo/digital-media/video-games/worldwide.
[3] https://newzoo.com/resources/trend-reports/newzoo-global-games-market-report-2021-free-version.
[4] theesa.com/resource/2022-essential-facts-about-the-video-game-industry.
[5] D. Staat, G. Wallner, and R. Bernhaupt, "Towards a community-based ranking system of overwatch players," in *Int. Conf. Entert. Comp.*, 2022.
[6] E. Kleinman and M. S. El-Nasr, "Using data to" git gud": a push for a player-centric approach to the use of data in esports," in *CHI*, 2021.
[7] P. P. Tricomi, L. Facciolo, G. Apruzzese, and M. Conti, "Attribute inference attacks in online multiplayer video games: A case study on dota2," in *ACM CODASPY*, 2023.
[8] D. Martinovic, V. Ralevich, J. McDougall, and M. Perklin, ""you are what you play": Breaching privacy and identifying users in online gaming," in *IEEE PST*, 2014.
[9] Z. Wang, A. Sapienza, A. Culotta, and E. Ferrara, "Personality and behavior in role-based online games," in *IEEE CoG*, 2019.
[10] R. Sifa, A. Drachen, and C. Bauckhage, "Profiling in games: Understanding behavior from telemetry," in *Social interactions in virtual worlds: An interdisciplinary perspective*, 2018.
[11] N. Z. Gong and B. Liu, "Attribute inference attacks in online social networks," *ACM TOPS*, 2018.
[12] ——, "You are who you know and how you behave: Attribute inference attacks via users' social friends and behaviors," in *USENIX Sec.*, 2016.
[13] G. Apruzzese, H. S. Anderson, S. Dambra, D. Freeman, F. Pierazzi, and K. Roundy, ""Real attackers don't compute gradients": Bridging the gap between adversarial ml research and practice," in *IEEE SaTML*, 2023.
[14] https://docdroid.net/ZeJTLar/rdota2-demographics-report-2021-pdf.
[15] D. Kao, R. Ratan, C. Mousas, A. Joshi, and E. F. Melcer, "Audio matters too: How audial avatar customization enhances visual avatar customization," in *CHI*, 2022.
[16] A. Eidelberg, C. Jacob, and J. Denzinger, "Using active probing by a game management ai to faster classify players," in *IEEE CoG*, 2019.
[17] A. Canossa, D. Salimov, A. Azadvar, C. Harteveld, and G. Yannakakis, "For honor, for toxicity: Detecting toxic behavior through gameplay," in *CHI PLAY*, 2021.
[18] R. Kowert and C. Cook, "The toxicity of our (virtual) cities: prevalence of dark participation in games and perceived effectiveness of reporting tools," in *HICSS*, 2022.
[19] G. Johnson, J. Runge, and E. Seufert, "Privacy-centric digital advertising: Implications for research," *Customer Needs and Solutions*, 2022.
[20] P. C. Ferreira, A. M. V. Simão, A. Paiva, C. Martinho, R. Prada, A. Ferreira, and F. Santos, "Exploring empathy in cyberbullying with serious games," *Computers & Education*, 2021.
[21] D. Antognoli and J. Fisher, "The purposes and meanings of video game bathrooms," in *IEEE CoG*, 2021.
[22] J. T. Bowey, J. Frommel, B. Piller, and R. L. Mandryk, "Predicting beliefs from npc dialogues," in *IEEE CoG*, 2021.
[23] J. Jia and N. Z. Gong, "Attriguard: A practical defense against attribute inference attacks via adversarial machine learning," in *SEC*, 2018.
[24] S. Tekofsky, P. Spronck, A. Plaat, J. Van den Herik, and J. Broersen, "Psyops: Personality assessment through gaming behavior," in *International Conference on the Foundations of Digital Games*, 2013.
[25] T. Kennedy *et al.*, "Predicting mmo player gender from in-game attributes using machine learning models," in *Predicting real world behaviors from virtual world data*, 2014.
[26] P. Likarish, O. Brdiczka, N. Yee, N. Ducheneaut, and L. Nelson, "Demographic profiling from mmog gameplay," in *PETS*, 2011.
[27] A. C. Tally, Y. R. Kim, K. Boustani, and C. Nippert-Eng, "Protect and project: Names, privacy, and the boundary negotiations of online video game players," *Proc. ACM Human-Comp. Inter.*, 2021.
[28] https://netenrich.com/blog/fraudgpt-the-villain-avatar-of-chatgpt.
[29] T. Kohno, Y. Acar, and W. Loh, "Ethical frameworks and computer security trolley problems: Foundations for conversations," *SEC*, 2023.
[30] M. Bailey, D. Dittrich, E. Kenneally, and D. Maughan, "The Menlo report," *IEEE Security & Privacy*, 2012.
[31] https://bleepingcomputer.com/news/security/hackers-try-to-extort-survey-firm-questionpro-after-alleged-data-theft/.
[32] Z. Halim, M. Atif, A. Rashid, and C. A. Edwin, "Profiling players using real-world datasets: clustering the data and correlating the results with the big-five personality traits," *IEEE TAC*, 2017.
[33] S. Lesmana, O. Ariwana, R. P. Halim, and A. A. Gunawan, "Behavior correlation between games in first-person shooter genre based on personality traits," *Procedia Computer Science*, 2021.
[34] D. L. King, A. M. Russell, P. H. Delfabbro, and D. Polisena, "Fortnite microtransaction spending was associated with peers' purchasing behaviors but not gaming disorder symptoms," *Addictive Behaviors*, 2020.
[35] S. M. F. Gillani, "Evaluation of games monetization approaches: A case study on players unknown's battlegrounds (PUBG)," MSc. Thesis, 2021.
[36] T. Ide and H. Hosobe, "Supporting online game players by the visualization of personalities and skills based on in-game statistics." in *VISIGRAPP*, 2023.
[37] M. Krcmar, R. McGloin, K. M. Farrar, and S. Scott Li, "'what is my call of duty?': Exploring the importance of player experience in a first-person shooter video game," *Journal of Gaming & Virtual Worlds*, 2018.
[38] M. Schaekermann *et al.*, "Curiously motivated: profiling curiosity with self-reports and behaviour metrics in the game "Destiny"," in *CHI Play*, 2017.
[39] R. Heartfield and G. Loukas, "A taxonomy of attacks and a survey of defence mechanisms for semantic social engineering attacks," *ACM CSIR*, 2015.
[40] Repository of our paper: https://github.com/hihey54/cog24_aia, 2024.
[41] steamcommunity.com/app/1517290/discussions/0/3191360735183020391/.
[42] steamcommunity.com/app/730/discussions/0/2592234299558984436/.
[43] steamcommunity.com/app/570/discussions/0/3818529263636669669.
[44] steamcommunity.com/app/730/discussions/0/6993585599474786899/.
[45] A. J. Bisberg, J. Jiang, Y. Zeng, E. Chen, and E. Ferrara, "The gift that keeps on giving: Generosity is contagious in multiplayer online games," *CHI*, 2022.
[46] M. Choi, G. Ko, and S. K. Cha, "Botscreen: Trust everybody, but cut the aimbots yourself," in *USENIX SEC*, 2023.