

Quando as Avaliações Viram Bombas: Explorando a Dinâmica do Review Bombing nos Jogos no Metacritic

Marcus Vinicius Guerra Ribeiro
marcus.guerra@ufv.br
Universidade Federal de Viçosa - UFV
Florestal, Minas Gerais

Clara Andrade Pimentel
clara.pimentel@cefetmg.br
CEFET-MG
Belo Horizonte, Minas Gerais

Philippe de Freitas Melo
philipe.freitas@ufv.br
Universidade Federal de Viçosa - UFV
Florestal, Minas Gerais

ABSTRACT

This study investigates the phenomenon of “Review bombing” on Metacritic, a platform for aggregating media reviews. Review bombs involve the deliberate posting of a large volume of negative reviews to alter the reputation of a media work, often driven by political, social, or creator-related disagreements. We collected extensive data from Metacritic, comprising 145K games and 1M reviews to understand and identify the characteristics of review bombing online. Using clustering techniques and analyzing various user and game-related metrics, we found that targeted games experience a significant increase in review volume, often with similar content and slightly higher toxicity, reflecting coordinated dissatisfaction campaigns. Additionally, the clustering achieved an F1-score of 0.5263, revealing particular challenges to categorizing this kind of attacks, such as the removal of evidence by platform moderation and difficulty in labeling minor attacks. The study highlights the need for robust mechanisms to detect and mitigate review bombs to preserve the integrity of review platforms and consumer trust.

KEYWORDS

jogos, videogames, Metacritic, review bombing, clusterização, discurso de ódio, toxicidade

1 INTRODUÇÃO

Atualmente, plataformas digitais desempenham um papel crucial na maneira como as pessoas se relacionam, consomem e interpretam conteúdo. Neste sentido, *websites* que reúnem revisões e avaliações de produtos e mídias, como Rotten Tomatoes, Metacritic e IMDB¹ se tornaram referências essenciais no universo do entretenimento [32]. Esses sites fornecem espaços para que indivíduos expressem suas avaliações sobre diversas mídias, estabelecendo-se como ferramentas indispensáveis para consumidores em busca de orientação sobre produtos culturais. Conseqüentemente, um número crescente de pessoas consulta essas plataformas para verificar críticas, comentários e notas antes de investir seu tempo e dinheiro em filmes, séries, jogos ou outras formas de entretenimento [29].

Entretanto, a ascensão destas plataformas não vem sem desafios. Pontuações e opiniões em plataformas agregadoras de críticas podem influenciar diretamente as vendas de jogos e outras mídias [12].

¹Available at: <https://www.rottentomatoes.com/>, <https://www.metacritic.com/> and <https://www.imdb.com/>

Um dos principais problemas derivados disso é o fenômeno de *review bomb* (RB), prática que tem se tornado uma preocupação crescente na indústria de jogos. *review bombing* refere-se ao ataque de usuários de inundar plataformas de avaliação com críticas negativas de forma coordenada, geralmente em resposta a alguma controvérsia relacionada ao jogo, desenvolvedores ou empresas. Essa prática vai além da avaliação da qualidade da obra e objetiva ativamente distorcer a percepção pública sobre um produto e alterar a reputação da obra [16], muitas vezes motivados por questões políticas, sociais ou de desavenças com criadores de conteúdo. O *review bombing* também prejudica a credibilidade das plataformas de avaliação e pode ter impactos financeiros significativos para os desenvolvedores.

Identificar esse fenômeno, entretanto, não é uma tarefa simples, dada a complexidade em distinguir entre críticas legítimas daquelas criadas para sabotar a reputação do produto [22]. A facilidade com que as pessoas podem criar múltiplas contas e mobilizar campanhas coordenadas através de fóruns como *Reddit*, *Chans* e outros, aumenta ainda mais o impacto desses ataques enquanto torna ainda mais difícil detectar e banir usuários.

Tendo em vista detectar e classificar esse fenômeno, assim como compreender as características dos usuários do Metacritic, esse artigo investiga as seguintes perguntas de pesquisa: **RQ1**: Quais são as características de um jogo que sofre *review bombing* online? **RQ2**: Como se comportam os usuários que atacam a plataforma em suas interações? **RQ3**: Quais as principais motivações dos comentários desses ataques?

Nesse contexto, o objetivo deste estudo é compreender as características campanhas de *review bombing* saber como identificar esse fenômeno em plataformas online. Para tal, fizemos uma coleta em larga escala na plataforma Metacritic de jogos e seus comentários/avaliações. Abrangemos em nossa coleta desde informações gerais sobre as obras como título e data de lançamento, até informações mais específicas sobre os comentários, seus autores e suas notas, resultando em um *dataset* com mais de 145k jogos e 1 milhão de comentários. A partir desses dados, filtramos 12.200 jogos distintos relevantes para o estudo, selecionando apenas os jogos que possuem ao menos 1 comentário e 1 nota, e rotulamos manualmente títulos de jogos de destaque que sofreram ataques de RB entre aqueles coletados. Em seguida, fizemos uma análise aprofundada sobre este fenômeno, levando em consideração fatores como discurso de ódio e toxicidade dos comentários. Com isso utilizamos de métodos de clusterização para agrupar as diferentes categorias de jogos, identificando particularidades distintas dos jogos que sofreram algum ataque. Por fim, aprofundamos no conteúdo dos comentários destes jogos para discutir quais os principais tópicos presentes no

ato do *review bombing*, assim como para identificar as principais motivações por trás destes ataques.

Como resultado, pudemos observar que jogos que sofrem de *RB* possuem características únicas, além das notas baixas, que nos permite diferenciar dos demais jogos na plataforma, como o volume gigantesco de comentários, maior engajamento dos usuários em comentários negativos, comentários mais tóxicos e com uso de termos específicos. Também encontramos nos comentários ataques direcionados à empresa e aos desenvolvedores, um apelo ao uso de termos técnicos para dar maior validade às críticas, além de motivações carregadas de discursos de ódio, como sexismo e homofobia.

2 TRABALHOS RELACIONADOS

Avaliar e identificar *review bombing* é uma tarefa complexa. Os usuários almejam criar a impressão de que os comentários e notas negativos são orgânicos, e não recorrem diretamente ao discurso de ódio ou a comentários diretos sobre o ataque, porque assim as plataformas conseguiriam banir com facilidade o conteúdo que viola os termos de uso. Desse modo, os usuários manipulam as ferramentas disponíveis nas plataformas digitais de maneira coordenada, como as notas em sites de críticas ou botões de *dislike*, como os do YouTube [19] [4].

Observando as relações entre o número de vendas e as notas atribuídas às obras, é possível notar que as avaliações de produtos ou mídias exercem uma influência direta e significativa sobre o comportamento dos consumidores, influenciando diretamente o volume de vendas [4] [2]. Especialmente no contexto das mídias eletrônicas, conforme demonstrado por [9], parece existir uma relação geométrica entre o número de vendas desses jogos e suas notas no Metacritic, destacando que pontuações mais altas promovem um aumento de vendas. Também existem indícios que destacam as limitações das pontuações atribuídas no site como representações completas e precisas das experiências dos jogadores [13]. O estudo de [28] traz um interessante contraste entre as avaliações feitas por amadores e por profissionais no Metacritic, onde se demonstra que amadores tendem a fornecer avaliações mais polarizadas e favoráveis a obras antigas, mesmo as avaliando surpreendentemente muito tempo após o lançamento. Em contraste, especialistas geralmente emitem avaliações menos polarizadas, concentrando-se em aspectos técnicos e contextuais dos jogos e tendem a avaliá-los logo após o lançamento.

Os resultados obtidos por [31] revelam a ineficácia da maioria dos classificadores de sentimento na avaliação precisa de resenhas de jogos. Uma análise mais detalhada apontou duas causas principais para as classificações incorretas: a presença de resenhas que destacam tanto os pontos positivos quanto os negativos, bem como resenhas que fazem comparações entre diferentes obras, o que pode confundir o classificador.

Diferente dos trabalhos apresentados, este se distingue, pois utiliza técnicas de clusterização e busca caracterizar padrões de *review bombing*. Assim como também foi realizada uma análise de como os usuários interagem e comentam com jogos que sofrem ataques, e a relação deles com toxicidade e discurso de ódio.

3 METODOLOGIA

O Metacritic reúne críticas de uma variedade de mídias de entretenimento, incluindo filmes, programas de televisão, álbuns de música

e jogos. Na plataforma, os usuários podem fornecer uma pontuação com ou sem comentário para mídias específicas. No sistema de avaliação, os avaliadores são categorizados como crítica especializada ou usuários amadores. Os críticos fazem parte de uma seleção cuidadosa e sua média ponderada resulta em um valor numérico singular, chamada de *Metascore*. Enquanto isso, as pontuações dos usuários amadores contribuem para o *UserScore*. Para a análise dos ataques de *review bombing* na plataforma, dada a ausência de uma API própria, foi necessário o desenvolvimento de um coletor e a criação de um novo conjunto de dados foi considerada necessária para o estudo.

3.1 Coleta de Dados

Para obter os dados, foi implementado um *web scraper*, conforme mostrado na Figura 1, que manipula o navegador através do uso do Selenium ² e coleta os dados nas páginas do Metacritic. O primeiro passo envolve a navegação pelo catálogo da plataforma. Este catálogo contém uma lista completa dos jogos disponíveis no site. A navegação é feita percorrendo todas as 6068 páginas existentes, manipulando a URL para iterar por todas as páginas.

A cada página visitada, realiza-se uma extração do HTML presente. Durante essa etapa, todos os títulos dos jogos são coletados e armazenados. O objetivo dessa etapa é compilar uma lista de todos os jogos disponíveis. Uma vez obtidos todos os títulos, o próximo passo é acessar cada URL individualmente. Isso é realizado concatenando o título do jogo com a URL base do Metacritic, formando uma URL no formato `https://www.metacritic.com/game/gameName`, na qual são extraídas todas as informações desejadas. Dessa maneira são coletadas as seguintes informações sobre os jogos: (1)plataforma, (2)classificação etária, (3)título, (4)editora, (5)desenvolvedor, (6)data de lançamento, (7)gênero, (8)notas, (9)quantidade de notas positivas, negativas e neutras. Quanto aos usuários, tanto críticos quanto amadores, foram coletados seus (i) comentários, (ii) notas, (iii) nomes, (iv) data dos comentários, (v) plataforma da qual o comentário foi feito e o (vi) tipo de usuário. Cada jogo é caracterizado por 27 atributos distintos, abrangendo desde informações sobre usuários e seus comentários até detalhes intrínsecos das mídias, como data de lançamento e desenvolvedores. Além disso, implementamos outras métricas como a variância entre as notas dentro de um jogo, a relação entre o *UserScore/Metascore*, contagem de palavras-chave relacionadas ao contexto de discurso de ódio, totalizando mais 8 atributos derivados.

Como resultado, coletamos **145.523 jogos** e **1.017.504 comentários** distintos. Para garantir a qualidade dos dados, foi elaborada uma filtragem, onde foram removidas todos os jogos que não tinham nenhum comentário, ou que não possuísem *UserScore* ou *Metascore*, ou que faltasse algum dado relevante para análise. Com isso, o conjunto de dados final contém **12.200 jogos** propriamente avaliados no Metacritic e os mesmos **1.017.504 comentários** distintos.

3.2 Rotulando Jogos alvos de *Review Bombing*

Dado nosso objetivo de estudo de identificar ataques de *RB*, também precisamos estabelecer um conjunto bem definido de jogos que foram alvos de ataques. Para levantar os jogos alvos de *RB*,

²Available at: <https://selenium-python.readthedocs.io/>

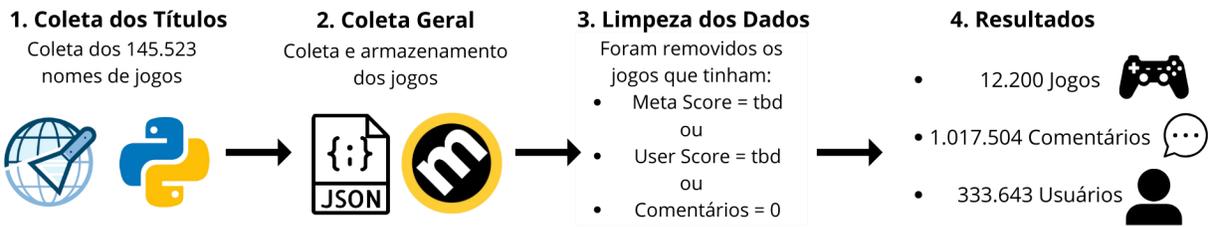


Figura 1: Diagrama geral da coleta dos dados dos jogos do Metacritic.

rotulamos manualmente os jogos baseado em fontes externas. Para isso, foram consultados três sites distintos e bem citados na comunidade *gamer*, que produziram matérias sobre casos notórios de *review bombing* [25, 33, 34]. A seleção dos candidatos foi realizada com o seguinte critério nessas fontes: um jogo é marcado como *review bombing* caso alguma fonte tenha explicitamente mencionado que o jogo foi alvo de um ataque no Metacritic; ou caso mais de uma única fonte tenha simplesmente citado o jogo. Além disso, também foram incluídos manualmente outros casos de destaque na comunidade como: *Horizon Forbidden West* [27], *Battlefield V* [18], e *Diablo Immortal* [5].

4 ANÁLISE DOS DADOS

Após limpar e analisar o *dataset*, partimos para uma análise mais aprofundada dos dados. Nesta seção, utilizamos técnicas de ciência de dados, como derivação de *features*, utilização de modelos de aprendizado de máquina e clusterização para podermos compreender quais características definem um ataque de *review bombing*. Nesta seção, portanto, analisamos como as características dos jogos no Metacritic (e.g., comentários, curtidas, toxicidade, sentimento) se comportam quando os jogos sofrem ataques, comparadas àquelas dos que não sofreram. Dessa forma, conseguimos mensurar as atividades relacionadas à prática de *review bombing* online e discutir as motivações e implicações de tais ataques para as comunidades online.

4.1 Engajamento no Metacritic

Ao analisar os dados, é possível perceber como a seção de jogos do Metacritic é composta. Onde apenas 8.3% dos jogos tem tanto um comentário, um *userScore* e um *Metascore*. Além disso, conforme a Figura 2a, mais de 60% dos jogos possuem menos de 10 avaliações. Isso evidencia que uma parte significativa dessa seção é vazia, com inúmeros jogos pouco avaliados, ou com um engajamento muito baixo. Também é importante ressaltar o limite superior do gráfico de

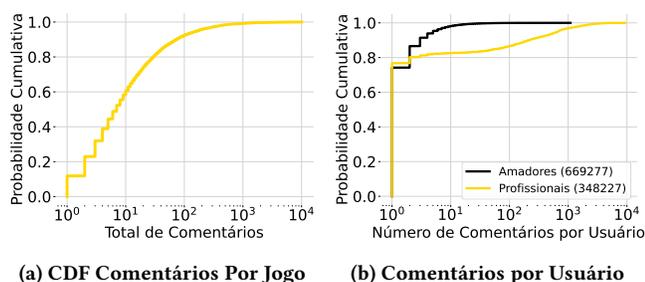


Figura 2: Estrutura geral dos dados do Metacritic

10.000 comentários. Esse limite é imposto pela própria plataforma como o número máximo de comentários visíveis por jogo, mas demonstra que alguns jogos atraem uma grande atenção do público.

Em relação aos autores de comentários, os dados mostram a presença de 333.624 usuários únicos. Destes, 99,08% são classificados como amadores e apenas 0,92% são críticos especializados. Dentro dessa minoria, aproximadamente 23% deles produziram mais de um comentários. Contudo, essa pequena parcela de profissionais é responsável por 34,22% de todos os comentários produzidos no site, indicando um nicho de altíssima produtividade. com alguns usuários escrevendo quase 10.000 comentários, como visto na CDF da Figura 2b. Já os amadores, mais de 98% possuem até 10 comentários, demonstrando que a maioria deles não participam tão ativamente na plataforma do Metacritic, apenas opinando em jogos de seu interesse. Entretanto, também temos usuários altamente ativos, com alguns chegando a perto de 100 jogos avaliados.

4.2 Análise de Toxicidade

Além das informações já obtidas, foi também processada a toxicidade de cada comentário, através da ferramenta do Perspective API³. Esse modelo analisa textos e identifica níveis de toxicidade numa pontuação que varia de 0 a 1. Baseando-se em outros estudos [17], nós consideramos que um comentário é determinado tóxico, se ele está acima do limiar de 0.75. Além disso, mantivemos o idioma original de cada comentário, uma vez que o modelo mostra bom desempenho no cenário com múltiplos idiomas [14]. Assim, processamos todos os comentários dos usuários amadores (669.277) a fim de mensurar a toxicidade dos jogos coletados.

Nossa primeira análise focou nos principais plataformas no qual os jogos são publicados, PlayStation, Nintendo, PC e Xbox (Fig. 3a). É possível observar que, as curvas apresentadas são notavelmente similares, indicando que a plataforma não ser um fator determinando na toxicidade. Mesmo que existam rivalidades nas comunidades fãs de cada plataforma e jogos exclusivos, tais características não refletem, de forma geral, em como os jogos são recebidos no Metacritic. Outra característica interessante presente nos dados é o gênero de cada jogo avaliado. A Figura 3b contabiliza a quantidade de comentários tóxicos (> 0.75) por gênero. Observa-se que “*Esportes*” possui relativamente uma taxa de toxicidade muito maior que as demais categorias (1.1%), muito superior a “*Simulação*”, que vem em segundo lugar com 0.5%. Essa alta pode ser causada pela natureza competitiva desse gênero, em que frustração dos jogadores podem se refletir em críticas mais contundentes ao jogo em questão. Em contrapartida, gêneros como, “*Estratégia*”, “*Aventura*” e “*RPG*” apresentam toxicidade menor, próximo de 0.2%. Essas mídias tendem a ser mais focados em resolução de problemas, narrativa ou

³<https://perspectiveapi.com/>

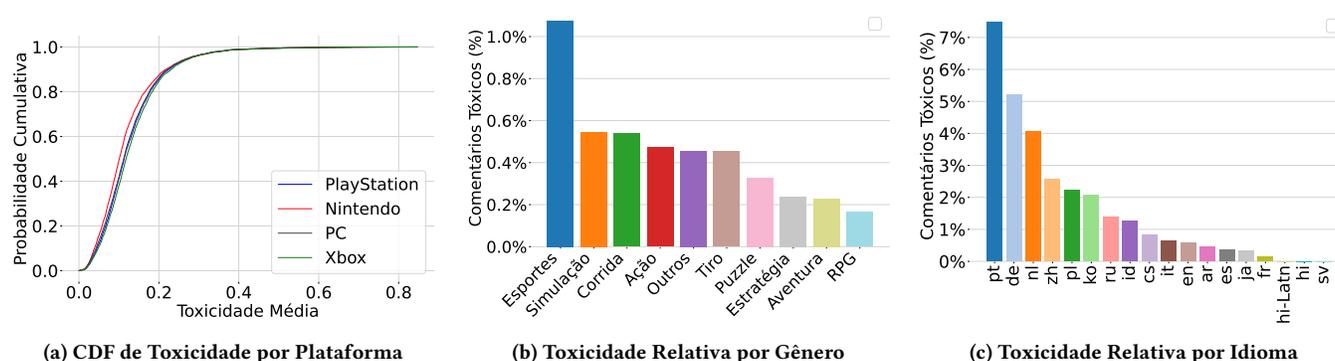


Figura 3: Análises de toxicidade e idioma dos comentários dos jogos do Metacritic.

cooperação, podendo assim desmotivar sentimentos negativos. Por fim, o gráfico 3c mostra a distribuição percentual relativa de comentários tóxicos por diferentes idiomas. Notavelmente, português têm a maior taxa de toxicidade, atingindo quase 8%. Essa diferença pode ser um reflexo das normas de comunicação do idioma, onde uma linguagem considerada tóxica por um povo, pode ser considerada normal para outro.

5 CARACTERIZANDO REVIEW BOMBING

Nossa primeira abordagem para compreender a natureza dos ataques de *review bombing* nos dados foi usar das *features* extraídas para agrupar os jogos e encontrar possíveis padrões. Para isso, realizamos diversos experimentos com diferentes algoritmos de clusterização. Este processo foi desenvolvido em várias etapas, com o objetivo de separar mais claramente as mídias que sofreram ataques de *review bomb* daquelas que não foram afetadas.

Dada a definição do *review bombing* que envolve as revisões e comentários publicados pelos usuários na plataforma, nesta etapa foi realizada uma filtragem dos jogos do *dataset* com apenas aqueles com pelo menos 50 comentários, possibilitando assim, uma análise mais aprofundada e com maior explicabilidade do conteúdo dos comentários em cada jogo, levando ao *dataset* final de 1645 jogos distintos. Esse valor foi alcançado baseado nos resultados apontados na Figura 2a, em que a média de comentários por jogos é de 47.

Com o conjunto de dados final em mãos, a próxima etapa foi definir quais *features*, dentre aquelas coletadas e derivadas descritas na Seção 3, serão utilizadas no algoritmo de clusterização, visando uma separação mais clara entre os jogos que sofreram ataque daqueles que não sofreram.

Tendo em vista limitar ainda mais esse conjunto, aprimorando assim ainda mais a precisão dos *clusters* e reduzindo processamento desnecessário, optou-se por utilizar o algoritmo do Mutual Information (MI) [15] do sklearn para identificar quais parâmetros são mais relevantes para conseguir discernir um alvo, no caso em questão, jogos que sofreram *review bomb*. Ao aplicá-lo, foi possível selecionar os atributos que apresentam maior relevância e capacidade explicativa no contexto em questão. Assim, foram escolhidos os seguintes parâmetros, que serão utilizados na clusterização: “totalComments” o número total de comentários feitos em um jogo, “meanToxicity”

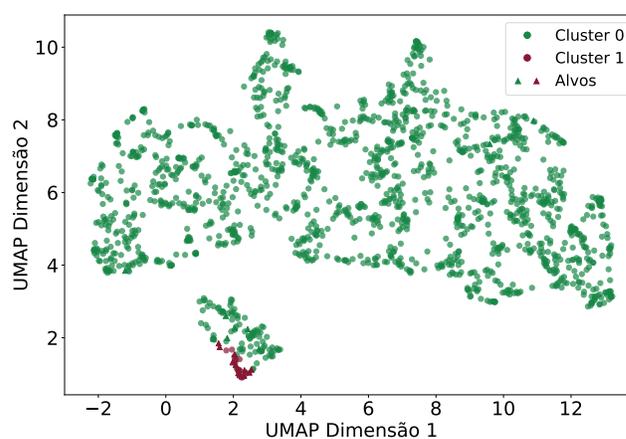


Figura 4: Clusterização de jogos.

a toxicidade média dos comentários, “totalAlike” o número total de comentários semelhantes (avaliados utilizando a biblioteca *fuzzywuzzy*⁴ com um *threshold* de 0.8), “userRtotal” o conjunto total de notas atribuídas por usuários amadores, “totalCurse” e o número total de comentários que contêm palavrões (considerando como palavrão qualquer comentário que contenha **** dentro dele). No qual valores de toxicidade, similaridade e palavrões são expressos em percentuais.

Para os algoritmos de clusterização, foram considerados o K-means, HDBSCAN, DBSCAN, Gaussian Mixture. Foram realizados testes de desempenho em todos, avaliando a capacidade de cada um em separar os jogos normais daqueles que sofreram *RB*. Como resultado, calculamos a métrica de F1-score, considerando o *cluster* com a maior proporção de jogos que sofreram ataques, como o de *review bomb*, e o outro, como sendo livre de ataque. Para casos com mais de 2 *clusters*, foi escolhida a configuração que resultaria no maior F1. Após a realização dos testes, foi observado que, o k-means com $k = 2$ apresentou os resultados mais satisfatórios.

Além disso, também foi aplicado um método de redução de dimensionalidade com o UMAP, para melhorar a visualização dos *clusters*. Os resultados são apresentados na Figura 4. Nela, todos os itens pertencentes ao *cluster 1* (representado pela cor vermelho)

⁴Available at: <https://pypi.org/project/fuzzywuzzy/>

Predição\Real	Sim	Não	Total
Sim	15	11	26
Não	16	1603	1619
Total	31	1614	1645

Tabela 1: Matriz de confusão da clusterização

foram classificados como alvos de ataques, enquanto os itens do *cluster 0* (representado pela cor verde) foram classificados como livres de ataques.

5.1 Matriz de Confusão

Na matriz de confusão, exibida na Tabela 1, observa-se um desempenho medido pelo F1-score de 0.5263. Este resultado, apesar de modesto à primeira vista, merece uma discussão detalhada sobre as peculiaridades de cada mídia e revela características e desafios interessantes do problema analisado.

Focando inicialmente nos falsos negativos, ou seja, os jogos que sofreram *Review bombing*, mas que não foram identificados pelo modelo de detecção. Alguns exemplos notáveis incluem *Bastion* e *Toy Soldiers*, cujos comentários negativos foram removidos pelo próprio site do Metacritic [24], dificultando a identificação desses casos pelo algoritmo. Além disso, existem situações como a de *Borderlands 3* [33], no qual a onda de comentários ocorreu na plataforma da Steam, diferente do Metacritic, levando a uma classificação incorreta do jogo. Examinando também os falsos positivos, isto é, os jogos que foram classificados como alvos de *Review Bombing*, embora não estivessem presentes na nossa rotulação inicial. Como é o exemplo de *FIFA 20* [6] e *Baldur's Gate 3*, que não entrou na nossa rotulação inicial, mas recentemente também foi alvo de ataques de *Review Bombing* [3], porém, como esses ataques tiveram uma menor repercussão ou foram mais recentes, acabaram não incluídos na rotulação original.

Portanto, vemos que a moderação ativa do Metacritic e variedade de outras plataformas de avaliação podem apresentar um obstáculo grande a análise do *review bombing*, uma vez que elas podem ocultar as evidências dos ataques que esses jogos sofreram. Entretanto, observando tanto a figura quanto os resultados, vemos que grande parte dos jogos no *cluster* de *Review bombing* Compartilham características semelhantes que sugerem que esse fenômeno pode ser identificado e, portanto, combatido dentro das plataformas.

5.2 Atributos do Cluster Review Bombing

Seguindo os resultados da clusterização, podemos observar os centroides de cada *cluster* na Figura 5 para entender melhor as características individuais dos jogos normais daqueles que sofreram ataque.

Primeiramente, ao examinar o número total de comentários (*totalComments*), tanto os valores absolutos dos centroides quanto a distribuição mostrada no boxplot indicam uma diferença substancial. O número médio de comentários para os jogos afetados é quase 19 vezes maior em comparação com os jogos não afetados. Sugerindo assim, um aumento massivo de comentários em jogos que sofreram *review bomb*. Outro atributo relevante é o *userRtotal*, que indica a quantidade total de notas atribuídas a um jogo. A quantidade de notas que um jogo que sofreu ataque recebe também é significativamente maior, quase 17 vezes mais. Reforçando ainda

mais a ideia de que jogos atacados, recebem um aumento massivo no número de comentários.

Abordando também a média de toxicidade (*meanToxicity*), e o total de comentários contendo palavrão (*totalCurse*), tem-se uma pequena diferença no valor de um centroide para o outro, embora os *Review Bomb* sejam um pouco mais elevado. Ao comparar ambos, é possível notar uma variação muito menor dos quartis de um *cluster* para outro, demonstrando que os jogos atacados são mais consistentes. Um fator que pode explicar a semelhança entre os dois grupos é o sistema de moderação do Metacritic que apaga comentários muito tóxicos ou não-apropriados. Ainda assim, ambos atributos se mostraram relevantes na diferenciação dos jogos.

A métrica *totalAlike*, que representa quantos comentários são similares entre si, traz um debate interessante. A grande diferença nos valores sugere que os comentários dos ataques são geralmente compostos por um argumento único ou repetitivo dentro da comunidade, impulsionando a onda de comentários negativos, enquanto jogos normais possuem comentários mais individualizados. Um exemplo notável é o jogo *Madden 21*, em que os fãs reclamaram da falta de mudanças significativas de uma versão para outra, resultando em uma avalanche de comentários com textos semelhantes [33].

A definição de um *review bomb* como: “a inundação de um site com avaliações negativas (seja de um produto ou negócio)” [20] se alinha com os resultados encontrados. Conforme discutido anteriormente, uma mídia que sofreu um ataque tende a apresentar um aumento massivo no número de comentários, predominantemente negativas. No entanto, é possível expandir essa definição para incluir também a similaridade entre os comentários e uma leve elevação na toxicidade. Essas outras características refletem o fato do *Review bombing* ser também uma campanha coordenada e de insatisfação, gerando esse efeito negativo e homogêneo nos comentários.

6 ANÁLISE DE CONTEÚDO DOS COMENTÁRIOS

Nesta seção voltamos nossas análises para o conteúdo dos comentários dos jogos que sofreram *review bombing* no Metacritic. Utilizando-se de nuvens de palavra, buscamos observar tendências gerais e tópicos pertinentes que servem de motivação para esses ataques. Também foi realizada uma análise de discurso de ódio presente nos comentários, relacionando com uma lista de palavras-chave criada a partir do contexto de ódio presente nos jogos.

6.1 Word Clouds

As *Word Clouds* são ferramentas analíticas úteis que permitem uma visualização abrangente de tópicos presentes em grandes volumes de texto. No contexto deste estudo, foram geradas nuvens de palavras a partir de todos os jogos que sofreram *review bomb*, com o objetivo de identificar críticas recorrentes e tópicos emergentes. A análise dessas nuvens de palavras revelou a existência de três tópicos comuns, que se repetiram em diversos jogos, fornecendo uma visão clara das principais áreas de insatisfação dos usuários que praticam *review bombing*. O primeiro tópico identificado envolve críticas diretas às empresas responsáveis pelo desenvolvimento e

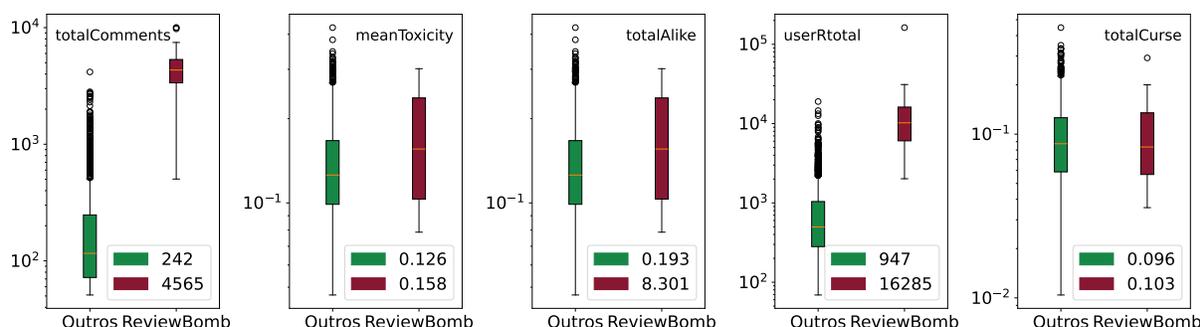


Figura 5: Comparação de atributos de jogos com *Review Bombing* e demais jogos, sendo o verde o centroide do *cluster* de jogos normais e o vermelho o *cluster* dos jogos que sofreram *review bombing*.

lançamento dos jogos. Nesses casos, jogadores frequentemente expressam insatisfação com decisões corporativas que, segundo eles, afetam negativamente a qualidade dos jogos. Dois exemplos emblemáticos dessa situação são os casos de *Grand Theft Auto: The Trilogy – The Definitive Edition* [25] e *Diablo IV* [26]. Conforme visto nas Figuras 6c e 6f, os nomes das empresas (Rockstar e Blizzard) aparecem em destaque nas nuvens de palavras, junto a palavras como "bug", "money", "release", "content" e "patch". Observando os comentários e buscando por notícias relacionadas aos jogos, podemos ver que ambas as empresas foram alvos de críticas que as acusam de ter lançado os jogos inacabados ou com muitos *bugs* [7, 8, 21, 23, 30].

O próximo tópico comum das críticas são o conjunto história e jogabilidade, como visto nas Figuras 6a e 6b. Esses pontos são abordados principalmente nas mídias que sofreram *review bombing* por motivos políticos, como é o caso dos jogos *The Last of Us 2* e *Horizon Forbidden West*. Os ataques aos jogos se deram principalmente pela orientação sexual das personagens principais [27]. Como demonstrando anteriormente por [19], os usuários que participam de ataques ligados a questões sociais tendem a demonstrar seu ódio indiretamente para escapar da moderação de conteúdo, pois comentários com discurso de ódio explícito são detectados com mais facilidade. Desse modo, focam suas críticas na história e jogabilidade de maneira geral.

O terceiro tópico envolve o julgamento sobre jogos percebidos pelo público como cópias pouco inovadoras de versões anteriores, um problema particularmente comum em mídias de esporte, como a franquia Madden. Na Figura 6d, podemos ver que "franchise" e "money" aparecem em destaque, relacionando o lançamento de novos títulos da franquia com o intuito de apenas ganhar dinheiro. Como já citado, este jogo de futebol americano, lançado anualmente, foi alvo de críticas dos usuários que reclamam da falta de mudanças significativas no título de um ano para o outro.

Por fim, existem as críticas diversas, que tratam de contextos específicos. Um desses casos é o jogo *Call of Duty: Modern Warfare*, que enfrentou uma enxurrada de comentários negativos devido à representação controversa dos russos [10], o que é constatado na Figura 6e, em que "Russian" aparece em destaque. O jogo foi criticado por insinuar que os russos estavam envolvidos no "Highway of Death", um ataque ocorrido durante a Guerra do Golfo, realizado por tropas aliadas aos Estados Unidos no Iraque, denunciado pelo uso excessivo de força [11].

6.2 Discurso de Ódio

Identificar discurso de ódio é uma tarefa complexa, especialmente em plataformas como o Metacritic, que possuem sistemas de moderação que já removem diretamente conteúdos mais ofensivos. No entanto, usuários muitas vezes conseguem contornar esses sistemas, utilizando linguagem que sugere, de forma indireta, discurso de ódio.

Por esse motivo, foi compilada uma lista extensiva de palavras-chave que, no contexto da plataforma, são geralmente utilizadas em enunciados com discurso de ódio. A lista começou a ser elaborada com a palavra "woke", que é majoritariamente utilizada em contextos de crítica à justiça social e racial [1]. A partir dessa palavra foi realizada uma análise manual de todos os comentários que continham ela, buscando outras palavras que a acompanham normalmente. Encontramos, por exemplo, as siglas SSJ (*social justice warrior*) e LGBTQIA+ (comumente utilizada de forma negativa no site). Também foram adicionadas à lista, pelos autores, algumas outras palavras consideradas explicitamente ofensivas.

Foram identificadas 22 palavras-chave em inglês. Considerando que 87.6% do conteúdo do site está nesse idioma, decidimos aplicar essa lista a todos os idiomas que representam mais de 1% do site. Isso inclui Espanhol com 4.5%, Russo com 2.8%, e Português com 2.3%. Logo, todas as palavras-chave foram traduzidas para cada um desses idiomas, cabendo adaptá-las devido à nuances dos idiomas e gírias, que variam conforme contexto linguístico. Cabe ressaltar que, devido à falta de conhecimento dos autores sobre russo, foi consultado um fórum do site Reddit, especializado no idioma, onde a tradução foi feita por terceiros.

Dessa maneira, foi possível elaborar a Tabela 2, que mostra as 5 palavras-chave mais utilizadas e os cinco jogos que tem mais palavras-chave, sendo esse percentual relativo ao total de comentários que possuem. É possível perceber que os jogos com a maior quantidade de palavras-chave são justamente os jogos que sofreram *review bombing* por questões sociais, como o gênero e sexualidade das personagens. Como observado anteriormente nas nuvens de palavras, os usuários tendem a esconder seu discurso de ódio, utilizando críticas gerais a história e aos personagens para justificar notas baixas e porque não gostaram do jogo.

Ao organizar o conjunto de dados em duas categorias, comentários com as palavras-chave e comentários sem as palavras-chave, foi possível notar uma diferença estatisticamente significativa na toxicidade. O conjunto com as palavras-chave tem uma média de



(a) The Last of Us Part II



(b) Horizon Forbidden West



(c) Grand Theft Auto: The Trilogy - The Definitive Edition



(d) Madden NFL 21



(e) Call of Duty: Modern Warfare



(f) Diablo IV

Figura 6: Comparação entre outros jogos

Palavra	Total	Jogo	Total
woke	1005	The Last of Us Part II	792
agenda	912	Mortal Kombat 11	200
sjw	776	Ghost of Tsushima	175
lgbt	473	Horizon Forbidden West DLC	157
lesbian	229	Battlefield V	133

Tabela 2: Jogos e Palavras-Chave

toxicidade de 0.291 ± 0.209 , enquanto o conjunto sem as palavras apresenta uma média de 0.118 ± 0.182 . Esses resultados reforçam a eficácia do método.

6.3 Discussão

Chegando ao final do estudo, foi finalmente possível responder às perguntas de pesquisa.

RQ1: Quais são as características de um jogo que sofreu *review bombing* online?

Com o auxílio dos centroides, foi possível determinar que um jogo que sofreu *review bombing* tem as seguintes características: Um número de comentários e *reviews* muito maior que a média. Uma grande similaridade entre os comentários, indicando um reclamação em comum entre os usuários. E por fim, uma toxicidade média e um número de palavras de baixo calão levemente maior que o normal.

RQ2: Como se comportam os usuários da plataforma e qual é o nível de toxicidade nas suas interações?

Em geral, os usuários comuns do Metacritic não são muito ativos, com a grande maioria tendo menos de 10 comentários. No entanto, ao observarmos os críticos, percebemos uma realidade diferente, onde eles tendem a ser muito mais ativos na plataforma. Quanto à toxicidade, o site consegue moderar bem, apresentando poucos casos extremos. Contudo, ainda há espaço para melhorias, pois é possível observar discrepâncias no gênero de esportes e no idioma português, que apresentaram muito mais toxicidade, conforme visto nas Figuras 3b e 3c. **RQ3:** Quais as principais motivações dos

comentários desses ataques?

As motivações dos ataques variam de jogo para jogo, mas é possível separar em 3 tópicos diferentes. (1) Críticas diretas à empresa e suas decisões relativas ao desenvolvimento e vendas, como lançar um jogo inacabado ou adicionar microtransações abusivas. (2) História e Jogabilidade, tópico geralmente ligados ao discurso de ódio contra minorias, onde os usuários disfarçam seu ódio em críticas gerais à história, personagens e jogabilidade. (3) Franquias repetitivas, em que as empresas desenvolvedoras lançam jogos anualmente sem muitas novidades e melhorias técnicas, porém cobrando preços caros, o que ocorre principalmente em franquias de esporte. Por fim, existem as outras críticas, que abordam temas específicos, que não compartilham um tópico em comum. Abaixo, é possível observar algumas dessas reclamações.

- “What a disgusting cash grab. It makes me so angry that rockstar removed the original games from the stores.”
- “they ruined this game. focused on political correctness. lgbt agenda. not worth the money. story is very weak.”
- “The same game as Madden 20. The face of the franchise is trash.”

7 CONCLUSÃO

A motivação deste trabalho é analisar, caracterizar e pensar metodologias para identificar o fenômeno de *review bomb* em plataformas de avaliação de mídias. Para tal, realizamos uma vasta coleta do Metacritic. A partir da limpeza e caracterização desses dados, conseguimos resultados relevantes a respeito do comportamento dos usuários e também percebemos que, mediante uma estratégia de aferição de toxicidade aliada à clusterização, é possível aferir quais jogos sofreram *review bombing*. Dentre nossos resultados, foi possível caracterizar o ataque de *Review bombing* como sendo uma grande onda de comentários altamente tóxicos e similares uns aos outros, refletindo a característica de campanha coordenada de insatisfação que eles representam. Também identificamos três categorias diferentes de tópicos que aparecem em *review bombings* no Metacritic: reclamações sobre as empresas desenvolvedoras e suas decisões empresariais, reclamações quanto a história e personagens, que geralmente estão permeadas por discurso de ódio e, por fim, reclamações sobre jogos de franquias muito semelhantes a seus antecessores.

Embora não encontrada diferenças significativas entre as plataformas PlayStation, Nintendo, PC e Xbox, vimos que o gênero do jogo pode ser um atributo relevante na caracterização dos ataques, com a categoria “Esportes” sendo o tipo com comentários mais tóxicos. Além disso, vimos que os comentários no Metacritic feitos em português exibem níveis de toxicidade mais elevados em comparação aos demais. Também notamos que discurso de ódio é bem presente dentro das comunidades, e é direcionado principalmente à comunidade LGBTQIA+, e se disfarça como críticas genéricas sobre a história, personagens e jogabilidade. Com nossos resultados, vimos que é possível identificar ataques aos jogos no Metacritic, e até mesmo suas motivações. Isso pode servir como uma ferramenta para ajudar na moderação dessas plataformas, reduzindo o impacto que as *review bombings* geram. Em trabalhos futuros, serão introduzidas novas métricas, como análises temporais e de sentimento, para identificar ataques, com o objetivo de conseguir detectar esses ataques enquanto eles ocorrem e alertar as plataformas.

REFERÊNCIAS

- [1] 2022. O que é 'woke' e por que o termo gera uma batalha cultural e política nos EUA. BBC. <https://www.bbc.com/portuguese/internacional-63547369> Online; Publicado em 8 de novembro de 2022.
- [2] Nikolay Archak, Anindya Ghose, and Panagiotis G Ipeirotis. 2011. Deriving the pricing power of product features by mining consumer reviews. *Management science* 57, 8 (2011), 1485–1509.
- [3] Suman Biswas. 2023. Why is Baldur's Gate 3 suddenly getting review bombed? SportsKeeda. <https://www.sportskeeda.com/esports/why-baldur-s-gate-3-suddenly-getting-review-bombed-10/08/2023>.
- [4] Giulio Giacomo Cantone, Venera Tomaselli, and Valeria Mazzeo. 2021. Ideology-driven polarisation in online ratings: the review bombing of The Last of Us Part II. *arXiv preprint arXiv:2104.01140* (2021).
- [5] Dalton Cooper. 2022. Diablo Immortal is Being Review Bombed. GAMERANT. <https://gamerant.com/diablo-immortal-review-bomb-metacritic/> Online; Publicado em 7 de Janeiro de 2022.
- [6] Ryan Craddock. 2019. Metacritic Review-Bombers Strike Again With FIFA 20 On Switch. Nintendo Life. https://www.nintendolife.com/news/2019/09/metacritic_review-bombers_strike_again_with_fifa_20_on_switch Online; Publicado em 30 de setembro de 2019.
- [7] John Esposito. 2023. Diablo 4 players refuse to buy expansion for an “unfinished” game. Dexerto. <https://www.dexerto.com/diablo/diablo-4-players-refuse-to-buy-expansion-for-an-unfinished-game-2368084/> Online; Publicado em 04 de novembro de 2023.
- [8] Raphael Giannotti. 2023. <https://www.adrenaline.com.br/games/pc-games/jogadores-continua-reclamando-do-remaster-de-gta-trilogy-mas-dessa-vez-no-steam/> Gamerant. Jogadores continuam reclamando do remaster de GTATrilogy, mas dessa vez no Steam Online; 20/01/2023.
- [9] ADAMS GREENWOOD-ERICKSEN, SCOTT R POORMAN, and ROY PAPP. 2013. On the Validity of Metacritic in Assessing Game Value. *Eludamos. Journal for Computer Game Culture* 7, 1 (2013), 101–127.
- [10] Charlie Hall. 2019. Call of Duty: Modern Warfare review bombed over Russian portrayal. Polygon. <https://www.polygon.com/2019/10/28/20936496/call-of-duty-modern-warfare-review-bombing-russian-military-highway-of-death> Online; Publicado em 28 de Outubro de 2019.
- [11] Charlie Hall. 2019. Call of Duty: Modern Warfare's Highway of Death controversy, explained. Polygon. <https://www.polygon.com/2019/10/30/20938550/call-of-duty-modern-warfare-highway-of-death-controversy> Online; 30/10/2019.
- [12] Nan Hu, Ling Liu, and Jie Jennifer Zhang. 2008. Do online reviews affect product sales? The role of reviewer characteristics and temporal effects. *Information Technology and Management* 9 (2008), 201–214.
- [13] Daniel Johnson, Christopher Watling, John Gardner, and Lennart E Nacke. 2014. The edge of glory: the relationship between metacritic scores and player experience. In *Proceedings of the first ACM SIGCHI annual symposium on Computer-human interaction in play*. 141–150.
- [14] Jordan K. Kobellarz and Thiago H. Silva. 2022. Should We Translate? Evaluating Toxicity in Online Comments when Translating from Portuguese to English. In *Proceedings of the Brazilian Symposium on Multimedia and the Web (Curitiba, Brazil) (WebMedia '22)*. ACM, 89–98. <https://doi.org/10.1145/3539637.3556892>
- [15] Alexander Kraskov, Harald Stögbauer, and Peter Grassberger. 2004. Estimating mutual information. *Physical review E* 69, 6 (2004), 066138.
- [16] Ben Kuchera. 2017. The anatomy of a review bombing campaign. Polygon. <https://www.polygon.com/2017/10/4/16418832/pubg-firewatch-steam-review-bomb> Online; Publicado em 4 de Outubro de 2017.
- [17] Lucas Lima, Julio CS Reis, Philipe Melo, Fabricio Murai, Leandro Araujo, Pantelis Vikatos, and Fabricio Benevenuto. 2018. Inside the right-leaning echo chambers: Characterizing gab, an unmoderated social system. In *2018 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM)*. IEEE, 515–522.
- [18] Júlia Mazuco. 2018. DICE sobre Battlefield 5: personagens femininas vieram para ficar. Voxel. www.tecmundo.com.br/voxel/195674-dice-battlefield-5-personagens-femininas-vieram.htm Online; 28/05/2018.
- [19] Philipe Melo and Clara Pimentel. 2022. A Campanha de Ódio contra The Last of Us Part II. In *Anais Estendidos do XXI Simpósio Brasileiro de Jogos e Entretenimento Digital*. SBC, 428–437.
- [20] Merriam-Webster.com Dictionary. 2024. Review bombing. Merriam-Webster. <https://www.merriam-webster.com/dictionary/review%20bombing> Online.
- [21] Rafael Monteiro. 2023. Diablo 4 tem bug que impede jogador de avançar no jogo; saiba evitar. TechTudo. <https://www.techtudo.com.br/noticias/2023/06/diablo-4-tem-bug-que-impede-jogador-de-avancar-no-jogo-saiba-evitar-edjogos.ghtml> Online; Publicado em 27 de junho de 2023.
- [22] João Gabriel Nogueira. 2023. Review bombing: o que é, como reconhecer e... Tem jeito de impedir? meuplaystation. <https://meups.com.br/especiais/review-bombing-o-que-reconhecer-impedir/> Online; Publicado em 23 de junho de 2023.
- [23] Tristan Ogilvie. 2021. Grand Theft Auto: The Trilogy - The Definitive Edition Review. IGN. <https://www.ign.com/articles/grand-theft-auto-the-trilogy-the-definitive-edition-review> Online; Publicado em 17 de novembro de 2021.
- [24] Tom Phillips. 2011. Metacritic culls review-bombing users. EUROGAMER. <https://www.eurogamer.net/metacritic-culls-review-bombing-users> Online; Publicado em 23 de Setembro de 2011.
- [25] Johnny Reynolds. [n. d.]. Top 10 Times Video Games Got Review Bombed. watchmojo. <https://www.watchmojo.com/articles/top-10-times-video-games-got-review-bombed> Online.
- [26] José V. Rossi. 2023. Diablo 4 is Being Review Bombed on Steam. GAMERANT. <https://gamerant.com/diablo-4-steam-review-bombed-why/> Online; Publicado em 10 de Outubro de 2023.
- [27] Gabriel Sales. 2023. Metacritic mudará estratégia de moderação após "review bomb" contra DLC de Horizon Forbidden West. IGN. <https://br.ign.com/horizon-forbidden-west/108210/news/metacritic-mudara-estrategia-de-moderacao-apos-review-bomb-contra-dlc-de-horizon-forbidden-west> Online; Publicado em 25 de Abril de 2023.
- [28] Tiago Santos, Florian Lemmerich, Markus Strohmaier, and Denis Helic. 2019. What's in a review: Discrepancies between expert and amateur reviews of video games on metacritic. *Proceedings of the ACM on human-computer interaction* 3, CSCW (2019), 1–22.
- [29] Amanda Sharkey, Balázs Kovács, and Greta Hsu. 2023. Expert critics, rankings, and review aggregators: The changing nature of intermediation and the rise of markets with multiple intermediaries. *Academy of Management Annals* 17, 1 (2023), 1–36.
- [30] Bruna Telles. 2023. Another Diablo 4 Bug is Killing Players' Progress. Gamerant. <https://gamerant.com/diablo-4-bug-helltides-cinders-lost-progress/> Online; Publicado em 14 de julho de 2023.
- [31] Markos Viggiano, Dayi Lin, Abram Hindle, and Cor-Paul Bezemer. 2021. What causes wrong sentiment classifications of game reviews? *IEEE Transactions on Games* 14, 3 (2021), 350–363.
- [32] Asher Weiss. 2018. 21st century film criticism: the evolution of film criticism from professional intellectual analysis to a democratic phenomenon. (2018).
- [33] Wikipedia. 2024. List of review-bombing incidents. Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/wiki/List_of_review-bombing_incidents Online; Atualizado em 8 de maio de 2024.
- [34] Preston Willke. 2021. 10 Games That Got Review-Bombed Before They Were Even Released. cbr. <https://www.cbr.com/games-review-bombed-before-release/> Online; Publicado em 25 de Agosto de 2021.