

Um panorama sobre *Community Smells* e suas Relações com Aspectos Sociais em Repositórios Open Source de Jogos Indie

An Overview about Community Smells and their Relationship with Social Aspects in Open Source Repositories of Indie Games

Gustavo Eufrausino de Medeiros¹, Gabriel Zoéga Fernandes¹, Estela Miranda Batista¹, Otávio Santos Gomes¹, Gláucia Braga e Silva¹

¹Universidade Federal de Viçosa (UFV) Campus Florestal - MG

{gustavo.medeiros, gabriel.z.fernandes, estela.batista, otavio.s.gomes, glaucia}@ufv.br

Abstract. *Introduction:* The analysis of social aspects such as social debt, diversity, knowledge concentration, and turnover is essential to identify the main impacts on team collaboration and project success. **Objective:** To present an overview of open-source repositories of indie games regarding the presence of community smells and their relationships with knowledge concentration, diversity, and team turnover. **Methodology:** Data collection from issue tracking in 15 open-source indie game repositories on GitHub, followed by the application of metrics related to the identification of community smells, team diversity calculation, knowledge concentration, and turnover rates. **Results:** The results showed the presence of various community smells in almost all the analyzed repositories. Low female representation was also observed across all repositories, along with a clear culture of knowledge concentration in the hands of a few developers and significant turnover rates within the teams.

Keywords Social Debt, Gender Diversity, Time Zone Diversity, Knowledge Concentration, Turnover.

Resumo. *Introdução:* A análise de aspectos sociais, como débito social, diversidade, concentração de conhecimento e rotatividade, é fundamental se identificar os principais impactos sobre a colaboração entre os membros da equipe e o sucesso dos projetos. **Objetivo:** apresentar um panorama de repositórios open source de jogos indie no que compete à presença de community smells e suas relações com concentração de conhecimento, diversidade e rotatividade nas equipes. **Metodologia:** Coleta de dados de issue tracking, a partir de 15 repositórios open source de jogos indie no GitHub, com posterior aplicação de métricas relacionadas à identificação de community smells, ao cálculo de diversidade das equipes, à concentração de conhecimento e aos índices de rotatividade. **Resultados:** Os resultados mostraram a presença de diversos community smells em praticamente quase todos os repositórios analisados. Também foi observada baixa representatividade feminina em todos os repositórios, uma cultura evidente de concentração de conhecimento na mão de poucos desenvolvedores e índices expressivos de rotatividade nas equipes. **Palavras-Chave** Débito Social, Diversidade de gênero, Diversidade de fuso horário, Concentração de conhecimento, Rotatividade

1. Introdução

Desenvolvimento de *software* é uma atividade social e humano-dependente que exige equipes multidisciplinares a fim de produzir *software* de qualidade. Devido à sua natureza colaborativa, projetos de *software* tendem a apresentar melhor desempenho, quando há constante debate de ideias oriundas de perspectivas diversas. Nesse contexto, surgem os chamados débitos sociais, conhecidos como *community smells*. [Caballero-Espinosa et al. 2023] apontam que esses *community smells* são causadores de dívidas sociais em repositórios de *software* que podem levar a dívidas técnicas. Esses *smells* podem influenciar ou ser influenciados por aspectos como comunicação, taxas de diversidade e rotatividade dos times. A ausência de estruturas de comunicação eficazes pode levar à concentração do conhecimento sobre o desenvolvimento em um número reduzido de pessoas — ou até mesmo em um único desenvolvedor —, elevando o risco de descontinuação do projeto. Para mensurar esse risco, [Avelino et al. 2016] abordam o conceito de *truck factor*, uma métrica que determina o número mínimo de desenvolvedores cuja saída comprometeria a continuidade do projeto.

No que compete à diversidade de gênero nas equipes, [Catolino et al. 2019] destacam a importância de equipes diversas na melhoria das estruturas comunicacionais do projeto, apontando que o aumento da diversidade de gênero em repositórios *software* reduz a incidência de débitos sociais associados a falhas na comunicação. No entanto, existem outras categorias de diversidade, como fuso horário, neurodivergência, raça, etnia, orientação sexual, dentre outros, que também têm influência tanto sobre a comunicação quanto sobre a produtividade da equipe [Motta e e Silva 2022] quanto sobre os resultados do projeto como um todo.

A presença de *community smells* também pode influenciar o quanto atrativos são os projetos de *software*. [Ferreira et al. 2020] avaliaram o grau de atratividade ou repulsividade de repositórios de *software*, analisando a rotatividade de *core developers* em repositórios de *software* de código aberto. Esses desenvolvedores possuem papel central no projeto, sendo, por muitas vezes, responsáveis por um volume significativo de contribuições e detentores de conhecimento essencial sobre seu desenvolvimento. Baixa diversidade entre membros do projeto e comunicação ineficiente favorecem a concentração de conhecimento, aumento da presença de *community smells* e, consequentemente, o afastamento de desenvolvedores.

Problemáticas de cunho social presentes no desenvolvimento de *software open source* também se manifestam na indústria de jogos digitais. Considerando-se o recorte dos jogos *indie*, que tiveram um crescimento expressivo em 2024 [VGI 2024], os principais gatilhos são os longos ciclos de desenvolvimento dessas aplicações [Politowski et al. 2021], as estruturas colaborativas menos consolidadas [Chueca et al. 2024, Freeman e McNeese 2019], bem como a predominância masculina nas equipes [Prieto e Nesteriuk 2021, Drummond et al. 2022, Batista et al. 2024]. [Freeman et al. 2023] destacam a baixa representatividade feminina nas equipes de desenvolvimento de jogos *indie*: dos 28 participantes do estudo, apenas 4 se identificaram como mulheres. Esses fatores contribuem para a concentração de conhecimento em um número restrito de desenvolvedores e aumento da rotatividade nas equipes.

Este trabalho apresenta um panorama de repositórios *open source* de jogos *indie* no que compete à presença de *community smells* e suas relações com concentração

de conhecimento, diversidade e rotatividade nas equipes. O estudo é orientado pelas seguintes questões de pesquisa:

- **QP1.** Quais são os *community smells* mais frequentes nos repositórios open source de jogos *indie*?
- **QP2.** Qual o panorama da concentração de conhecimento em repositórios open source de jogos *indie*?
- **QP3.** Quais são os índices de diversidade de gênero e de fuso horário nos repositórios *open source* de jogos *indie*?
- **QP4.** Quais são as taxas de rotatividade nos repositórios de jogos *indie*?

Orientado por estas questões de pesquisa, este estudo objetiva obter um conjunto de dados que expresse os problemas relacionados às estruturas de comunicação em projetos de *software*, a partir da análise de repositórios *open source* de jogos *indie*. As contribuições pretendidas incluem: 1. Analisar o panorama dos *community smells* presentes nesses projetos; 2. Identificar e avaliar os valores de *truck factor*; 3. Medir os índices de diversidade nas equipes de desenvolvimento; 4. Calcular as taxas de rotatividade de *core developers*; 5. Fornecer uma análise quantitativa dos fatores que podem acentuar ou mitigar os problemas nas estruturas de comunicação observadas.

O restante deste trabalho está organizado como segue: a seção 2 traz os trabalhos relacionados; a seção 3 apresenta os materiais e métodos; os resultados obtidos são discutidos na seção 4; por fim, as considerações finais são apresentadas na seção 5.

2. Trabalhos Relacionados

A análise de *community smells* em comunidades *open source* também foi conduzida por [Catolino et al. 2019], explorando, em especial, o impacto da diversidade de gênero neste contexto. Foram analisados dados de 40 projetos, a partir da ferramenta CODEFACE4SMELLS. O estudo destaca que a diversidade de gênero possui um impacto positivo na mitigação de problemas de comunicação em equipes de desenvolvimento de *software*. [Almarimi et al. 2021, Almarimi et al. 2023] também investigaram a presença de *community smells* em projetos de *software open source* em diferentes contextos, por meio do desenvolvimento e da aplicação da ferramenta csDetector. [Almarimi et al. 2023] conduziram um estudo empírico aplicando a ferramenta para analisar dados de 143 projetos *open source* hospedados no GitHub. Dentre os principais achados, destacam-se como características influentes, o número de desenvolvedores por fuso horário e a proporção de issues com sentimentos negativos.

Dentre os *smells* identificados pela csDetector [Almarimi et al. 2021], destaca-se o *truck factor*, abordado na literatura como *smell* ([Almarimi et al. 2021, Almarimi et al. 2023] ou como métrica de quantificação de concentração de conhecimento em projetos [Zazworka et al. 2010, Avelino et al. 2016]. Considerando o método XP, [Zazworka et al. 2010] propuseram uma abordagem automatizada para estimar o *truck factor*, baseada na propriedade de código, simulação da saída de desenvolvedores e cobertura de conhecimento sobre componentes do sistema. Posteriormente, [Avelino et al. 2016] aprimoraram essa proposta utilizando o *degree-of-authorship* (DOA), análise do histórico de contribuições, remoção iterativa de desenvolvedores e um critério de cobertura mínima de 50% dos arquivos. Enquanto a abordagem de [Zazworka et al. 2010] apresentaram resultados

preliminares, [Avelino et al. 2016] trouxeram contribuições mais robustas, influenciando no desenvolvimento da ferramenta truckfactor de [Pfeiffer 2021].

Tendo em vista que a análise da atração e repulsão de repositórios *open source* também é fundamental para compreender sua continuidade e desempenho. Como apontado por [Ferreira et al. 2020], a baixa taxa de contribuição pode comprometer a precisão das análises, tornando essencial o foco na rotatividade de *core developers*. O estudo investiga a saída desses desenvolvedores em janelas temporais de aproximadamente um ano, buscando identificar os fatores que influenciam a atratividade dos repositórios e seu impacto no desenvolvimento de *software*.

No que compete à análise sobre fatores sociais em comunidades de jogos *indie*, [Freeman e McNeese 2019] abordaram práticas e experiências sociais em times de desenvolvimento de jogos *indie*, explorando as características dos desenvolvedores e as interações entre os mesmos. Através do estudo empírico e qualitativo, baseado em fóruns online, entrevistas e eventos como a Global Game Jam¹, [Freeman e McNeese 2019] identificaram que equipes *indie* se formam por laços de amizade ou interesses técnicos, promovendo aprendizagem colaborativa. Também foram observados efeitos sociais variados, como isolamento de alguns membros e aumento de autoconfiança ou soberba em outros. Contudo, esses aspectos não foram caracterizados como *community smells*, pois o foco do estudo foi compreender as dinâmicas sociais e organizacionais dos times de desenvolvimento de jogos *indie*.

A pesquisa realizada por [Catolino et al. 2019] se assemelha ao estudo aplicado no presente trabalho, uma vez que [Catolino et al. 2019] analisaram o impacto da diversidade de gênero e da participação de mulheres em comunidades *open source* na presença de *community smells*. Assim, ao integrar a análise de *community smells*, taxas de rotatividade, diversidade de gênero e *truck factor*, este estudo busca aprofundar a compreensão sobre os fatores que influenciam a sustentabilidade de comunidades de desenvolvimento de *software*, através do estudo das estruturas socio-colaborativas no contexto de jogos *indie* em repositórios *open source*. Com base nas abordagens existentes e nas lacunas identificadas na literatura, investigar como esses elementos se relacionam pode evidenciar quais fatores podem garantir a continuação ou descontinuação de projetos de *software*.

3. Materiais e métodos

Para responder às questões de pesquisa propostas, foram executadas três etapas: seleção de repositórios; coleta de dados; análise de dados com aplicação de métricas.

3.1. Seleção dos repositórios

A escolha dos repositórios foi baseada em uma busca direcionada por projetos *open source* de jogos *indie* hospedados no GitHub que atendessem aos seguintes critérios: (a) possuir entre 1.000 e 11.000 submissões; (b) contar com uma equipe de, no mínimo, 10 desenvolvedores; (c) apresentar histórico de atividade recente, com ao menos uma contribuição registrada até um ano antes da coleta dos dados; (d) possuir um tempo de vida entre 5 e 10 anos. Tais critérios foram definidos com o intuito

¹<https://v3.globalgamejam.org/>

de garantir que os repositórios analisados apresentassem uma base de colaboradores relativamente estável e um volume significativo de atividade ao longo do tempo. Em especial, o critério de atividade recente — ao menos uma contribuição no último ano — é respaldado pela literatura, sendo adotado em estudos anteriores como os de [Cánovas Izquierdo et al. 2017], [Coelho et al. 2018] e [Khondhu et al. 2013], o que justifica sua inclusão neste trabalho. Após a aplicação dos critérios, 15 repositórios foram selecionados para o estudo.

3.2. Coleta de dados

A coleta de dados foi realizada por meio da ferramenta csDetector [Almarimi et al. 2023], e complementada por uma busca manual de perfis de desenvolvedores no GitHub e em mídias sociais associadas. A escolha da ferramenta csDetector foi guiada pelo estudo comparativo de [André Filipe Queiroz de Melo e Soares 2023] e por outros estudos em que a mesma foi adotada, como é o caso de [Voria et al. 2022].

A csDetector foi utilizada para a extração automatizada da maior parte dos dados, incluindo a detecção de *community smells* e informações dos desenvolvedores, como identificador, número de submissões e fuso horário. Além da csDetector foi utilizada a ferramenta truckfactor [Pfeiffer 2021] que identifica e quantifica o mínimo de desenvolvedores capazes de darem continuidade ao projeto. Dados relacionados a desenvolvedores não obtíveis pela csDetector e pela truckfactor foram coletados manualmente através da busca em mídias sociais ao quais os desenvolvedores estavam envolvidos.

O processo de coleta pela csDetector consistiu em 2 fases: (1) executar a csDetector sobre cada repositório remoto selecionado sem determinar intervalo de tempo, ou seja, extrair os dados de todo o tempo de vida do repositório; (2) executar a csDetector apenas para intervalos de tempo de aproximadamente 1 ano a fim de estudar a rotatividade de *core developers* dentro do repositório. Enquanto a truckfactor foi executada sobre todo período de vida do repositório.

3.3. Análise de dados

Para responder as questões de pesquisa propostas, os dados obtidos foram avaliados pela ferramenta: Jupyter Notebook². Nesta etapa, são conduzidas análises sobre os dados coletados dos 15 repositórios: a quantidade de *community smells* identificados por repositório; a ocorrência de diferentes tipos de *smells* em cada projeto; o número de desenvolvedores associados a cada repositório; a taxa média de rotatividade; os valores de *truck factor*; os índices de diversidade de gênero e de fuso horário; e a distribuição de homens e mulheres por repositório.

Para a extração dos *Community Smells* foi utilizada a ferramenta csDetector [Almarimi et al. 2023]. A ferramenta é capaz de capturar até 10 smells: *Organizational Silo Effect* (OSE), *Black-cloud Effect* (BCE), *Prima-donnas Effect* (PDE), *Sharing Villainy* (SV), *Organizational Skirmish* (OS), *Solution Defiance* (SD), *Radio Silence* (RS), *Truck Factor Smell* (TFS), *Unhealthy Interaction* (UI), *Toxic Communication* (TC). Importante ressaltar que, embora a csDetector trate o *truck factor* como um *community smell*, no contexto deste estudo, ele será usado como uma métrica de concentração de

²<https://jupyter.org/>

conhecimento, de forma similar a [Zazworka et al. 2010, Avelino et al. 2016] através da ferramenta *truckfactor* de [Pfeiffer 2021].

A fim de avaliar a diversidade da equipe dentro dos repositórios escolhidos, duas métricas foram selecionadas para a análise: diversidade de gênero [Catolino et al. 2019] e de fuso horário [Holmstrom et al. 2006] e [Al-Ani e Redmiles 2009]. Para o cálculo da diversidade de gênero, utiliza-se o índice de diversidade de Blau [Blau et al. 1977], cujo cálculo é apresentado na Equação 1, onde a partir de um total de N, a porcentagem P de indivíduos em cada categoria i. Neste trabalho, N = 2 (quando analisado diversidade de gênero) e N = M (quando analisado fuso horários), sendo M a quantidade de fuso horários obtidos dentro de cada repositório. Vale destacar que, neste trabalho, considerou-se N = 2 devido à limitação das ferramentas disponíveis na literatura para a identificação de gênero, que permitem apenas a classificação binária entre os gêneros masculino e feminino.

$$Blau = 1 - \sum_{i=1}^N P_i^2 \quad (1)$$

O índice de diversidade de Blau varia entre 0 (ausência total de diversidade) e 1 (distribuição perfeitamente equilibrada entre grupos, embora esse valor não seja atingível na realidade). No caso da diversidade por fuso horário, o índice foi calculado com base na diferença de tempo (em segundos) fornecida pela csDetector, posteriormente convertida para horas. A partir disso, determinou-se a proporção de desenvolvedores em cada fuso horário e, então, aplicou-se a Equação 1. Para a diversidade de gênero, o cálculo segue a mesma lógica. Contudo, foi necessário utilizar também a ferramenta Namsor³, que infere gênero a partir do nome passado como entrada.

Para o cálculo da rotatividade nas equipes, este estudo analisou a rotatividade de *core developers* em projetos de *software*, seguindo a abordagem de [Ferreira et al. 2020]. Para isso, foi desenvolvido um algoritmo (10.5281/zenodo.1579984) baseado na heurística proposta por [Ferreira et al. 2020] para identificação e análise da rotatividade desses desenvolvedores. O algoritmo segue uma métrica de duas etapas: (1) seleção dos *core developers*, com base na quantidade de submissões; (2) cálculo da rotatividade desses desenvolvedores ao longo do tempo. A identificação considerou os responsáveis por 80% das submissões em um período de aproximadamente um ano, excluindo aqueles cuja contribuição foi inferior a 5% do total.

A partir do conjunto de *core developers* identificados em intervalos anuais consecutivos, define-se o conjunto B (*Set B*) como o grupo obtido no período de $n - 1$, onde n representa o ano de interesse. Já o conjunto A (*Set A*) corresponde ao intervalo de $n - 1$ a $n - 2$. Esse processo se repete alternadamente até alcançar o ano da primeira submissão registrada no repositório.

Primeiramente, define-se o conjunto de desenvolvedores que deixam um projeto (*Core Developers Leavers* (CDL) dentro do intervalo de interesse. A Equação 2 exprime como é definido esse conjunto.

$$CDL(n) = SetA - SetB \quad (2)$$

³<https://namsor.app/>

Em seguida, calcula-se a taxa de saída no intervalo de interesse. A Equação 3 apresenta o cálculo para definir a saída de *core developers* do repositório ($CDL_{Rate}(n)$).

$$CDL_{Rate}(n) = \frac{|CDL(n)|}{|Set(A)|} \times 100 \quad (3)$$

A partir do resultado obtido através da Equação 2, é possível calcular a taxa de rotatividade de *core developers* ($CDT_{Rate}(n)$) dentro do intervalo de interesse como apresentado na Equação 4.

$$CDT_{Rate}(n) = \frac{|CDL(n)|}{avg(|SetA| + |SetB|)} \times 100 \quad (4)$$

Após calcular a $CDT_{Rate}(n)$ para todos as janelas de tempo em um repositório, basta realizar a média dessas taxas para obter a taxa de rotatividade média de *core developers* do repositório.

4. Resultados e Discussão

4.1. QP1. Quais são os *community smells* mais frequentes nos repositórios open source de jogos indie?

A Figura 1 mostra a distribuição dos *community smells* nos repositórios analisados. Os *smells* mais recorrentes — UI, SV, SD, TC e PDE — estão diretamente ligados a estruturas comunicacionais frágeis e ineficientes.

Dentre os nove *community smells* estudados, a menor ocorrência foi de quatro *smells* por repositório (cerca de 44.45% do total), sugerindo que projetos desse nicho compartilham características organizacionais e comunicacionais que resultam em uma dívida social significativa.

O único *smell* ausente foi OS, indicando que, apesar das deficiências na comunicação, o nível técnico dos desenvolvedores parece equilibrado. No entanto, a persistência da dívida social pode comprometer a qualidade do projeto e favorecer a ocorrência de *code smells*, conforme apontado por [Palomba et al. 2021].

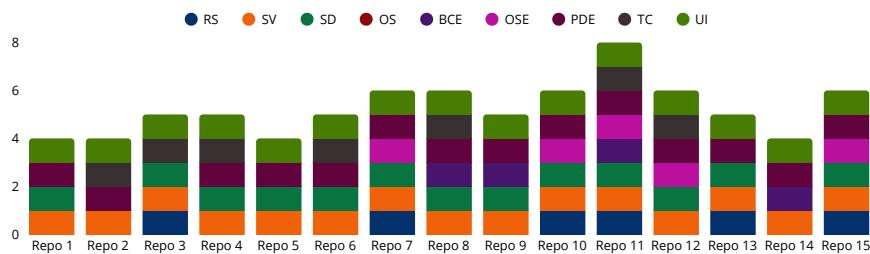


Figura 1. Distribuição dos *community smells* ao longo dos repositórios

4.2. QP2. Qual o panorama da concentração de conhecimento em repositórios open source de jogos indie?

A distribuição do conhecimento, medida através da métrica *truck factor* (TF), nos repositórios pode ser vista na Tabela 1.

Como é possível observar, o maior valor de *truck factor* obtido ao longo dos 15 repositórios foi de apenas 2. Vale destacar que cada repositório conta com, no mínimo, 10 desenvolvedores, o que indica que, na maioria dos casos, menos de 20 por cento dos desenvolvedores possuem conhecimento suficiente para dar continuidade ao projeto por conta própria.

É importante ressaltar que o repositório 14 conta com mais de nove membros; no entanto, foi possível identificar o gênero de apenas nove integrantes. Assim, os desenvolvedores cujo gênero não pôde ser inferido foram desconsiderados da análise deste e dos outros repositórios.

Outro ponto importante destacar é o repositório 6 ser o único a apresentar *truck factor* feminino diferente de 0. Isso pode indicar que o repositório possivelmente protagoniza o envolvimento de mulheres na tomada de decisões do projeto, embora, ainda de forma muito limitada.

Além disso, a elevada incidência de *community smells* relacionados a estruturas de comunicação frágeis pode ser um fator determinante para a baixa distribuição de conhecimento nos repositórios. A concentração de conhecimento pode ser atribuída a diversos fatores, entre eles: a falta de escrita de documentação técnica, o isolamento social de membros da equipe, ambientes de desenvolvimento tóxicos e estruturas de comunicação falhas evidenciados principalmente pelos *smells Sharing Villainy* (presente em todos os repositórios analisados) e *Solution Defiance* (presente em 13/15 repositórios).

Tabela 1. Distribuição do conhecimento ao longo dos repositórios

Repositórios	Valor TF	Valor TF Mulheres	Valor TF Homens
Repo 1	1	0	1
Repo 2	1	0	1
Repo 3	2	0	2
Repo 4	2	0	2
Repo 5	1	0	1
Repo 6	1	1	0
Repo 7	1	0	1
Repo 8	2	0	2
Repo 9	1	0	1
Repo 10	1	0	1
Repo 11	1	0	1
Repo 12	1	0	1
Repo 13	2	0	2
Repo 14	1	0	1
Repo 15	1	0	1

4.3. QP3. Quais são os índices de diversidade de gênero e de fuso horário nos repositórios open source de jogos indie?

A partir dos dados coletados, foi possível usar o índice de diversidade de Blau para medir a diversidade de gênero e de fuso horário (*time zones*) nos repositórios. Vale destacar que alguns repositórios apresentaram índices de diversidade nulos em função da ausência de mulheres nestes repositórios.

Ao observar a Figura 2a, é possível notar que a diversidade de gênero nesses repositórios é extremamente baixa, o que é enfatizado quando se observa a disparidade

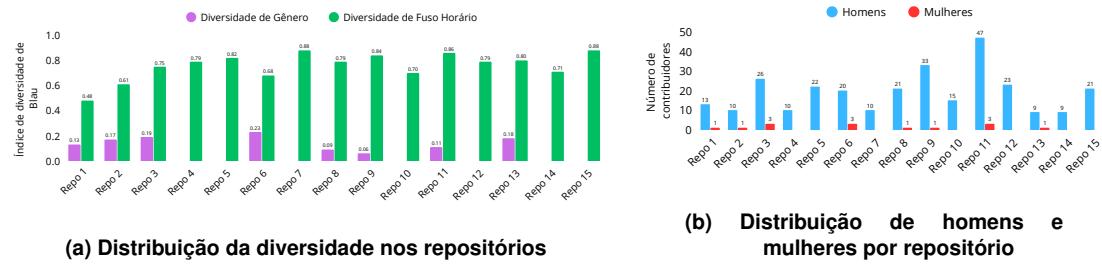


Figura 2. Panorama de diversidade nos repositórios

na distribuição de homens e mulheres nas equipes por repositório (Figura 2b). Entre os 15 repositórios analisados, foram observados: 7 repositórios (46.67%) com equipes inteiramente masculinas; 5 repositórios (33.33%) com apenas 1 mulher; 3 repositórios (20%) com 3 mulheres. Além disso dentre aqueles que possuem mulheres, o índice de diversidade de Blau para gênero não ultrapassa metade do máximo possível (0.5), sendo o maior deles apenas 0.23. Vale destacar que por se tratar de uma variável binária, o valor máximo do índice é 0.5, indicando equilíbrio perfeito entre homens e mulheres.

Em contrapartida, o índice de diversidade de fuso horário está bem diversificado, o que pode ser um indicativo de que essas comunidades de desenvolvedores são consideravelmente abertas e tolerantes a membros de outras regiões do globo. O valor máximo do índice de diversidade de Blau depende da quantidade de fusos horários distintos em cada repositório, aumentando conforme essa diversidade. Com 5 fusos horários uniformemente distribuídos (20% cada), o valor máximo do índice é 80%. Desta maneira, é possível supor que comunidades que atuam nesse nicho podem ser de fato abertas a receber membros diversos, mas que algo nessas comunidades provoque a evasão de desenvolvedoras, tendo em vista a grande quantidade de homens nesses repositórios. Nesta análise, foram considerados 303 perfis dos quais 289 foram classificados como homens e 14 como mulheres.

4.4. QP4. Quais são as taxas de rotatividade nos repositórios de jogos indie?

A Tabela 2 mostra as taxas de rotatividade de cada repositório, cruzadas com o índice de Blau correspondente. Observa-se que o repositório 6 apresenta o maior índice de diversidade, apesar de baixo, e também a maior taxa de rotatividade (55.00).

Em particular, o repositório 3, que possui alta rotatividade, indica que a combinação de *community smells* como UI, SV, SD, TC e PDE pode aumentar a repulsividade do projeto para os desenvolvedores. Esse comportamento é semelhante ao observado no repositório 6, que compartilha os mesmos *smells*, com exceção de RS.

Por outro lado, os repositórios 7 e 15 não apresentaram rotatividade e compartilham seis *community smells* (RS, SV, SD, BCE, PDE e UI). Isso pode indicar que os desenvolvedores se acomodaram à presença desses problemas ou que sua manifestação ocorre de forma mais branda. Além disso, a ausência do *smell* TC em alguns repositórios (1, 5, 7, 9, 10, 13, 14 e 15) sugere que a toxicidade na comunicação pode ser um fator de evasão de desenvolvedores e também impactar a distribuição de conhecimento, dado que apenas um desses repositórios apresentou *truck factor* superior a 1.

Tabela 2. Rotatividade, Diversidade de Gênero e *Truck Factor*

Repositórios	Contribuidores	Rotatividade média de <i>core developers</i>	Div. de Gênero	TF Total
Repo 1	14	16.67	0.13	1
Repo 2	11	25.00	0.17	1
Repo 3	29	53.65	0.19	2
Repo 4	10	44.45	0	2
Repo 5	22	42.50	0	1
Repo 6	23	55.00	0.23	1
Repo 7	10	0.00	0	1
Repo 8	22	40.95	0.09	2
Repo 9	34	14.29	0.06	1
Repo 10	15	9.52	0	1
Repo 11	50	21.70	0.11	1
Repo 12	23	29.17	0	1
Repo 13	10	33.34	0.18	2
Repo 14	9	13.33	0	1
Repo 15	21	0.00	0	1

5. Conclusão

Esta pesquisa analisou a distribuição de conhecimento em repositórios *open source* de jogos *indie*, abordando aspectos como diversidade de gênero, fuso horário, concentração de conhecimento entre desenvolvedores e rotatividade de *core developers*.

Para a amostra analisada, os resultados mostraram a presença de diversos *community smells* em praticamente quase todos os repositórios. Também foi observada baixa representatividade feminina, sendo que 7 dos 15 tinham equipes compostas por apenas homens e os demais com uma ou três mulheres apenas. Com relação ao *truck factor*, os índices variaram entre 1 e 2, o que indica uma concentração de conhecimento preocupante. Por fim, também foram observados índices expressivos de rotatividade, sendo que os maiores índices ocorreram nos repositórios com maior diversidade de gênero e em 2 repositórios, majoritariamente masculinos, obteve-se rotatividade zero.

Entre as limitações do estudo, destaca-se a natureza binária dos resultados fornecidos pela ferramenta *csDetector*, que restringe análises mais detalhadas sobre os *smells*. Para mitigar essa limitação, foram aplicadas métricas externas como *truck factor*, diversidade de gênero, fuso horário e rotatividade. Outra limitação diz respeito à adivinhação de gênero dos desenvolvedores, por meio da ferramenta Namsor. Apesar da ferramenta apresentar boa precisão dos resultados, conforme já apontado por [Sebo 2023], o gênero de algumas pessoas desenvolvedoras não puderam ser identificados, sendo rotulados como indefinidos. Para contornar essa limitação, foi feita uma busca manual nos perfis dessas pessoas desenvolvedoras, com o intuito de identificar o gênero pelo nome/nickname, pela foto ou por outros atributos explícitos nos perfis.

Trabalhos futuros poderiam incluir o aumento da amostra de projetos para viabilizar a aplicação de análises estatísticas sobre os dados de forma a encontrar correlações que permitam a identificação de novos atributos que possam ser usados em modelos de previsão de rotatividade, em função da presença de *smells*, por exemplo. Além disso, uma pesquisa futura poderia investigar outras ferramentas de identificação de *community smells* que fornecam resultados mais detalhados e não apenas binários.

Referências

Al-Ani, B. e Redmiles, D. (2009). In strangers we trust? findings of an empirical study of distributed teams. In 2009 Fourth IEEE International Conference on Global Software Engineering, pages 121–130.

Almarimi, N., Ouni, A., Chouchen, M., e Mkaouer, M. W. (2021). csdetector: an open source tool for community smells detection. In Proceedings of the 29th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering, ESEC/FSE 2021, page 1560–1564, New York, NY, USA. Association for Computing Machinery.

Almarimi, N., Ouni, A., Chouchen, M., e Mkaouer, M. W. (2023). Improving the detection of community smells through socio-technical and sentiment analysis. Journal of Software: Evolution and Process, 35(6):e2505.

André Filipe Queiroz de Melo e Soares, Tiago Lima Massoni, F. V. B. M. A. d. B. (2023). Community smells in software teams: a systematic literature mapping.

Avelino, G., Passos, L., Hora, A., e Valente, M. T. (2016). A novel approach for estimating Truck Factors . In 2016 IEEE 24th International Conference on Program Comprehension (ICPC), pages 1–10, Los Alamitos, CA, USA. IEEE Computer Society.

Batista, E., Silva, T., Silva, G., e Aylon, L. (2024). Mais que um jogo: Explorando a diversidade de gênero e sua influência na produção de jogos indie no brasil. In Anais do XXIII Simpósio Brasileiro de Jogos e Entretenimento Digital, pages 640–651, Porto Alegre, RS, Brasil. SBC.

Blau, P. M. et al. (1977). Inequality and heterogeneity: A primitive theory of social structure, volume 7. Free Press New York.

Caballero-Espinosa, E., Carver, J. C., e Stowers, K. (2023). Community smells—the sources of social debt: A systematic literature review. Inf. Softw. Technol., 153(C).

Catolino, G., Palomba, F., Tamburri, D. A., Serebrenik, A., e Ferrucci, F. (2019). Gender diversity and women in software teams: How do they affect community smells? In 2019 IEEE/ACM 41st International Conference on Software Engineering: Software Engineering in Society (ICSE-SEIS), pages 11–20.

Chueca, J., Verón, J., Font, J., Pérez, F., e Cetina, C. (2024). The consolidation of game software engineering: A systematic literature review of software engineering for industry-scale computer games. Information and Software Technology, 165:107330.

Coelho, J., Valente, M. T., Silva, L. L., e Shihab, E. (2018). Identifying unmaintained projects in github. In Proceedings of the 12th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement, ESEM '18, New York, NY, USA. Association for Computing Machinery.

Cánovas Izquierdo, J. L., Cosentino, V., e Cabot, J. (2017). An empirical study on the maturity of the eclipse modeling ecosystem. In 2017 ACM/IEEE 20th International Conference on Model Driven Engineering Languages and Systems (MODELS), pages 292–302.

Drummond, B., Salgado, L., e Viterbo, J. (2022). What are the challenges faced by women in the games industry? In Anais do XVI Women in Information Technology, pages 167–178, Porto Alegre, RS, Brasil. SBC.

Ferreira, F., Silva, L. L., e Valente, M. T. (2020). Turnover in open-source projects: The case of core developers. In Proceedings of the XXXIV Brazilian Symposium on Software Engineering, SBES '20, page 447–456, New York, NY, USA. Association for Computing Machinery.

Freeman, G., Li, L., Mcneese, N., e Schulenberg, K. (2023). Understanding and mitigating challenges for non-profit driven indie game development to innovate game production. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, CHI '23, New York, NY, USA. Association for Computing Machinery.

Freeman, G. e McNeese, N. (2019). Exploring indie game development: Team practices and social experiences in a creativity-centric technology community. Computer Supported Cooperative Work (CSCW), 28.

Holmstrom, H., Conchuir, E. O., Agerfalk, P. J., e Fitzgerald, B. (2006). Global software development challenges: A case study on temporal, geographical and socio-cultural distance. In 2006 IEEE International Conference on Global Software Engineering (ICGSE'06), pages 3–11.

Khondhu, J., Capiluppi, A., e Stol, K.-J. (2013). Is it all lost? a study of inactive open source projects. In Petrinja, E., Succi, G., El Ioini, N., e Sillitti, A., editors, Open Source Software: Quality Verification, pages 61–79, Berlin, Heidelberg. Springer Berlin Heidelberg.

Motta, T. e e Silva, G. B. (2022). Análise quali-quantitativa sobre a influência da diversidade na produtividade de equipes ágeis: um estudo na indústria. In Anais do VII Workshop sobre Aspectos Sociais, Humanos e Econômicos de Software, pages 31–40, Porto Alegre, RS, Brasil. SBC.

Palomba, F., Andrew Tamburri, D., Arcelli Fontana, F., Oliveto, R., Zaidman, A., e Serebrenik, A. (2021). Beyond technical aspects: How do community smells influence the intensity of code smells? IEEE Transactions on Software Engineering, 47(1):108–129.

Pfeiffer, R.-H. (2021). Identifying critical projects via pagerank and truck factor. In 2021 IEEE/ACM 18th International Conference on Mining Software Repositories (MSR), pages 41–45.

Politowski, C., Petrillo, F., Ullmann, G. C., e Guéhéneuc, Y.-G. (2021). Game industry problems: An extensive analysis of the gray literature. Information and Software Technology, 134:106538.

Prieto, D. e Nesteriuk, S. (2021). Indie games br: estado da arte das pesquisas sobre jogos independentes no brasil. In Anais Estendidos do XX Simpósio Brasileiro de Jogos e Entretenimento Digital, pages 745–754, Porto Alegre, RS, Brasil. SBC.

Sebo, P. (2023). How well does namsor perform in predicting the country of origin and ethnicity of individuals based on their first and last names? Plos one, 18(11):e0294562.

VGI (2024). Global Indie Games Market Report 2024. <https://vginsights.com/insights/article/global-indie-games-market-report-2024>.
Online; Publicado em 13 de outubro de 2024.

Voria, G., Pentangelo, V., Porta, A. D., Lambiase, S., Catolino, G., Palomba, F., e Ferrucci, F. (2022). Community smell detection and refactoring in slack: The cadocs project. In 2022 IEEE International Conference on Software Maintenance and Evolution (ICSME), pages 469–473.

Zazworka, N., Stapel, K., Knauss, E., Shull, F., Basili, V. R., e Schneider, K. (2010). Are developers complying with the process: an xp study. In Proceedings of the 2010 ACM-IEEE International Symposium on Empirical Software Engineering and Measurement, ESEM '10, New York, NY, USA. Association for Computing Machinery.