

Perfil de Participantes em Grupo de Dúvidas no Telegram: Estudo Inicial

Leonardo Gargano¹, Gabriel F. Pereira², Adriana Vivacqua¹

¹Universidade Federal do Rio de Janeiro (PPGI/UFRJ)–RJ–Brasil

²Instituto de Computação –Universidade Federal Fluminense –Niterói –Brasil

leonardo_gargano@ufrj.br, gabrielferraz@id.uff.br,
avivacqua@dcc.ufrj.br

Resumo. *Muitos interessados em Ciência de Dados usam aplicativos de mensagens como o Telegram para tirar dúvidas, mas o motivo dessa escolha ainda não está claro. Compreender quem são essas pessoas e quais são seus objetivos pode direcionar a uma melhor compreensão dessas comunidades e do uso do aplicativo. Em um estudo exploratório, analisamos um grupo do Telegram focado em dúvidas sobre Ciência de Dados e caracterizamos 473 participantes. Descobrimos que o Telegram é uma importante ferramenta para usuários iniciantes e que muitos deles estão em busca de migração de carreira.*

1. Introdução

Em um mundo altamente conectado, os dispositivos móveis se tornaram companheiros constantes, tornando-se parte da vida diária de grande parcela da população. Plataformas digitais, como redes sociais e aplicativos de trocas de mensagens, estão presentes em mais de 70% das telas iniciais dos smartphones [Angeli e Reis, 2022]. No Brasil¹, os dois principais aplicativos de troca de mensagens são o Whatsapp e o Telegram, largamente utilizados para comunicação e disseminação de conteúdo [Ribas et. al. 2022]. Esses aplicativos rapidamente se tornaram o meio mais frequente de comunicação através de dispositivos móveis, trazendo maior praticidade para a divulgação e acesso a informações das mais variadas fontes [Dehon, 2018].

A nova geração de desenvolvedores é adepta ao uso de mídias sociais e está fortemente ligada aos seus dispositivos, não apenas para a comunicação, mas também para aprender com outras pessoas [Marques e Silva, 2018]. Diversos trabalhos têm utilizado mensagens trocadas no Whatsapp e Telegram como fontes de dados para os mais variados temas como análise de desinformação [Reis e Benevenuto, 2021] ou sistemas de recomendação [Karimpour et. al. 2021; Rezai e Chahooki 2022], entre outros [Resende et. al. 2018]. Entretanto, encontramos poucos trabalhos que analisassem os aplicativos de troca de mensagens enquanto grupos de discussão para dúvidas. Identificando esse espaço na literatura, esse trabalho busca apresentar resultados iniciais de uma análise de comunidades do Telegram como espaços para debate do tema de Ciência de Dados no público brasileiro.

¹ <https://www.mobilettime.com.br/pesquisas/uso-de-apps-no-brasil-maio-de-2023/>

Com o reconhecimento das empresas a respeito do valor dos dados em diversos temas, como otimização de operações ou apoio na tomada de decisões, o tema Ciência de Dados se tornou popular nos últimos anos [Kim et. al. 2018]. Davenport e Patil afirmam que o cientista de dados seria "o trabalho mais sexy do século 21" [2012]. Neste trabalho buscamos um entendimento inicial da população dos grupos de dúvidas em Ciência de Dados do Telegram e de suas motivações.

Neste trabalho buscou-se investigar as motivações dos desenvolvedores em usar os grupos do Telegram, dado que as pessoas que compartilham interesses em comum têm buscado cada vez mais utilizar aplicativos de mensagens instantâneas para aprender, tirar dúvidas, esclarecer questionamentos e colaborar de forma conjunta. As comunidades presentes nessas plataformas têm se destacado como um espaço propício para a reunião dessas pessoas, proporcionando interações significativas. Diante disso, torna-se relevante investigar como esses grupos são utilizados, e para isso é necessária a identificação dos perfis de usuários presentes nos mesmos.

2. Trabalhos Relacionados

O trabalho de Bougie et. al. [2011] conduziu uma análise de mais de 11.000 tweets de desenvolvedores e descobriu que esses tweets envolvem conversas e compartilham informações com mais frequência do que outros trabalhos com populações mais generalizadas. No trabalho de Shayegan e Valizadeh [2022], os autores buscaram identificar automaticamente as personalidades dos usuários com base no seu comportamento no Telegram, para isso foram coletadas as mensagens dos usuários do grupo Telegram e em seguida, foram identificados os traços de personalidade de cada membro de acordo com o NEO Personality Inventory (NEO PI-R).

O trabalho de Nam et. al. [2009] analisou as características e o comportamento de usuários na maior comunidade online de perguntas e respostas da Coreia do Sul, Naver Knowledge-iN. Os autores coletaram 2,6 milhões de pares de perguntas e respostas em cinco anos e entrevistaram vinte e seis usuários para obter insights sobre suas motivações, funções, uso e experiência.

3. Metodologia

Para caracterizar a população de grupos de dúvidas em Ciência de Dados de aplicativos de mensagens instantâneas, a pesquisa foi desenhada como um estudo de caso único e exploratório. Tal estudo compreende a investigação empírica de um fenômeno contemporâneo dentro de seu contexto real, especialmente quando os limites entre o fenômeno e o contexto não estão totalmente definidos. De acordo com Yin [2001], o estudo de caso é apropriado para questões de pesquisa do tipo “como” e “por que”.

Baseado em Walsham [2006], para esse trabalho foi realizado um estudo aprofundado em que o primeiro autor se envolveu com o objeto de estudo por um determinado intervalo de tempo. A pesquisa acompanhou o histórico de mensagens desde o ano de 2017 até o início de 2023. Nesse período um total de 163.000 mensagens foram coletadas com o objetivo de melhor compreender o cenário social de uso da tecnologia [Walsham, 2006]. A análise das mensagens foi feita em duas etapas: Inicialmente foi criado um script em python para selecionar as mensagens que poderiam ter a descrição do usuário, após esse filtro, foi feita uma análise manual das mensagens. Questões norteadoras:

1. Quem são os participantes de grupos de dúvidas sobre Ciência de Dados (demografia, formação, área de trabalho)
2. Quais seus objetivos ao participar do grupo

Optamos por utilizar o Telegram na escolha do grupo, pela maior facilidade de busca dos grupos, o WhatsApp não disponibiliza ferramentas para busca de grupos públicos (tal como ocorre no Telegram) e pela maior facilidade em obter o histórico das mensagens.

Para a escolha do grupo de estudo foram utilizados os seguintes critérios: ser um grupo público, ter sido criado antes da pandemia (mais de 3 anos de existência) possuir mais de 50.000 mensagens em seu histórico, possuir troca de mensagens diárias, além de ser um grupo criado por brasileiros e para brasileiros.

O grupo selecionado para este estudo contava com quase 5.000 participantes até a data de análise, março de 2023. Trata-se de um grupo aberto, que pode ser facilmente encontrado na barra de pesquisa do aplicativo, por meio de um link público ou através da divulgação feita pelos participantes em outros grupos e em redes sociais.

Nestes espaços os participantes podem fazer perguntas, esclarecer dúvidas, realizar contribuições, refutar outros participantes, entre outras ações. Tudo ocorre de forma assíncrona e as mensagens ficam à disposição para acesso de todos os participantes.

Para o grupo escolhido, foram coletadas mais de 160.000 mensagens enviadas entre 2017 e 2023, que foram analisadas para caracterizar os participantes do mesmo. A análise foi feita a partir das informações de autodescrição dos participantes e dividida em dois níveis. No primeiro nível foi realizada uma separação por formação e, no segundo nível, foram analisadas as mensagens para verificar quais os tipos de dúvidas mais comuns em cada grupo. Com as informações expostas, se tornou possível capturar os cargos, grau de escolaridade e área de formação dos participantes. Um total de 473 indivíduos formaram a base da caracterização.

4. Resultado da Análise

De um universo de quase 5.000 participantes, conseguimos caracterizar 473 através das análises de mensagens introdutórias. Esses usuários expuseram suas áreas de estudo e/ou atuação, além de seus níveis de conhecimento e, a partir desses, foram coletadas as informações necessárias. A caracterização foi possível devido às informações fornecidas pelos próprios usuários, geralmente antecedendo as suas primeiras perguntas. A Figura 1 ilustra o total de participantes.

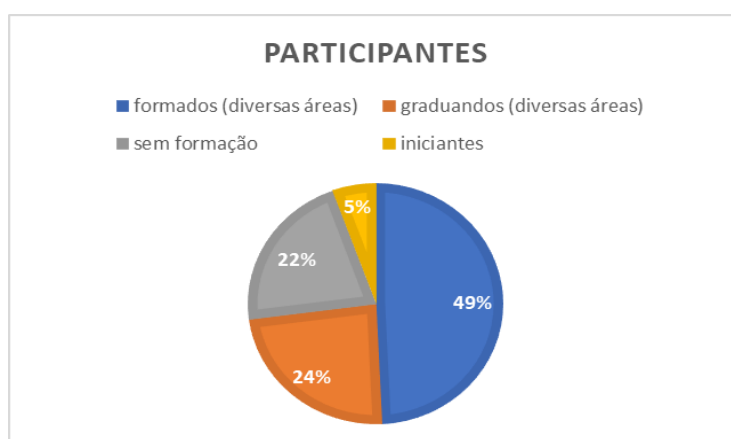


Figura 1. Participantes

Dos 473 indivíduos, 233 participantes possuem graduação concluída em formações distintas, como: Farmácia, Física, Engenharias, Tecnologia da Informação (TI), entre outros. Deste subgrupo de 233, há 30 participantes trabalhando na área de dados e 11 participantes que cursam doutorado nos cursos de Computação, Física e Biologia. Dos que cursam doutorado, há apenas 1 que atua concomitantemente trabalhando na área de dados. Ainda considerando este universo com graduação concluída, foram identificados outros 11 integrantes que já concluíram o doutorado, nos seguintes cursos: Física, Geografia, Música, Bioinformática e Economia. Destes, apenas 2 estão atualmente trabalhando na área de dados. Continuando a análise dos graduados, foram identificados 31 que cursam mestrado em diferentes áreas (desde Computação a Telemedicina), os cursos também são bem variados desde: Oceanografia, Biologia, e Ciência de Dados. Destes mestrandos, apenas 2 participantes atuam em paralelo no trabalho com a área de dados. Complementando a análise deste primeiro grupo, também foram identificados 32 integrantes com mestrado concluído. Destes mestres, apenas 3 indivíduos atuam trabalhando na área de dados. Por fim, dos graduados há 28 participantes cursando pós-graduação/MBA, principalmente em Ciência de Dados. Destes, apenas 3 indivíduos atuam trabalhando na área de dados. Também foram identificados 14 participantes com a pós concluída, novamente com cursos para Ciência de Dados predominando. Destes, há 2 participantes trabalhando na área de dados.

Há 112 participantes cursando alguma graduação. Os cursos se concentram nas áreas exatas, mas aparecem outros, como Biologia e Biblioteconomia. Destes, apenas 4 participantes trabalham na área de dados.

Foram identificados 102 participantes sem nenhuma formação mencionada, mas com as descrições dos seus cargos: 20 deles trabalham na área de dados; 52 em outras áreas de TI (front-end, dba, infraestrutura entre outros); 32 trabalham nas mais diversas áreas como marketing, mercado financeiro e até funcionalismo público. E por fim, 26 participantes do grupo que se declararam iniciantes, com nenhum deles trabalhando na área.

Dos participantes sem a formação descrita, grande parte está buscando uma migração da sua área de atuação para a área de dados. Os mesmos buscam se aprofundar em cursos e bibliotecas relacionadas à Ciência de Dados. Os exemplos abaixo reforçam as afirmações:

☞ *“Boa noite a todos, iniciando na área agora... fazendo migração de carreira.” [P60]*

☞ *“Tô querendo migrar pra área de Data Science, já tenho conhecimentos básicos em programação e também entendo um pouco de Python. Alguém pode indicar algum material?” [P43]*

Os já formados manifestam interesse na migração de carreiras para ciência de dados, além da necessidade de conhecimento de programação para algumas ocasiões no trabalho ou em projetos. Outra ocorrência era a busca pelo acréscimo de skills no currículo.

☞ *“Opa glr. Sou jornalista, entrei no grupo pra poder aprender a programar em python e utilizar a linguagem em análise de dados” [P367]*

☞ *“Boa noite pessoal! Sou analista de performance e estou querendo entrar no Data Science. Alguém indica algum bom curso com certificado?” P[184]*

Devido à diversidade de cursos de formação, as dúvidas dos graduandos foram relacionadas aos primeiros passos que devem ser dados, conselhos sobre como começar os estudos, bases para criar conhecimento, cursos e materiais que ensinem as ferramentas mais utilizadas no momento, dúvidas de programação, além de questionamentos sobre o mercado de trabalho no presente e futuro.

☞ *“Boa tarde pessoal, sou estudante de economia, mas tenho me interessado pela Data Science após uma boa conversa com amigos. Por onde posso começar meus estudos ?” [P400]*

☞ *“Olá, galera. Sou estudante de engenharia eletrônica e me interessei por ciência de dados. Gostaria de saber se vocês tem dicas de cursos e por onde começar nessa área?” [P245]*

Com os pós-graduandos é visível um padrão dentre os integrantes desse âmbito a respeito do conhecimento ao lidar com dados e, em alguns casos, até mesmo das ferramentas computacionais, há também procura de cursos para aprofundar conhecimentos, busca por materiais e dúvidas em linguagens de programação.

☞ *“Alguém conhece um bom livro de estatística com exemplos em Python aí? pode ser um material na web também. Obs não vale documentação do Scipy e Statsmodels rrsrs” “[P78]*

☞ *“Boa tarde, galera vcs recomendam a utilização do fastapi ou do flask para publicar os modelos?”[P17]*

Os mestrandos e doutorandos apresentam um padrão de dúvidas predominantemente apresentadas sobre conceitos, ferramentas e nos casos de áreas que não são relacionadas a TI, ocorrem dúvidas relacionadas à linguagem de programação.

☞ *“Eu tbm sou da Geografia, estou terminando mestrado e estou estudando Python e ciência de dados. Estou incluindo alguns desses conhecimentos na minha pesquisa de mestrado. [P300]*

☞ *“Boa tarde, alguém tem livros ou apostilas sobre Python para Data Science, ML, DL e IA? Estou precisando pro meu mestrado.” [P309]*

5. Discussão

A larga adoção de ferramentas de mensagens e aumento de funcionalidades presentes nestas, tem levado a sua utilização para novos objetivos. A proposta inicial de troca de mensagens diretas entre conhecidos agora se estende a conversas em grupos, muitas vezes temáticos onde nem sempre os participantes se conhecem. Grupos abertos vêm sendo adotados por desenvolvedores para sanar dúvidas e discutir soluções.

Cabe ressaltar que esse aprendizado e socialização do conhecimento sempre existiu, na pergunta ao colega mais experiente e, mais abrangente, em fóruns de discussão (ex: reddit, stackoverflow). Se antes essa necessidade de socialização da informação era suprida por fóruns na web, hoje ferramentas de mensagens vêm ganhando espaço, juntamente com outros recursos.

O compartilhamento de conhecimento por meio de mensagens pode ser uma experiência enriquecedora e útil para diversos tipos de usuários. Mesmo para aqueles que apresentam um bom nível de experiência no assunto em questão, a troca de informações e abordagens diferentes pode oferecer novas perspectivas e soluções mais eficientes para determinadas tarefas. Além disso, para os usuários menos experientes, acompanhar as discussões e dúvidas de outros participantes pode proporcionar aprendizado e aquisição de novos conhecimentos. O ambiente colaborativo dos grupos de mensagens permite uma interação dinâmica, proporcionando uma fonte constante de aprendizado e oportunidades de melhorar as habilidades relacionadas à ciência de dados. Dessa forma, é importante conhecer os perfis desses usuários, para melhor projetar tecnologias que apoiem o seu trabalho.

As redes sociais e aplicativos de mensagens atraem usuários de diversas faixas etárias devido à sua facilidade de uso e flexibilidade na comunicação com outras pessoas [Rezai e Chahooki, 2022]. O presente trabalho reforça a afirmação dos autores ao demonstrar que os participantes do grupo de dúvidas são das mais diversas idades, desde vestibulandos a doutores, de recém-formados a profissionais com vários anos de experiência.

No contexto de aprendizado, o Telegram pode ser especialmente útil para conectar especialistas em determinado assunto com iniciantes que desejam aprender e se aprofundar nesse tema. Os grupos de estudo e discussão no Telegram proporcionam um ambiente virtual propício para compartilhar informações, tirar dúvidas e colaborar na resolução de problemas.

As atividades nos aplicativos de troca de mensagens em dispositivos móveis estão disponíveis 24 horas por dia, sete dias por semana, proporcionando aos usuários acesso a qualquer momento. Essa disponibilidade contínua permite que os usuários se beneficiem da flexibilidade dos horários de envio de mensagens e da capacidade de estudar e visualizar o conteúdo em qualquer lugar, já que os smartphones estão sempre ao alcance das mãos [Martins e Gouveia, 2019].

Um ponto importante que deve ser considerado é a prática autodidata dos participantes, onde muitos buscam sanar suas dúvidas e aprender não apenas para a resolução de um trabalho acadêmico ou uma tarefa do trabalho, mas para aprendizado e

mudança de carreira. Além disso, o Telegram permite que especialistas se comuniquem entre si, possibilitando a troca de conhecimentos e experiências, a discussão de tópicos relevantes e o compartilhamento de recursos valiosos. Isso cria oportunidades de aprendizado contínuo e enriquecimento mútuo.

Nos próximos passos desta pesquisa, investigaremos a motivação para adoção de diferentes ferramentas, e a organização interna dos grupos de mensagens, buscando um melhor entendimento das estruturas que os sustentam.

6. Conclusão e Trabalhos Futuros

O avanço tecnológico e a ampla adoção de smartphones têm contribuído para a popularização e intensificação do uso de aplicativos de trocas de mensagens. Esses aplicativos oferecem uma forma rápida, conveniente e acessível de se comunicar, permitindo que as pessoas troquem mensagens em tempo real, independentemente da sua localização geográfica.

Com o uso generalizado de smartphones, as pessoas possuem esses aplicativos sempre à mão, o que facilita a comunicação instantânea em qualquer momento e em qualquer lugar. Além disso, os aplicativos de trocas de mensagens evoluíram para oferecer recursos avançados, como chamadas de voz e vídeo, compartilhamento de mídia, criptografia de ponta a ponta e integração com outros serviços e aplicativos.

O aplicativo Telegram oferece recursos que facilitam a ação comunicativa e a colaboração entre diferentes usuários, sejam eles especialistas, iniciantes ou ambos. Essa plataforma permite a troca de mensagens de texto, áudio, imagens, vídeos e arquivos, o que possibilita uma comunicação rica e diversificada.

Storey et. al [2014] em seu trabalho já sinalizava que as mensagens instantâneas são usadas por muitos desenvolvedores para apoiar discussões e trocar conhecimentos. Este trabalho buscou caracterizar quem são os desenvolvedores e quais suas motivações que os levaram a utilizar os grupos de dúvidas do Telegram. Como trabalhos futuros, buscamos entrevistar esses participantes e aprofundar o estudo. Também é necessário caracterizar outros grupos para comparações se os perfis dos participantes variam nos grupos.

Referências

- Angeli, P. and Reis, J. (2022). Analyzing the Potential of Feature Groups for Misinformation Detection in WhatsApp. In Anais Estendidos do XXVIII Simpósio Brasileiro de Sistemas Multimídia e Web, novembro 07, 2022, Curitiba, Brasil. SBC, Porto Alegre, Brasil, 45-48. DOI: https://doi.org/10.5753/webmedia_estendido.2022.227031.
- Ribas, L. R.; Gomes-Jr, L.; Silva, T. H. (2020). Rechat: Ferramenta para Estudo do Comportamento de Usuários em Sistemas de Bate-papo do Estilo WhatsApp. In: WORKSHOP DE FERRAMENTAS E APLICAÇÕES - SIMPÓSIO BRASILEIRO DE SISTEMAS MULTIMÍDIA E WEB (WEBMEDIA), 26. , 2020, São Luís. Anais [...]. Porto Alegre: Sociedade Brasileira de Computação, 2020 . p. 79-83. ISSN 2596-1683. DOI: https://doi.org/10.5753/webmedia_estendido.2020.13067.
- Dehon, P., Silva, A., Inocêncio, A., Castro, C., Costa, H., and Júnior, P. (2018). Cvchatbot: Um chatbot para o aplicativo facebook messenger integrado ao ava

- moodle. Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação - SBIE), 29(1):1623.
- Marques, J. B. da S.; De Godoi e Silva, K. A. (2021). Aplicativos de Mensagens Instantâneas em Contextos de Aprendizagem. *Revista de Ensino, Educação e Ciências Humanas*, [S. l.], v. 22, n. 1, p. 38–42, DOI: 10.17921/2447-8733.2021v22n1p38-42. Disponível em: <https://revistaensinoeducacao.pgsscogna.com.br/ensino/article/view/8896>. Acesso em: 30 jun. 2023.
- Karimpour, D. ; Zare Chahooki , M. A. ; and Hashemi, A. (2021). "Telegram group recommendation based on users' migration," 26th International Computer Conference, Computer Society of Iran (CSICC), Tehran, Iran, 2021, pp. 1-6, doi: 10.1109/CSICC52343.2021.9420581.
- Resende, G.; Messias, J.; Silva, J.; Almeida, J.; Vasconcelos, M. and Benevenuto. F. (2018). "A System for Monitoring Public Political Groups in WhatsApp", in *Anais do XXIV Simpósio Brasileiro de Sistemas Multimídia e Web*, Salvador, pp. 387-390.
- Rezai, M. M. and Chahooki, M. A. Z. (2022). "Group recommendation in Telegram by membership graph analyzing," 27th International Computer Conference, Computer Society of Iran (CSICC), Tehran, Iran, Islamic Republic of, 2022, pp. 1-7, doi: 10.1109/CSICC55295.2022.9780527.
- G. Bougie, J. Starke, M.-A. Storey, and D. M. German (2011). Towards understanding twitter use in software engineering: preliminary findings, ongoing challenges and future questions. In *Proc. of the 2nd Int. Workshop on Web 2.0 for Software Engineering, Web2SE '11*, pages 31–36, New York, NY, USA, 2011. ACM
- Shayegan, M. J. and Valizadeh, M. (2022). "A Method for Identifying Personality Traits in Telegram," 2022 8th International Conference on Web Research (ICWR), Tehran, Iran, Islamic Republic of, pp. 88-93, doi: 10.1109/ICWR54782.2022.9786253.
- Nam, K. K; Ackerman, M. S. and Adamic, L. A. (2009). Questions in, knowledge in? a study of naver's question answering community. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. Association for Computing Machinery, New York, NY, USA, 779–788. <https://doi.org/10.1145/1518701.1518821>
- Yin, R. (2001), *Estudo de caso: planejamento e métodos*. Bookman, Porto Alegre.
- Walsham, G. (2006), *Doing interpretive research*. *European Journal of Information Systems*, 320-330
- Storey, M. A; Singer, L; Cleary, B; Filho, F. F. and Zagalsky, A. (2014). The (R) Evolution of social media in software engineering. In *Future of Software Engineering Proceedings (FOSE 2014)*. Association for Computing Machinery, New York, NY, USA, 100–116 <https://doi.org/10.1145/2593882.2593887>
- Martins, E.; Gouveia, L. (2019). Uso do WhatsApp em Atividades Educativas Extraclasse On-line no Ensino de Programação. In: *WORKSHOP SOBRE EDUCAÇÃO EM COMPUTAÇÃO (WEI)*, 27., Belém. *Anais [...]*. Porto Alegre: Sociedade Brasileira de Computação, 2019. p. 141-150. ISSN 2595-6175. DOI: <https://doi.org/10.5753/wei.2019.6625>.

- Reis, J. C. S. and Benevenuto F. (2021). Supervised Learning for Misinformation Detection in WhatsApp. In Proceedings of the Brazilian Symposium on Multimedia and the Web (WebMedia '21). Association for Computing Machinery, New York, NY, USA, 245–252. <https://doi.org/10.1145/3470482.3479641>
- Kim M. , Zimmermann T. , DeLine R. and Begel A. (2018). "Data Scientists in Software Teams: State of the Art and Challenges," in IEEE Transactions on Software Engineering, vol. 44, no. 11, pp. 1024-1038, 1 Nov. (2018), doi: 10.1109/TSE.2017.2754374.
- Davenport T. H. and Patil D. (2012). “Data scientist: The sexiest job of the 21st century,” Harvard Business Review, pp. 70–76 .