

# **Proposta de Trabalho para Edital nº 09/2020 Prevenção e Combate a Surtos, Endemias, Epidemias e Pandemias**

**Programa Pós Graduação em Informática**

**Aluno: Silas Pereira Lima Filho**

**Orientadoras: Jonice de Oliveira Sampaio, Mônica Ferreira da Silva**

**Linha de Pesquisa: Monitoramento da Saúde Mental em População Afetada pelo COVID-19**

## **Introdução**

A depressão é uma das doenças mentais mais relatadas no mundo. Algumas pessoas chamam de doença do século devido ao seu risco<sup>1</sup>. A Organização *Global Burden Disease* indica os transtornos depressivos como a terceira principal causa de incapacidade [1]. Desde 1990, foi a quarta causa principal. Transtornos depressivos são indicados como a terceira principal causa de incapacidade<sup>2</sup>, como confirmado também por [2]. No entanto, em contraste com essa estabilidade, a depressão afeta certos grupos mais do que outros. Como mulheres para homens, europeus para africanos e pessoas com mais renda. É apresentado no Apêndice a Figura 4, demonstrando essa estabilidade na evolução dos casos nos diferentes continentes.

A Organização Mundial de Saúde (OMS) apresenta que cerca de 300 milhões de pessoas de diferentes idades sofrem de algum nível de depressão<sup>3</sup>. Alguns desses sintomas são, por exemplo, humor deprimido na maior parte do dia, perda de interesse em atividades regulares, perda de peso e insônia. O Ministério da Saúde no Brasil apresenta que 11,5 milhões de pessoas são afetadas pela depressão<sup>4</sup>.

É um desafio identificar pessoas doentes na fase inicial da depressão. Alguém depressivo pode enfrentar impedimentos como custo, preconceito social e até uma obstrução pessoal. Além disso, o caminho inverso também é um problema. Devido ao grande número de ocorrências de depressão, pode ser um obstáculo para institutos e profissionais alcançar aqueles que enfrentam uma doença mental e suas variantes. [3] destaca a urgência na identificação e previsão precoces da depressão e seus sintomas devido às dificuldades em detectar esses sintomas nos estágios iniciais. [4] apresenta que 34% da pesquisa em saúde é feita nas mídias sociais e 59% dos adultos procuram informações sobre saúde na Internet. Às vezes, devido à localização ou cenário econômico, o custo em obter ajuda de um profissional impossibilita alguém a procurar ajuda especializada.

O objetivo desse documento é apresentar informações e fatos que embasam a proposta de pesquisa que visa utilizar dados de mídias sociais para conseguir mensurar a saúde mental de pessoas usuárias de mídias sociais.

## **COVID-19 e suas Implicações**

O ano de 2020 é marcado pelo impacto do coronavírus na população mundial. O então definido pela comunidade científica, COVID-19, fez com que distintos grupos, economias e países adotassem o isolamento social como forma de controle da propagação da doença. Dado que sua tal doença é de alto contágio, atualmente, segundo o instituto John Hopkins, mais de 3 milhões de pessoas foram contaminadas e mais de 100 mil pessoas faleceram devido à doença.<sup>5</sup>

Por conta da fácil propagação do COVID-19, os diferentes níveis de estado (Federal, Estado e Municipal) adotaram medidas para controle da doença. O que acarretou em uma adaptação da população criando então novos comportamentos. Comportamentos relacionados ao trabalho, formas de consumo, modos de relacionamento e outros tipos de atividades foram modificados por conta dos novos hábitos de comportamento delegados à população. Tais hábitos rodeiam o isolamento de pessoas e famílias.

<sup>1</sup> [www.theguardian.com/news/2018/jun/04/what-is-depression-and-why-is-it-rising](http://www.theguardian.com/news/2018/jun/04/what-is-depression-and-why-is-it-rising)

<sup>2</sup> Disponível em <http://ihmeuw.org/51aj>

<sup>3</sup> [www.who.int/en/news-room/fact-sheets/detail/mental-disorders](http://www.who.int/en/news-room/fact-sheets/detail/mental-disorders)

<sup>4</sup> [www.blog.saude.gov.br/index.php/materias-especiais/52516-mais-de-onze-milhoes-de-brasileiros-tem-depressao](http://www.blog.saude.gov.br/index.php/materias-especiais/52516-mais-de-onze-milhoes-de-brasileiros-tem-depressao)

<sup>5</sup> Dados conferidos em 13/08.

As variáveis que envolvem a pandemia ainda são analisadas tanto pela comunidade científica, mas também é debatida pelas pessoas fora dessa comunidade. Na população brasileira, existe a discussão popular sobre a eficácia das medidas adotadas e em muitos momentos as medidas tomadas pelas autoridades governamentais nem sempre se baseiam no conhecimento gerado pela comunidade científica. Em meio a tantos fatores, de informação, conhecimento relacionados a doença, ainda é estudado que tipo de consequências cognitivas e psicológicas são geradas por conta da pandemia do COVID-19. Problemas como ansiedade, depressão vem a tona por conta tanto do medo de contaminação da doença que é facilmente propagável. Uma pessoa pode ter preocupações sobre o risco de contaminar-se, ou então ter preocupações sobre familiares ou pessoas próximas estarem em situação de delicada. O Instituto Brasileiro de Geografia e Estatística (IBGE) aponta que boa parte das empresas brasileiras tiveram um impacto negativo durante o período da pandemia<sup>6</sup>. O que gera um impacto na empregabilidade e consequentemente criando uma preocupação em parte da população.

[5] realiza um estudo utilizando questionários para medir o nível de estresse, ansiedade e depressão da população chinesa no período da pandemia. Os autores informam medidas que podem ser tomadas para minimizar o impacto na saúde mental, e também informa que dos grupos analisados, o que tem maior propensão a ter a saúde mental afetada é o de jovens e estudantes. [6] também mede o impacto na saúde mental em relação ao nível de stress, ansiedade e depressão de grupos da população chinesa no período da pandemia.

Em ambos os trabalhos citados, [5] e [6], os autores sugerem que a identificação prévia dos grupos mais propensos a desenvolverem alguma doença psicológica, pode ajudar na adoção de medidas preventivas. Independentemente da doença destacada nesse documento, o uso inteligente de informação e dados pode ajudar a identificar grupos de risco. Isso permite com que as devidas autoridades possam embasar sua tomada de decisões.

Alguns casos podem ser citados onde a tecnologia foi utilizada como forma de atenuar os problemas correlatos à pandemia. Em [7], os autores medem a eficácia da utilização de mensagens SMS para dar apoio às pessoas e como forma de diminuir o stress. Já os autores de [8] medem as mudanças no comportamento de alunos universitários no período da pandemia. Os autores utilizam questionários para medir o nível de ansiedade e stress dos alunos, mas também medem o uso dos smartphone e.g. número de vezes em que o celular é destravado, distância percorrida etc. os autores analisam se tais medidas possuem correlação com as notícias sobre COVID-19. Ye em [9] lista diversas formas de utilização da computação e informática no domínio da saúde. O autor sugere um framework chamado *Tecnologia da Informação na Saúde* onde são mapeados as aplicações da computação. A Figura 1 exibe tais divisões.

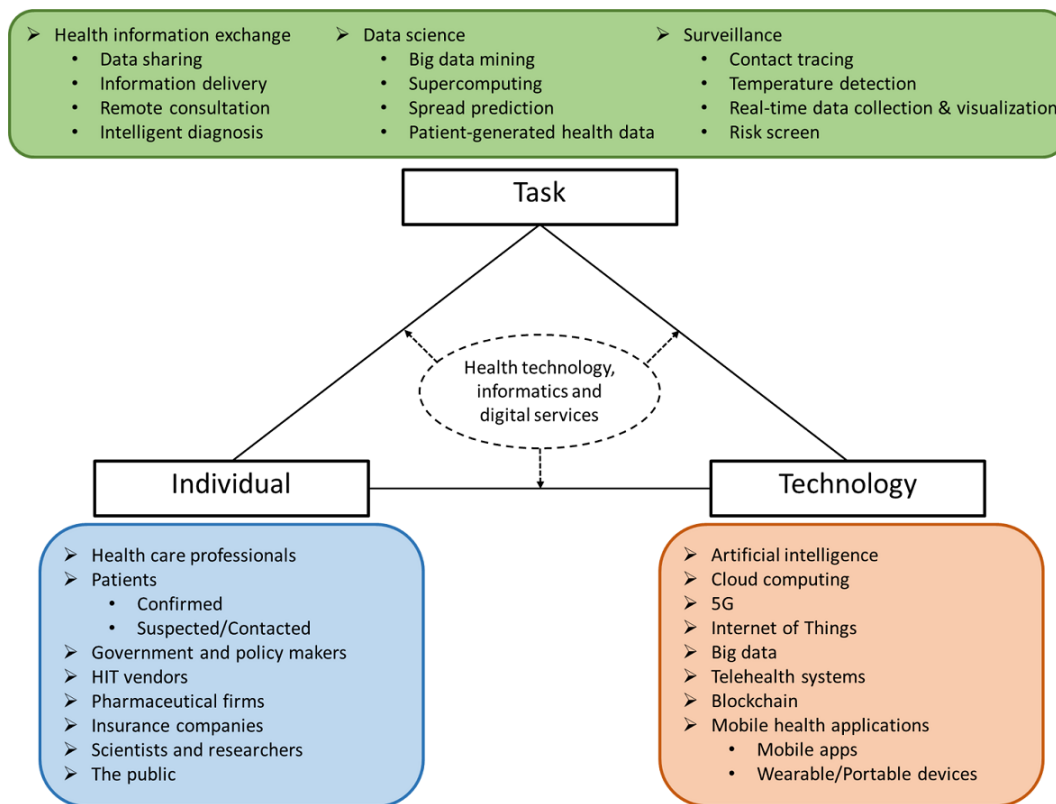
## Monitoramento em Mídias Sociais

*Infodemiologia e Detecção Digital de Doenças* são termos correlatos para descrever o uso de plataformas e ferramentas digitais para melhorar a saúde social. Eles podem ser traduzidos como esforços para combater epidemias, identificar indivíduos em risco e comunicar doenças urgentes de candidatos. O uso da tecnologia apóia diretamente instituições, profissionais e até ajuda as pessoas a se conscientizarem de algumas doenças [10].

Baseado em [9], uma das tarefas listas é a análise de grandes conjuntos de dados. Com a coleta de informações sobre uma população e grupos, cria-se a chance de aproximar os recursos necessários para indivíduos em risco. Uma das tecnologias de coleta, e análise de grandes conjuntos de dados e que podem ser aplicados no domínio da saúde são as mídias sociais. Nelas, seus usuários publicam seus interesses e preferências sociais para então compartilhar com outros usuários. Elas se encaixam nos dois conceitos definidos acima. Podendo assim servir como fonte de informação e monitoramento para autoridades políticas e também para a comunidade científica. Podendo Através da mídia social, um usuário pode se conectar a seus amigos, parentes e até desconhecidos. O conteúdo gerado nessas plataformas se parece com boas fontes de informação que podem ajudar ao lidar com a detecção ou conexão de doenças entre um psicólogo e um paciente depressivo.

---

<sup>6</sup>covid19.ibge.gov.br



**Figure 1:** Framework sugerido por [9].

## Depressão

A 11ª Classificação Internacional de Doenças (CID 11) classifica a depressão como uma doença diagnosticada em relação ao comportamento de alguém. Um evento em que a pessoa perdeu algo, trabalho, alguma pessoa próxima, etc. podem iniciar sintomas de depressão. Esta doença também é perigosa por causa de suas consequências extremas. A depressão, segundo a CID 11, pode levar à ideação suicida e em consequência ao suicídio [11].

Embora depressão seja o nome comum na sociedade, o Manual de Diagnóstico de Transtornos Mentais (DSM-V) detalha diferentes tipos de depressão. O mais comum e mais geral é o transtorno depressivo maior (TDM), embora cada variante da depressão seja coberta pelo termo *Transtorno Depressivo*. As alternativas desse filho do transtorno são *transtorno disruptivo da desregulação do humor*, *Transtorno Depressivo Persistente (distímia)* e *transtorno disfórico pré-menstrual*. O DSM-V também lista as características de diagnóstico de cada variante. Podemos listar critérios de diagnóstico de TDM, por exemplo insônia ou excesso de sono, humor depressivo na maior parte do dia, perda de interesse pelas atividades e perda de peso. O DSM-V também destaca que um grupo de pelo menos 5 sintomas deve ocorrer em um período de duas semanas. Os sintomas e características da depressão são muito semelhantes à descrição de Freud da melancolia [12].

## Dificuldades Tratamento da Depressão

O diagnóstico da patologia se dá geralmente pela terapia, que é feita pela entrevista entre o profissional (psicólogo, psiquiatra ou médico). Atualmente, com a facilidade da tecnologia, profissionais da área da saúde adotaram ao *teleatendimento* para conseguir atender pacientes remotamente. Cada região do cenário do serviço de saúde brasileiro possui suas particularidades. Num mesmo estado é possível encontrar grupos com grande poder aquisitivo, e outros que vivem em situação com recursos limitados. Infelizmente, boa parte da população tem acesso limitado ao serviço de saúde. Muitas vezes por conta de desinformação ou por não conhecimento, não consegue tratamento médico profissional. Outras dificuldades para o tratamento de uma doença é a distância geográfica, ou então a disponibilidade de mão de obra em cidades nem sempre com muitos recursos, ou até

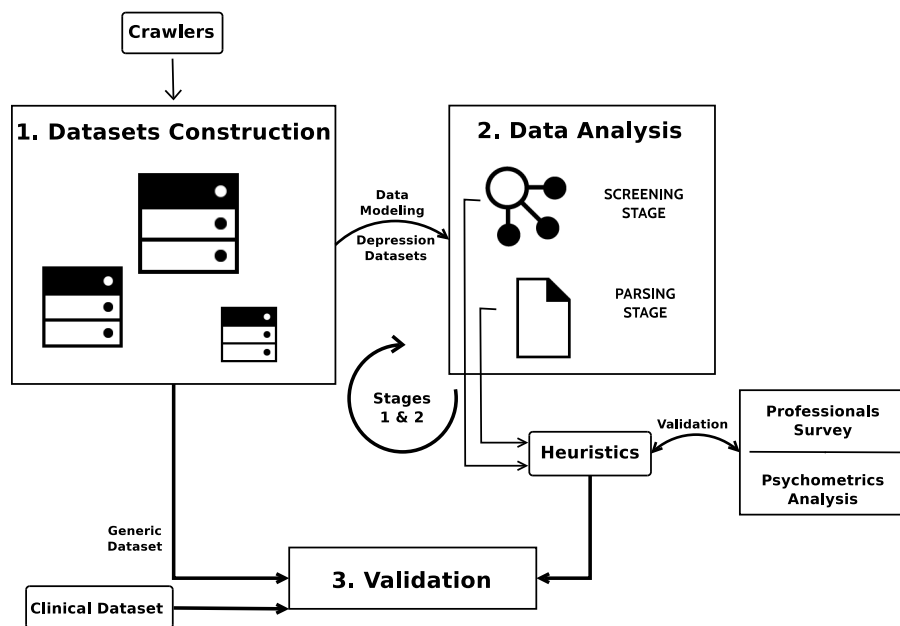
mesmo preconceito de como pessoas próximas encaram uma possível doença psicológica. Todos esses fatores dificultam o trabalho de profissionais.

## Detecção prévia de depressão

Devido ao cenário descrito acima, identificar e atender alguém que possa ser um potencial paciente depressivo, de maneira rápida e discreta, parece ser muito útil tanto para o paciente quanto para o profissional. A tarefa de identificar alguma doença, mesmo que não seja depressão, pode ser desafiadora, mas ao mesmo tempo relevante. Investigar se é possível identificar sinais, sintomas de comportamento depressivo nas plataformas de mídia social. Devido à abundância de dados, pode tornar-se difícil o dado representativo. Portanto, escolha uma técnica confiável e uma análise consistente do método pode exigir uma grande quantidade de pesquisas.

## Proposta de Pesquisa

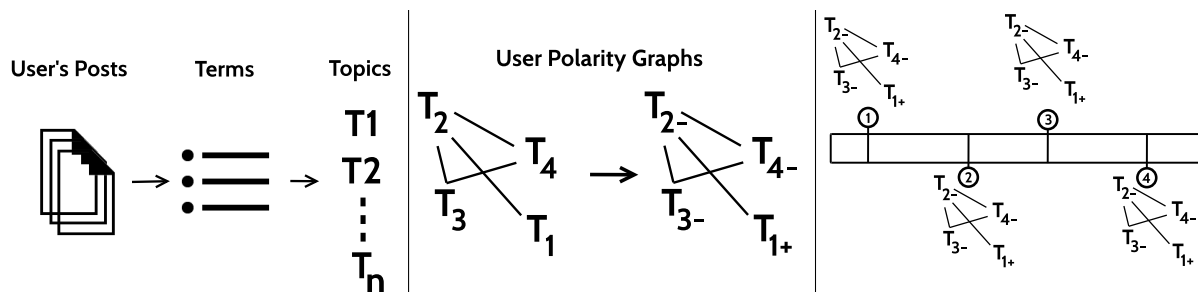
A solução proposta, para a criação de um artefato, enfrenta diferentes etapas que podem oferecer desafios à pesquisa a ser realizada. Já o produto final é a habilidade de identificar sintomas depressivos. Devemos passar por etapas como obter, processar, modelar e persistir os dados de sites de redes sociais em uma estrutura que possa permitir as próximas etapas para recuperar esses dados para análises futuras. Nosso método visa identificar um usuário depressivo inserido em uma rede social de forma confiável, consistente e discreta. A Figura 2 mostra um esquema conceitual de pretendemos organizar as etapas da pesquisa.



**Figure 2:** Esquema conceitual da abordagem proposta.

O Módulo 1 representa a fase em que construímos os conjuntos de dados a serem explorados nos estágios iniciais e posteriores. Este módulo abrange a construção de datasets a serem analisados. Sua construção se dá por meio de crawlers e nos fornece informações sobre o usuário depressivo como seu vocabulário, comportamento de conteúdo e suas preferências sobre relacionamentos com outros usuários. Tais datasets devem possuir distintas proveniências para assim evitar um viés do dataset e consequentemente nas análises realizadas. Atualmente, dois coletores foram desenvolvidos para obter dados dos sites *HealingWell* e *Reddit*. Por se tratar de uma pesquisa acadêmica, não temos a intenção de publicar os nomes dos autores e suas informações pessoais. No momento, temos um conjunto de dados com cerca de 1,5 GB de informações coletadas. O primeiro conjunto de dados está relacionado ao fórum de depressão do site da *HealingWell* <sup>7</sup>. Coletamos 3075 tópicos

<sup>7</sup>[www.healingwell.com/community/default.aspx?f=19](http://www.healingwell.com/community/default.aspx?f=19)



**Figure 3:** As fases da etapa de Análise. Respectivamente, extração dos tópicos do conteúdo de um usuário; Polaridade dos tópicos descobertos; Evolução dos grafos de polaridade ao longo do tempo.

postados e suas respectivas postagens, totalizando 18.450 postagens respondidas.

Posteriormente, a construção dos conjuntos de dados, por meio da modelagem de dados, o Modelo 1 fornece ao Módulo 2 conjuntos de dados de sites onde a depressão é um tópico chave. O segundo módulo compreende a análise dos dados e divide este processo em dois subestágios. Ambos os subestágios pretendem abstrair o conteúdo implícito e explícito do usuário. Os Módulos 1 e 2 são representados como um processo cíclico principalmente devido ao aprimoramento da análise e evolução dos conjuntos de dados em relação ao comprimento, complexidade e informações mais abundantes.

O primeiro subestágio, de triagem, compreende um conjunto de métricas baseadas na análise topológica dos dados. Por meio da análise de redes sociais, podemos entender como as entidades de uma rede social estão conectadas. O estudo da topologia, através do uso da análise de redes sociais, pode levar à escolha de usuários específicos para melhor compreensão deles e sua influência [13]. Podemos listar como exemplos três tipos diferentes de análises relevantes, *Análise de Influência Social*, *Classificação de Nó* e *Deteção de Comunidade* [14]. Com essas análises, seria possível ter um melhor entendimento das conexões em torno de uma pessoa depressiva. O trabalho de Rosenquist et.al. por exemplo consegue mapear dados de questionários psicométricos como grafos (que é a forma de representação de redes sociais). A partir de então os autores conseguem identificar o espalhamento da doença num determinado grupo [15].

O subestágio de análise é dividido em três subprocessos e é representado na Figura 3. O processamento de linguagem natural (NLP) tem mostrado grande avanço nos últimos anos na análise de dados textuais. Modelos de linguagem tem sido utilizados para tarefas como a geração de conhecimento de um determinado domínio, extração de conhecimento em grandes conjuntos de texto, dentre outros. Uma das tarefas de NLP é a identificação automática dos tópicos de um texto. No subestágio de análise, aplicamos então as técnicas de análise textual nos dados coletados, o que se assemelha com outras abordagens científicas. Com base no trabalho de [16], pretendemos identificar quais são os principais tópicos de cada usuário em uma plataforma de mídia social. Após a identificação do tópico, para cada usuário, encontraremos a correlação entre os tópicos e mapearemos os termos em um gráfico de polaridade. O gráfico de polaridade pode ajudar a identificar se as palavras e os termos mais usados por um usuário tendem a ser positivos ou negativos. Trabalhos relacionados mostraram que pessoas depressivas costumam manipular palavras mais negativas. Isso reflete a baixa autoestima desse grupo. Com os gráficos de polaridade de cada usuário, a terceira etapa pretende analisar como esses gráficos evoluíram. Dessa forma, será possível verificar se o discurso de alguém se tornou mais positivo ou negativo. A análise de séries temporais parece ser importante, visto que alguém é considerado deprimido se um conjunto de sintomas já ocorreu por um determinado período de tempo. Isso poderia ajudar a explicar melhor o fenômeno da depressão nas redes sociais.

À medida que as iterações progredirem, o Módulo 2 constrói um conjunto de heurísticas extraídas da análise. Essas heurísticas são testadas por profissionais de pesquisa e comparadas à psicomетria. O conjunto de heurísticas também é validado no Módulo 3. Neste módulo, pretendemos validar as heurísticas não apenas aplicando-as em um conjunto de dados mais genérico e mais amplo, mas também testando em um cenário clínico de psicologia. Dessa forma, pretendemos corroborar a dinâmica e o comportamento da primeira iteração que foi abstraída de conjuntos de dados particulares.

## Contribuições Esperadas

A contribuição final esperada aumentar a área de conhecimento da computação na saúde mental. Devido às amplas implicações da depressão na saúde das pessoas e ao grande uso das mídias sociais, visa-se utilizar os dados coletados com maior precisão e confiabilidade para assim conseguir auxiliar profissionais e autoridades na tomada de decisões. É proposto uma abordagem que difere sobre como lidar com a depressão com a ajuda de SNA e análise de texto. Ressaltamos que a análise de redes sociais compreende a estrutura topológica. Portanto, a principal diferença em relação às pesquisas anteriores e à nossa proposta, é usar o conhecimento e as idéias da psicologia para melhorar o desempenho da classificação com SNA e análise textual.

Com esse experimento, iríamos entender o ambiente onde a doença depressão é o principal tópico de discussão, conseqüentemente, reproduziríamos e validaríamos os resultados e comportamentos das conversas sobre a depressão. A validação é principalmente dirigida à análise textual e de vocabulário. Embora pretendamos explorar profundamente a análise de redes sociais.

As etapas da metodologia proposta As contribuições esperadas estão relacionadas à informática e à psicologia. Para a ciência da computação, esperamos que o emprego de métricas computacionais como análise de redes sociais e classificação de tópicos possa ajudar as pessoas a adquirir uma compreensão mais precisa de como o fenômeno da depressão acontece nas redes sociais e também como as redes sociais refletem a vida real. Do ponto de vista da pesquisa em saúde (psicologia, medicina), nossa abordagem poderia melhorar a forma como o diagnóstico de depressão é realizado. Podendo assim ajudar as pessoas a terem uma saúde melhor por meio do uso da tecnologia.

## Trabalhos Correlatos

Pode não parecer claro, mas podem ser elencados os principais objetivos dos artigos lidos são: identificar quais sintomas são pesquisáveis nas mídias sociais e irão compor um modelo como recursos; criar um modelo que classifique um usuário invisível como potencial depressivo ou não. O problema de lidar com a depressão e as redes sociais pode ser entendido como uma busca por pessoas que sofrem os sintomas da depressão.

Uma boa quantidade de artigos se baseia em técnicas de NLP para fazer uma análise sistêmica do texto em publicações de mídia social. Nem todas as pesquisas analisadas levam em consideração o ponto de vista da psicologia. O efeito de levar em consideração as abordagens existentes da psicologia é que a análise será mais robusta e confiável, uma vez que a área de pesquisa da psicologia já aborda problemas de doenças mentais. É um desafio alinhar a quantificação feita por métricas. NLP, análise de redes sociais e outras técnicas para a cognição de um psicólogo em tratamento clínico comum.

[17] desenvolveu muitos artigos e pesquisas sobre a medição da depressão na população usando informações de mídia social. Os autores deste trabalho fizeram uso de questionários psicométricos. A psicometria representa a teoria e a técnica de mensuração dos processos mentais e é aplicada na psicologia e na educação. É uma abordagem interessante, embora seja questionável por simplificar todo o processo de compreensão do comportamento de alguém. Em [17], uma coleta no estilo crowdsourcing é aplicada para obter dados do Twitter por pessoas que foram clinicamente diagnosticadas com depressão. Com esses dados, eles construíram um corpus e desenvolveram um modelo probabilístico. O modelo treinado classifica se um post indica depressão. Semelhante ao trabalho anterior, Tsugawa et al [18] aplicaram a mesma análise para replicar os resultados em um grupo de usuários do Japão.

[19] apresenta como as atividades no Facebook estão associadas a estados depressivos de usuários a fim de aumentar a conscientização sobre a depressão na universidade onde o estudo foi realizado, que havia visto um aumento na taxa de suicídio de seus alunos. [20] explore postagens de auto-revelação no Instagram. Neste artigo, os autores usaram conteúdo de postagens marcadas com *#depression* para entender quais divulgações bastante confidenciais as pessoas fazem no Instagram. O trabalho in [21] é um estudo qualitativo que busca compreender como se dá o comportamento e a compreensão da população chinesa sobre a depressão. É um estudo qualitativo e difere dos anteriores. [22] conduz um estudo observacional para compreender as interações entre usuários clinicamente deprimidos e sua rede de ego quando comparados com um grupo de usuários sem

depressão. Eles identificam sinais linguísticos e emocionais relevantes de intercâmbios de mídia social para detectar sinais sintomáticos de depressão. [23] aplicou a classificação de texto usando Redes Neurais Convolucionais para classificar a depressão usando análise de texto. [24] também usou redes neurais para identificar padrões em períodos de tempo em que o risco de tentativa de suicídio está aumentando em textos SMS. [25] incorpora análise temporal de conteúdo gerado pelo usuário nas redes sociais para capturar sintomas. Eles desenvolveram um modelo estatístico que emula estudos de coorte observacionais tradicionais conduzidos por meio de questionários online e extraem e categorizam diferentes sintomas de depressão e modelam conteúdo gerado pelo usuário nas mídias sociais. [26] detectou oito emoções básicas e calculou a intensidade geral (pontuação de força) das emoções extraídas de todos os tweets anteriores de cada usuário. Depois disso, eles geraram uma série temporal para cada emoção de cada usuário a fim de gerar uma seleção de estatísticas descritivas para esta série temporal.

Os artigos citados acima nem sempre levam em consideração como os psicólogos inferem se alguém é depressivo ou não. Ressaltamos também que muitas das contribuições reais dependem de informações textuais geradas por um usuário. Uma vez que um dos sintomas de depressão no CID 11 é a inatividade, poderíamos questionar se um sintoma depressivo geraria consistentemente conteúdo online. O contexto da psicologia lida regularmente com a subjetividade da informação. Com base nisso, acreditamos que informações relevantes podem ser extraídas de outros métodos, em vez de conteúdo de texto. Acreditamos que a classificação de potenciais usuários depressivos poderia ser mais confiável se combinada com “informações subjetivas”.

## Cronograma

O período estipulado para a proposta apresentada é de 36 meses. A Tabela 1 apresenta as etapas estipuladas para a pesquisa. Destaca-se que o cronograma apresentado poderá flexibilizar-se de modo a atender as demandas de tempo da chamada de trabalho. Basicamente é intencionado que a pesquisa se desenvolva no principalmente nos dois primeiros anos. Assim, nos últimos doze meses é esperado um trabalho mais focado na divulgação acadêmica e científica dos dados analisados e do conhecimento gerado por meio da pesquisa.

<b>Primeiro Ano - De 09/2020 a 09/2021</b>				
<b>Atividades</b>	<b>Quartis</b>			
	Q.1	Q.2	Q.3	Q.4
Revisão Bibliográfica	X			
Criação de coletores de dados		X		
Qualificação de Tese		X		
Aplicação de Survey com profissionais da saúde			X	
Análise Topológica dos Dados			X	
Análise de Dados Textuais				X
Validação de Análises Realizadas				X
<b>Segundo Ano - De 10/2021 a 10/2022</b>				
<b>Atividades</b>	<b>Quartis</b>			
	Q.1	Q.2	Q.3	Q.4
Atualização da Revisão Bibliográfica	X			
Criação de modelo de classificação inicial	X			
Validação de Performance do Modelo	X			
Validação de Análises Realizadas		X		
Aplicação de survey com pacientes depressivos		X	X	
Análise de dados do survey			X	X
Validação de Heurísticas				
Escrita de Tese	X	X	X	X
Defesa Tese				X
<b>Terceiro Ano - De 10/2022 a 09/2023</b>				
<b>Atividades</b>	<b>Quartis</b>			
	Q.1	Q.2	Q.3	Q.4
Atualização da Revisão Bibliográfica	X			
Submissão para Journal (A definir)	X	X	X	X
Submissão para Congresso (A definir)	X	X	X	X

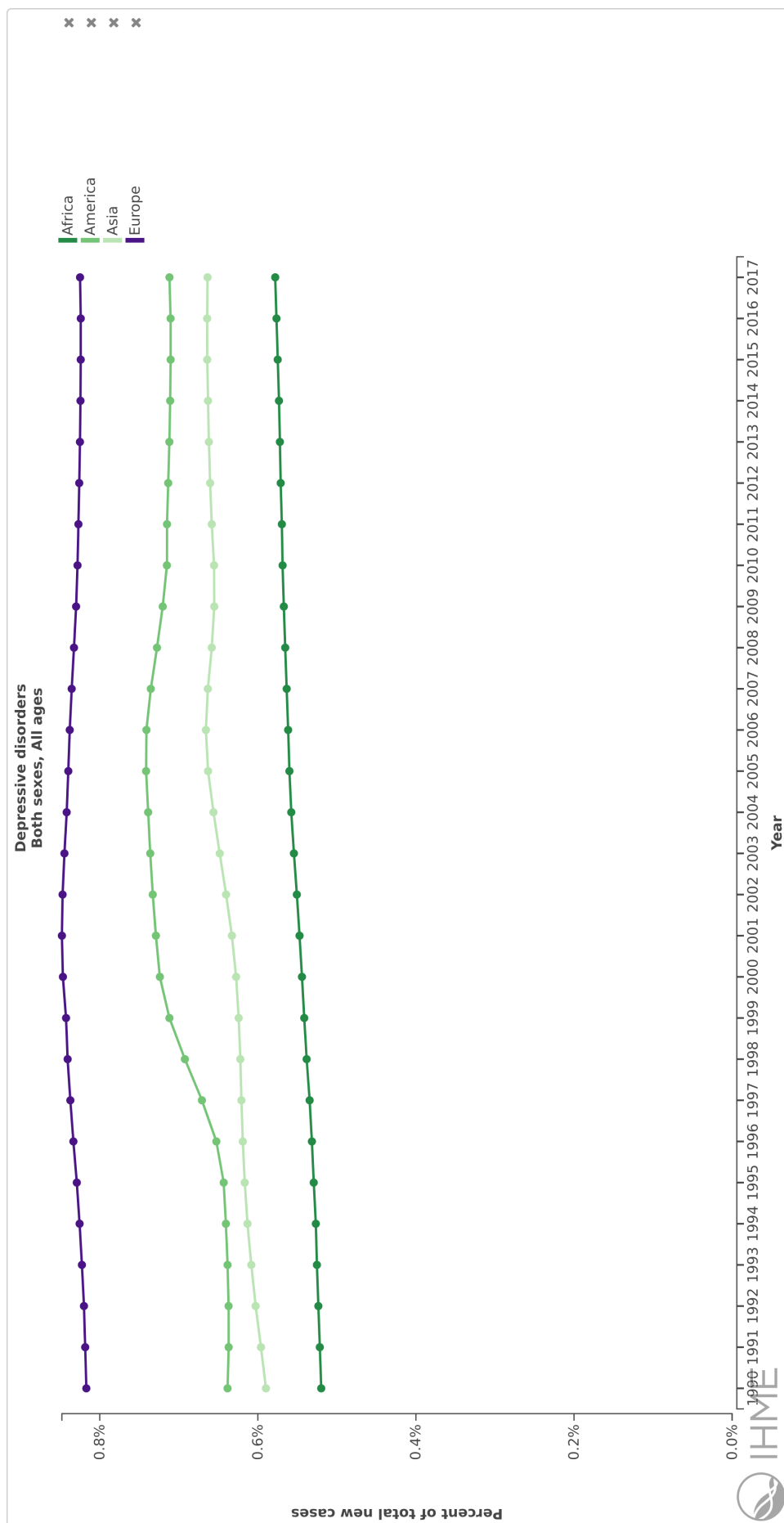
**Table 1:** Cronograma da proposta de pesquisa.



## References

- [1] S. L. James, D. Abate, K. H. Abate, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet*, 392(10159):1789–1858, 2018.
- [2] D. J. Brody, L. A. Pratt, and J. P. Hughes. Prevalence of depression among adults aged 20 and over: United States, 2013-2016. *National Center for Health Statistics*, 2018.
- [3] M. Lech, L.-S. Low, and K. E. Ooi. Detection and prediction of clinical depression. In *Mental Health Informatics*, pages 185–199. Springer, 2014.
- [4] N. Elkin. How America searches: Health and wellness. *Opinion Research Corporation: iCrossing*, pages 1–17, 2008.
- [5] C. Wang, R. Pan, X. Wan, Y. Tan, L. Xu, R. S. McIntyre, F. N. Choo, B. Tran, R. Ho, V. K. Sharma, and C. Ho. A longitudinal study on the mental health of general population during the COVID-19 epidemic in China. *Brain, behavior, and immunity*, 87:40–48, July 2020.
- [6] C. Wang, R. Pan, X. Wan, Y. Tan, L. Xu, C. S. Ho, and R. C. Ho. Immediate psychological responses and associated factors during the initial stage of the 2019 coronavirus disease (COVID-19) epidemic among the general population in China. *International journal of environmental research and public health*, 17(5):1729, 2020.
- [7] V. I. O. Agyapong, M. Hrabok, W. Vuong, A. Gusnowski, R. Shalaby, K. Mrklas, D. Li, L. Urichuk, M. Snaterse, S. Surood, B. Cao, X.-M. Li, R. Greiner, and A. J. Greenshaw. Closing the Psychological Treatment Gap During the COVID-19 Pandemic With a Supportive Text Messaging Program: Protocol for Implementation and Evaluation. *JMIR Res Protoc*, 9(6):e19292, Jun 2020.
- [8] J. F. Huckins, A. W. daSilva, W. Wang, E. Hedlund, C. Rogers, S. K. Nepal, J. Wu, M. Obuchi, E. I. Murphy, M. L. Meyer, D. D. Wagner, P. E. Holtzheimer, and A. T. Campbell. Mental Health and Behavior of College Students During the Early Phases of the COVID-19 Pandemic: Longitudinal Smartphone and Ecological Momentary Assessment Study. *J Med Internet Res*, 22(6):e20185, Jun 2020.
- [9] J. Ye. The Role of Health Technology and Informatics in a Global Public Health Emergency: Practices and Implications From the COVID-19 Pandemic. *JMIR Med Inform*, 8(7):e19866, Jul 2020.
- [10] E. Horvitz and D. Mulligan. Data, privacy, and the greater good. *Science*, 349(6245):253–255, jul 2015.
- [11] A. P. Association et al. *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub, 2013.
- [12] S. Freud. Mourning and Melancholia. U: The Standard Edition of the Complete Psychological Works of Sigmund Freud, Volume XIV (1914-1916): On the History of the Psycho-Analytic Movement, Papers on Metapsychology and Other Works, 1917.
- [13] G. Razis, I. Anagnostopoulos, and S. Zeadally. Modeling Influence with Semantics in Social Networks. *ACM Computing Surveys (CSUR)*, 53(1):1–38, feb 2020.
- [14] C. C. Aggarwal. *Social Network Data Analytics*. Springer Publishing Company, Incorporated, 1st edition, 2011.
- [15] J. N. Rosenquist, J. H. Fowler, and N. A. Christakis. Social network determinants of depression. *Molecular Psychiatry*, 16(3):273–281, 2011. PMID: 20231839.
- [16] D. N. F. Sousa. Automatic Research Areas Identification in C&T. Master’s thesis, 2016.

- [17] M. De Choudhury, S. Counts, and E. Horvitz. Social Media As a Measurement Tool of Depression in Populations. In *Proc. of the 5th Annu. ACM Web Sci. Conf.*, WebSci '13, pages 47–56, New York, NY, USA, 2013. ACM.
- [18] S. Tsugawa, Y. Kikuchi, F. Kishino, K. Nakajima, Y. Itoh, and H. Ohsaki. Recognizing Depression from Twitter Activity. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*, pages 3187–3196, 2015.
- [19] S. Park, I. Kim, S. W. Lee, J. Yoo, B. Jeong, and M. Cha. Manifestation of Depression and Loneliness on Social Networks: A Case Study of Young Adults on Facebook. In *Proc. of the 18th ACM Conf. on Comput. Supported Cooperative Work Social Comput.*, CSCW '15, pages 557–570, New York, NY, USA, 2015. ACM.
- [20] N. Andalibi, P. Ozturk, and A. Forte. Sensitive Self-disclosures, Responses, and Social Support on Instagram: The Case of #Depression. In *Proc. of 2017 ACM Conf. on Comput. Supported Cooperative Work and Social Comput.*, pages 1485–1500, Portland, OR, USA, 2017. ACM.
- [21] G. Li, X. Zhou, T. Lu, J. Yang, and N. Gu. SunForum: Understanding Depression in a Chinese Online Community. In *Proc. of the 19th ACM Conf. on Comput.-Supported Cooperative Work & Social Comput.*, pages 514–525, New York, NY, USA, 2016. ACM Press.
- [22] N. Vedula and S. Parthasarathy. Emotional and Linguistic Cues of Depression from Social Media. *Proceedings of the 2017 International Conference on Digital Health - DH '17*, pages 127–136, 2017.
- [23] X. Zhao, S. Lin, and Z. Huang. Text Classification of Micro-blog's "Tree Hole" Based on Convolutional Neural Network. In *Proc. of the 2018 Int. Conf. on Algorithms, Comput. and Artificial Intelligence*, ACAI 2018, pages 61:1—61:5, New York, NY, USA, 2018. ACM.
- [24] A. L. Nobles, J. J. Glenn, K. Kowsari, B. A. Teachman, and L. E. Barnes. Identification of Imminent Suicide Risk Among Young Adults Using Text Messages. In *Proc. of the 2018 CHI Conf. on Human Factors in Computing Systems*, CHI '18, pages 413:1—413:11, New York, NY, USA, 2018. ACM.
- [25] A. H. Yazdavar, H. S. Al-Olimat, M. Ebrahimi, et al. Semi-supervised approach to monitoring clinical depressive symptoms in social media. In *Proc. of the 2017 IEEE/ACM Int. Conf. on Advances in Social Networks Analysis and Mining 2017*, pages 1191–1198, 2017.
- [26] X. Chen, M. D. Sykora, T. W. Jackson, and S. Elayan. What about Mood Swings. In *Companion Proc. of the The Web Conf. 2018*, pages 1653–1660, New York, New York, USA, 2018. Int. World Wide Web Conf. Steering Committee.



**Figure 4:** Depression progression over the years by global regions