

NOTA METODOLÓGICA

Índice de Transparência da Covid-19

APRESENTAÇÃO

POR QUE É IMPORTANTE?

METODOLOGIA

CRITÉRIOS DE AVALIAÇÃO

SOBRE A OKBR


APRESENTAÇÃO

O **Índice de Transparência da Covid-19** é uma iniciativa da **Open Knowledge Brasil** (OKBR) para avaliar a qualidade dos dados e informações relativos à pandemia do novo coronavírus que têm sido publicados pela União e pelos estados brasileiros em seus portais oficiais. Para superarmos este momento tão desafiador para gestores públicos e população, acreditamos que nenhuma ferramenta é mais poderosa que a **colaboração**, e a **informação** é parte fundamental desse processo.

Ao definirmos um conjunto de dados essenciais e parâmetros para publicação, contribuímos para a padronização da divulgação de dados da pandemia no país, e, desta forma, podemos tanto agilizar o trabalho interno aos órgãos públicos, quanto convidar outros setores da sociedade para construir soluções conjuntamente. **Quando olhamos juntos para os mesmos problemas, somos mais fortes!**

POR QUE É IMPORTANTE?

Gestores públicos de todo o mundo têm corrido contra o tempo para monitorar a crise e implementar políticas eficazes de combate à pandemia. A gestão e divulgação de dados é fundamental nesse esforço. No Brasil, porém, ainda temos uma



heterogeneidade grande na forma de coleta e divulgação de dados, o que pode dificultar a realização de pesquisas e o monitoramento por parte dos próprios gestores e da sociedade como um todo.

Além de contribuir para o aprimoramento dos dados e informações sobre a pandemia no Brasil, o **Índice de Transparência da Covid-19** busca o compartilhamento de métodos de coleta e disponibilização desses dados, que têm sido empregados por gestores públicos de todo o país. Por essa razão, o Índice não deve ser interpretado como uma competição entre entes, mas como um **esforço coletivo de combate ao coronavírus por meio da produção e disponibilização de informações oficiais mais íntegras**.

Embora a construção deste Índice privilegie as definições de dados abertos, ela não exclui a preocupação com a **privacidade e a proteção de dados pessoais** das pessoas atendidas pelo sistema de saúde. A urgência por dados a respeito de medidas relacionadas ao enfrentamento do novo coronavírus não deve significar a flexibilização do direito à privacidade, tampouco estimular a adoção de tecnologias invasivas para coleta de dados pessoais. Todos os dados mantidos por órgãos públicos devem prezar os princípios definidos na Lei Federal n. 13.709/2018, e privilegiar o emprego de técnicas de anonimização às bases para assegurar a transparência das ações governamentais.

As dimensões avaliadas por este Índice não abarcam outras medidas de comunicação à população, como o uso de aplicativos e redes sociais, mas tal escolha não quer dizer que essas iniciativas não sejam valorizadas. O momento exige reforço na capacidade de comunicação e esses meios também têm valor, porém é indispensável que os gestores se atentem a manter todos os canais atualizados da mesma forma, evitando ruídos. A (re)organização dos portais oficiais, com indicações claras de hotspots, painéis e outras seções fixas para tratar de temas relativos ao coronavírus, também é uma medida que facilita a divulgação de informações. Por fim, às administrações que optarem por utilizar aplicativos e redes sociais, vale reforçar: atenção máxima à proteção de dados dos cidadãos.

A intenção da OKBR com a elaboração deste índice não estará satisfeita se o esforço se encerrar nesta publicação. Antes de revelar falhas ou méritos, a principal contribuição do Índice de Transparência da Covid-19 é **apontar caminhos viáveis para que as secretarias estaduais e o governo federal aprimorem a divulgação dos dados**, por meio da publicação padronizada de um conjunto mínimo de informações, definidos

neste documento. No contexto atual, esses dados podem ser determinantes não apenas para o exercício do controle social, mas para **aperfeiçoar políticas que vão salvar vidas**.

Diante deste potencial de impacto tão relevante, a Open Knowledge Brasil espera que os órgãos avaliados possam tomar medidas para contemplar os itens destacados e **se coloca à disposição para ajudá-los a desenvolver políticas de dados abertos e de proteção de dados pessoais**, tanto em relação aos dados da pandemia que ora enfrentamos, quanto para avançarmos nas políticas de transparência quando este desafio estiver superado.

METODOLOGIA


O **Índice de Transparência da Covid-19** é um indicador sintético composto por três dimensões: **Conteúdo**, **Granularidade** e **Formato**. Por sua vez, cada dimensão é constituída por um conjunto de aspectos avaliados separadamente, aos quais são atribuídos diferentes pesos para a construção da nota final. O Índice é representado em uma escala de 0 a 100, em que 0 é atribuído ao ente menos transparente, e 100 ao mais transparente.

Cálculo do Índice

$$\left(\frac{\textit{Conteúdo} + (2 \times \textit{Granularidade}) + (3 \times \textit{Formato})}{\Sigma \textit{Pontos possíveis}} \right) \times 100$$

A coleta dos dados avaliados pelo Índice é baseada nas últimas publicações periódicas de portais oficiais dos órgãos de saúde e controle dos governos estaduais e federal. A análise privilegia seções fixas e hotspots dedicados às informações sobre o novo coronavírus, onde devem ser publicados boletins e informativos epidemiológicos, bem como painéis, relatórios e outros tipos de divulgação de dados.

Embora reconheçamos esforços dos gestores em utilizar redes sociais e aplicativos para incrementar a comunicação com a população, esses meios não foram apreciados para a composição do Índice. Isso porque consideramos que a forma mais acessível e democrática de se compartilhar dados e informações oficiais ainda é por meio de portais oficiais. Sites são meios mais intuitivos e estruturados para buscar



informações, possibilitam o compartilhamento de arquivos, relatórios e análises de forma mais eficaz, além de serem públicos e abertos, no caso dos governamentais. As principais redes sociais, por sua vez, são mantidas por empresas privadas, de modo que muitas delas exigem cadastro, monitoram seus usuários e têm seu código fechado. Esses aspectos são particularmente importantes para pessoas e entidades que estão conduzindo pesquisas sobre a pandemia do novo coronavírus em todo o país e precisam coletar dados com confiança e agilidade.

CRITÉRIOS DE AVALIAÇÃO

As dimensões avaliadas pelo Índice são baseadas nos princípios dos dados abertos, principal tema de atuação da OKBR. Dados abertos são aqueles que podem ser livremente usados, reutilizados e redistribuídos por qualquer pessoa – sujeitos, no máximo, à exigência de atribuição da fonte e ao compartilhamento sob as mesmas regras.

Com base na definição de dados abertos e considerando a realidade dos gestores públicos neste momento de grave crise, o **Índice de Transparência da Covid-19** foi construído como um instrumento baseado em um conjunto de dados que já estão sendo coletados e processados pelas administrações públicas, conforme orientações das autoridades de saúde e órgãos internacionais. Nesse sentido, as dimensões analisadas pelo Índice são as seguintes:

CONTEÚDO

Parâmetros mínimos que podem auxiliar na produção de análises mais aprofundadas sobre a disseminação do novo coronavírus pelo país. Cada indicador recebe **Peso 1** na composição do Índice. São eles:

1. Idade ou Faixa Etária

Informações sobre idade ou faixa etária de pacientes são importantes para que pesquisadores possam realizar pesquisas mais profundas sobre a população afetada pelo vírus, e, assim, descobrir formas mais eficazes para nos tratarmos e evitarmos o contágio.

Tipo de interpretação que os dados permitem: “X casos confirmados na faixa etária A, Y na faixa etária B, Z na faixa etária C”.

2. Sexo

Assim como as informações sobre idade ou faixa etária, dados sobre sexo dos pacientes também são importantes para as pesquisas sobre formas de contágio e tratamento.

Tipo de interpretação que os dados permitem: “X casos confirmados do sexo masculino, Y do sexo feminino”.

3. Status de atendimento

Conhecer os status de atendimento (UTI; isolamento domiciliar; curado etc.) é importante para que possamos compreender a evolução dos quadros de contágio. Aliadas aos dados de características dos pacientes, essas informações nos auxiliam a entender e organizar as demandas de hospitais por leitos, medicamentos e profissionais da saúde.

Tipo de interpretação que os dados permitem: “Dentre os casos confirmados, X pessoas receberam o atendimento A, Y pessoas receberam o atendimento B”.

4. Doenças preexistentes

O novo coronavírus pode se manifestar de forma mais grave em pessoas com comorbidades, se comparado com casos de pessoas que não as têm. Por essa razão, é importante publicar informações sobre doenças preexistentes dos pacientes, de modo a colaborar com a produção de pesquisas sobre esses casos.

Tipo de interpretação que os dados permitem: “Dentre os casos confirmados, X pessoas têm a doença A, Y pessoas têm a doença B”.

5. Ocupação de leitos

Dados sobre ocupação de leitos são fundamentais para a avaliação e para o ajuste da política de atendimento aos casos de Covid-19. Por meio do compartilhamento desse dado, hospitais podem organizar melhor a distribuição de atendimentos, e pacientes podem se encaminhar às unidades mais adequadas a cada situação.

Tipo de interpretação que os dados permitem: “A unidade de saúde/município X possui Y leitos e Z estão ocupados, a unidade de saúde/município A possui B leitos, dos quais C estão ocupados”.

6. Outras doenças respiratórias

É comum que pessoas com suspeitas de Covid-19 sejam identificadas, primeiramente, como portadoras de sintomas comuns a outras doenças respiratórias, como a Síndrome Respiratória Aguda Grave (SRAG). Por esse motivo, a publicação de dados sobre essas doenças é importante tanto para conhecermos melhor o avanço do quadro do novo coronavírus no país, diferenciando-o de outras enfermidades, quanto para o acompanhamento de casos considerados suspeitos – sobretudo em um contexto de falta de insumos para testes.

Tipo de interpretação que os dados permitem: “Dentre os casos atendidos pela rede de saúde do município/estado, X pessoas foram diagnosticadas com a doença A, Y pessoas foram diagnosticadas com a doença B”.

7. Testes disponíveis

Confirmações de casos de Covid-19 têm se baseado principalmente nos resultados de exames laboratoriais. A disponibilização de dados sobre a quantidade de testes disponíveis é importante para se conhecer a infraestrutura que já está preparada e estimar a subnotificação, além de facilitar a distribuição de casos suspeitos no sistema de saúde.


Tipo de interpretação que os dados permitem: “A unidade de saúde/município/estado X possui Y testes de coronavírus disponíveis; a unidade de saúde/município/estado A possui B testes de coronavírus disponíveis”.

8. Testes aplicados

A quantidade de testes aplicados é fundamental para a composição do número de casos confirmados e descartados. Além disso, é um indicador de demanda de testes, contribuindo para o planejamento dos órgãos e unidades de saúde.

Tipo de interpretação que os dados permitem: “A unidade de saúde/município/estado X realizou Y testes, sendo Z positivos para coronavírus e W negativos”.

GRANULARIDADE



A Granularidade é a dimensão que avalia o detalhamento dos dados divulgados pelas autoridades, fornecendo uma camada extra de informação. Cada indicador recebe **Peso 2** na composição do Índice.

9. Microdado

A divulgação anonimizada de dados relativos a cada caso, separadamente, é importante para a construção de um panorama mais preciso sobre o avanço do novo coronavírus no país. Informações sobre sexo, idade, município, possível origem do contágio e tratamento conferido são exemplos de dados que podem ajudar a construir uma boa base de dados de monitoramento.

Novamente, vale ressaltar a obrigatoriedade de anonimização dos dados, isto é, retirar informações que permitam identificar o paciente, como nome, RG, CPF, data de nascimento e o hospital de atendimento. Quando o ente já divulgar dados sobre ocupação de leitos, não é necessário informar a localização dos casos na base de microdados, como medida para evitar a exposição dos pacientes.

Tipo de interpretação que os dados permitem: “A pessoa 1, confirmada com o coronavírus, é do sexo X, idade Y, residente do município Z e recebeu o tratamento W”.

10. Localização

Publicar informações de localização de forma detalhada é importante para a realização de diversos trabalhos durante a pandemia: órgãos públicos que precisam alocar recursos de forma eficiente; grupos e indivíduos que desejam enviar doações; pesquisadores que estão buscando focos potenciais de disseminação do vírus.

Quanto mais detalhamento neste critério, mais cuidados devem ser tomados com a proteção da privacidade dos pacientes. Por essa razão, é recomendável que a publicação ocorra de duas formas: i) quantidades de casos por bairros/distritos somente para as capitais e municípios de médio e grande portes; e ii) quantidades de casos por municípios para cidades de pequeno porte.

Tipo de interpretação que os dados permitem: “X casos foram identificados no município Y, A estão no bairro/distrito B, e C estão no bairro/distrito D”.

FORMATO

Em Formato, avaliamos de que maneira os dados foram disponibilizados. A forma de acessar, ler e compartilhar os dados é um determinante do potencial de análise dos mesmos. Cada indicador recebe **Peso 3** na composição do Índice.

11. Visualização

Painéis facilitam a visualização dos dados, favorecendo uma visão ampla do cenário de disseminação do novo coronavírus, bem como das ações tomadas pelos órgãos responsáveis. São ferramentas importantes tanto para gestores públicos, quanto para outros interessados no tema e o público em geral, por facilitar o acesso à informação.

12. Formato aberto

A disponibilização de bases de dados para *download* é fundamental para compartilhar informações de forma organizada com outros entes governamentais, organizações e indivíduos que possam ajudar no combate ao novo coronavírus. Quando os entes disponibilizam a maior parte das bases apresentadas nos boletins e painéis em formatos editáveis ou abertos, como o .CSV ou o .ODS, facilita-se o acesso por parte de pessoas que não utilizam softwares de edição de planilhas fechados e/ou pagos.

13. Série histórica

Séries históricas são formas de acompanhamento de fenômenos ao longo do tempo. No contexto da pandemia, a publicação de dados que permitam construir séries históricas auxilia nas análises sobre o comportamento do vírus e sobre o sucesso das políticas públicas de contenção, favorecendo tanto a gestão pública, quanto as pesquisas acadêmicas.

RESUMO DOS CRITÉRIOS

Dimensão	Critério	Descrição	Pontuação
Conteúdo (Peso = 1)	Idade ou Faixa Etária	Idade ou faixa etária das pessoas que pertencem ao grupo de, no mínimo, casos confirmados	0 = não apresenta; 0,5 = apresenta apenas uma média geral ou dados de algum grupo (ex. óbitos); 1 = apresenta idade ou faixa etária.

	Sexo	Sexo das pessoas que pertencem ao grupo de, no mínimo, casos confirmados	0 = não apresenta; 0,5 = apresenta para algum grupo (ex. óbitos); 1 = apresenta idade ou faixa etária.
	Status de atendimento	Especifica casos hospitalizados (internação e UTI) ou em isolamento domiciliar	0 = não apresenta; 1 = apresenta.
	Doenças preexistentes	Presença de doenças preexistentes/ comorbidades (diabetes, hipertensão etc.)	0 = não apresenta; 0,5 = apresenta em algum grupo (ex. óbitos) 1 = apresenta para todos os casos
	Ocupação de leitos	Quantidade de leitos ocupados no estado em relação ao total disponível. Pode ser apresentado como taxa (%) ou desagregado (capacidade e ocupação)	0 = não apresenta; 0,5 = apresenta somente leitos exclusivos de Covid-19; 1 = apresenta.
	Outras doenças respiratórias	Número de casos de Síndrome Respiratória Aguda Grave registrados ou outras enfermidades que possam indicar suspeita de Covid-19	0 = não apresenta; 1 = apresenta.
	Testes disponíveis	Quantidade de testes de que o estado dispõe para atender a determinado período de tempo	0 = não apresenta; 1 = apresenta.
	Testes aplicados	Quantidade de testes já realizados, incluindo detalhamento de resultados negativos e positivos.	0 = não apresenta; 1 = apresenta.
Granularidade (Peso = 2)	Microdado	Registros individualizados, detalhados e anonimizados dos casos confirmados.	0 = não apresenta; 0,5 = apresenta de um grupo de casos apenas (ex. óbitos); 1 = apresenta.
	Localização	Nível de agregação geográfica dos casos divulgados	estado = 0; cidade = 0,5; quantidade por hospital ou bairro/distrito = 1
Formato (Peso = 3)	Visualização	Painel para consulta do público em geral	0 = não apresenta; 1 = apresenta.
	Formato aberto	Maioria dos dados estruturados de painéis e boletins em ao menos uma	0 = não apresenta; 1 = apresenta.

		planilha em formato editável, de preferência aberto (CSV, ODS)	
	Série Histórica	Base de dados única e atualizada com o histórico completo de registro de casos do novo coronavírus	0 = não apresenta série; 0,5 = série em gráfico; 1 = série em uma base de dados completa para download.

SOBRE A OKBR

A OKBR, também conhecida como Rede pelo Conhecimento Livre, é uma organização da sociedade civil sem fins lucrativos e apartidária que atua no país desde 2013. Desenvolvemos e incentivamos o uso de tecnologias cívicas e de dados abertos, realizamos análises de políticas públicas e promovemos o conhecimento livre para tornar a relação entre governo e sociedade mais transparente e participativa.

Saiba mais no site: <http://br.okfn.org>

Equipe responsável pelo levantamento:

COORDENAÇÃO GERAL

Fernanda Campagnucci

COLETA E ANÁLISE DE DADOS

Camille Moura

Fernanda Campagnucci

REDAÇÃO E REVISÃO

Murilo Machado

Thiago Teixeira