



TOOLKIT #1 - TRANSPARÊNCIA COVID-19

Publicação de Microdados

Um dos principais desafios para a transparência da Covid-19, a disponibilização de Microdados é o instrumento com maior potencial para qualificar as informações que o Estado pode oferecer a pesquisadores, jornalistas, voluntários e demais protagonistas e interessados no enfrentamento à pandemia.

Aliada às devidas técnicas de anonimização e aos padrões abertos, a publicação de uma base de Microdados detalhada atende aos padrões internacionais de abertura, e permite o desenvolvimento de análises, comparações e projeções mais precisas, complexas e informativas. Além disso, a disponibilização da base pode fazer o ente pontuar em diversos critérios no Índice de Transparência da Covid-19, como os das dimensões Conteúdo e Formato, impulsionando o desempenho no ranking geral.

Nesta publicação, você conhecerá alguns passos para abrir um conjunto de Microdados. O Toolkit de Publicação de Microdados será atualizado com novas orientações conforme identificarmos os avanços dos estados nas avaliações. Não deixe de conferir também [nosso Template](#), que pode tanto te ajudar a organizar os dados que já possui, quanto vislumbrar aqueles que ainda precisa coletar.

ABRINDO E PROTEGENDO DADOS

Poucos estados avançaram na granularidade das informações sobre localização de casos suspeitos e confirmados do novo coronavírus até o momento. Este ponto parece ter suscitado dúvidas em alguns gestores devido ao fato de a publicação detalhada dos dados apresentar potencial para ferir a privacidade dos cidadãos a que os dados se referem. No entanto, também é fato que quanto mais detalhes tivermos sobre

os casos, mais análises de qualidade conseguimos produzir. Como resolver este aparente conflito?

Dados de saúde são considerados dados sensíveis pela Lei Geral de Proteção de Dados Pessoais, e, por isso, é necessário ter o máximo de cautela ao coletar e administrar esses dados. No caso da disponibilização de dados sobre a Covid-19, a solução mais simples para mitigar os riscos à privacidade se encontra nos diferentes **níveis de detalhamento e agregação** que podem ser empregados aos dados coletados. Quando a publicação vincula diretamente atributos de idade e sexo dos pacientes a seus respectivos locais de internação, por exemplo, isso pode facilitar a identificação de um determinado paciente, sobretudo em municípios pequenos. Esse tipo de publicação não é incentivada pelo ITC-19.

No entanto, uma opção para não expor os pacientes e nem deixar de publicar esses dados -- que são todos igualmente importantes -- é divulgar o quantitativo de casos por hospital, e, separadamente, publicar os Microdados de casos com as informações de idade e sexo. Afinal, não é tão importante saber especificamente se um homem de 35 anos está internado no Hospital das Flores, mas é urgente saber quantas pessoas estão sendo atendidas no Hospital das Flores, e é fundamental saber que o vírus está acometendo homens de 35 anos.

Lógica similar se aplica à disponibilização de dados de casos por bairro/distrito em municípios pequenos e médios, que ainda possuem poucas confirmações de Covid-19. Para essas situações, é necessário redobrar o cuidado com a exposição de pacientes, e muitas vezes analisar caso a caso, dado que é comum que muitos municípios se conheçam. Por este motivo, o ideal é publicar informações (qualitativas e quantitativas) de casos por bairro/distrito somente para capitais e grandes cidades.

ENCONTRANDO E ORGANIZANDO DADOS

Sistema e-SUS VE

Em todos os níveis de governo, diversos sistemas e tecnologias têm sido empregados para coletar dados de casos sobre o novo coronavírus. No entanto, o [e-SUS VE](#), sistema criado pelo Governo Federal para registro obrigatório dos casos de novo Covid-19, lança luz sobre quais tipos de dado a administração deve coletar, e, portanto, pode disponibilizar, após aplicar as devidas medidas de anonimização.

Conforme o [manual de navegação](#) do e-SUS VE publicado pelo Ministério da Saúde, gestores devem cadastrar os seguintes dados referentes às notificações de Covid-19:

- CPF
- Nome completo
- Data de nascimento
- Sexo
- Estado de residência
- Município de residência
- Telefone
- Data da notificação
- Sintomas (Dor de garganta, Dispneia, Febre, Tosse)
- Data do início dos sintomas
- Condições (Doenças respiratórias crônicas descompensadas; Doenças cardíacas crônicas; Diabetes; Doenças renais crônicas; Imunossupressão; Gestante de alto risco; Portador de doenças cromossômicas)
- Estado do teste (Solicitado; Coletado; Concluído)
- Data de realização do teste
- Tipo de teste (Teste rápido - Anticorpo; Teste rápido - Antígeno; RT-PCR)
- Resultado do teste (Negativo; Positivo)

Como, segundo o manual do Ministério da Saúde, o e-SUS VE permite a **exportação da planilha CSV** com o compilado de dados inseridos no sistema pela administração, facilita-se a disponibilização de uma base com Microdados. Além disso, o documento também destaca que é **possível baixar a base do e-SUS VE filtrando por CNES**, o que possibilita contabilizar a quantidade de casos por unidade de saúde.

Dentre os dados requeridos para a notificação no sistema do Governo Federal, é possível observar que ao suprimir somente os campos CPF, Nome completo e Telefone, os gestores já podem avançar significativamente na publicação periódica de um conjunto de informações bastante completo.

Recomendamos que sejam disponibilizados dados de **todos os casos notificados**, não somente os de casos confirmados. Isso porque, com esses números, os pesquisadores conseguem, por exemplo, acompanhar a capacidade de testagem dos

entes ou mesmo estimar com mais precisão a quantidade de casos que podem estar fora das estatísticas oficiais de confirmados.

Vale atentar ainda para o fato de o e-SUS VE permitir atualização caso a caso. Gestores que focarem a disponibilização de Microdados somente neste sistema deverão extrair e publicar as bases com frequência constante para permitir o acompanhamento da evolução do número de recuperados e óbitos. É de suma importância que o ente seja capaz de **informar com clareza a data de atualização dos dados publicados, o que pode ser feito adicionando a coluna de "Data de Atualização" na base de dados.** Recomendamos que a administração faça backup diário das bases de dados, uma vez que o sistema sobrescreve as informações.

BAIXE O MODELO DE PADRONIZAÇÃO DOS MICRODADOS

A OKBR produziu um template com sugestão de atributos e formato para padronização da base de microdados. [Acesse aqui.](#)

Código do Workset	Nome do Workset	Formato	Descrição	Exemplo
data_publicacao	Data de publicação da base	Data (DD-MM-AAAA)	Data em que o arquivo CSV foi atualizado. Alterar a cada publicação	-
data_mes	Data de início do mês	Data (DD-MM-AAAA)	Data que se registra a ocorrência dos primeiros casos de Covid-19	-
data_confirmacao	Data de confirmação	Data (DD-MM-AAAA)	Data que se registra a confirmação do caso de Covid-19. Se o caso não é confirmado, deixar em branco	-
data_incidencia	Data de início da doença	Data (DD-MM-AAAA)	Data que o usuário insere sobre a primeira informação sobre o caso	-
data_fim_doenca	Data de término da doença	Data (DD-MM-AAAA)	Data que o usuário insere sobre o fim da doença	-
classificacao	Classificação	Texto	Definição da doença e seu estado atualizado	Covid
endereco_bairro	Cidade/Bairro	Texto	Nome do bairro/cidade onde ocorreu o caso	Luzia
tipo_caso	Tipos de caso	Texto	Padrão de base utilizado para confirmação dos casos	Óbito
evolucao	Evolução	Texto	Evolução do caso até o momento	Cure
estado_cidade	Estado/Cidade	Texto	Nome do Estado e da Cidade onde ocorreu o caso	Paraná
municipio	Município (Nome)	Texto	Nome do Município onde ocorreu o caso	-
codUge	Município (Código)	Texto	Código UGE do Município	-
estado	Estado	Texto	Nome do Estado onde ocorreu o caso	-
municipio_estado	Município e Estado (Nome)	Texto	Nome do Município e do Estado onde ocorreu o caso	-
codUge_estado	Município e Estado (Código)	Texto	Código UGE do Município e do Estado onde ocorreu o caso	-
estado_cidade	Estado e Cidade	Texto	Nome do Estado e da Cidade onde ocorreu o caso	-

Se tiver comentários para aprimoramento ou dúvidas sobre como implementar o modelo, escreva para transparenciacovid19@ok.org.br



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://ok.org.br>.

Contato: transparenciacovid19@ok.org.br