



LAI: DADOS DE ATENDIMENTO NÃO REFLETEM REALIDADE

Novembro/2022

Órgãos federais apresentam alto índice de imprecisão na classificação de pedidos de informação quanto ao atendimento



Este trabalho está sob a licença [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/). Mediante atribuição de crédito à organização autora, pode ser copiado e redistribuído em qualquer suporte ou formato; remixado e adaptado para qualquer fim, inclusive comercial (nestes casos, as alterações feitas devem ser indicadas).

FICHA TÉCNICA

novembro/2022

DIREÇÃO EXECUTIVA

Juliana Sakai

SUPERVISÃO E EDIÇÃO

Marina Iemini Atoji

ANÁLISE DE DADOS

Raul Durlo

PESQUISA E REDAÇÃO

Nathália Mendes, Rebeca Lins

FINANCIAMENTO



FORD
FOUNDATION

Introdução

A Lei de Acesso à Informação (lei nº 12.527/2011) completa 11 anos de sua sanção neste mês. Seu cumprimento integral, especialmente em relação ao atendimento satisfatório dos pedidos de acesso à informação, sempre foi um desafio. A forma mais abrangente de verificar a evolução desse aspecto são os indicadores de respostas positivas informados pelos Poderes e órgãos.

Uma comparação dos indicadores oficiais de atendimento a 48.507 pedidos feitos a órgãos federais com indicadores obtidos por meio de avaliação independente indica que **20% dos pedidos cujas respostas foram classificadas oficialmente por órgãos federais como “acesso concedido” não foram de fato atendidos, ou seja, não forneceram a informação solicitada.** Outros 12% classificados como “acesso concedido” forneceram apenas parte da informação (foram “parcialmente atendidos”). Ou seja, **os dados oficiais mostram uma perspectiva melhor do que a realidade.**

Há indícios de que a proporção oficial de pedidos considerados concedidos nos últimos dois anos está inflada em mais de 20 pontos percentuais. Enquanto o índice oficial de atendimento a pedidos de informação para a amostra analisada é de 83%, o índice de atendimento obtido por meio da avaliação independente é de 61,5%.

Destaca-se o caso da Secretaria Geral da Presidência da República, que qualificou como “acesso concedido” centenas de pedidos aos quais respondeu que não dispunha das supostas provas sobre fraudes nas eleições de 2018 que o presidente Jair Bolsonaro (PL-RJ) alegou ter, em declaração feita [em março de 2020](#) e repetida em outras ocasiões.

A classificação independente foi obtida por meio do uso de um modelo de inteligência artificial que classificou as demandas relativas ao período de 1º de janeiro de 2021 a 7 de agosto de 2022 em três categorias: “atendidos”, “parcialmente atendidos” e “não atendidos”. **É a primeira vez em que a discrepância entre a classificação sinalizada pelo órgão público e o**

atendimento de fato aos pedidos é quantificada em uma grande amostra.

Até o momento, contava-se apenas com análises pontuais e qualitativas.

O modelo utiliza técnicas de processamento de linguagem natural (NLP-BERT, na sigla em inglês) para analisar a estrutura dos textos do pedido e da resposta para deduzir se a informação foi fornecida, ou seja, se o pedido foi atendido. A ferramenta foi treinada com uma base de 10 mil pedidos e respostas classificados manualmente em “atendidos”, “não atendidos” e “parcialmente atendidos”, em análise publicada em 2018¹.

O algoritmo foi aplicado na referida amostra da base de dados de pedidos de acesso à informação, extraída das bases disponibilizadas pela Controladoria Geral da União (CGU). Os resultados dessa análise ensejam contribuições para a categorização adotada pelo órgão, a qual é determinante para avaliar o desempenho da administração pública federal nas políticas de transparência passiva. Entender se determinado pedido de informação foi satisfatoriamente atendido, pelo olhar do cidadão, é essencial para aperfeiçoar a forma como a CGU orienta os órgãos para categorizar as suas respostas e propor medidas de aperfeiçoamento para a referida política pública. Esta análise coopera com esse objetivo ao fornecer uma avaliação alternativa à do governo federal sobre o desempenho do poder público federal na transparência passiva.

Metodologia

O modelo de inteligência artificial aplicado aos pedidos para classificar as respostas quanto ao atendimento foi criado especialmente para o projeto Achados e Pedidos, realizado pela Transparência Brasil em parceria com a Abraji (Associação Brasileira de Jornalismo Investigativo). Ele utiliza técnicas de processamento de linguagem natural (NLP-BERT, na sigla em inglês) para analisar a estrutura dos textos do pedido e da resposta para deduzir se o pedido foi “atendido” (ou seja, a informação foi fornecida), “parcialmente atendido” (só

¹ O que a população quer saber do poder público. Transparência Brasil. São Paulo, 2018. https://www.transparencia.org.br/downloads/publicacoes/RelatorioLAI_TransparenciaBrasil_2018_vf.pdf

parte da informação foi fornecida e o restante foi negado ou não entregue) ou “não atendido” (a resposta não fornece a informação, ou a resposta é negativa).

As três categorias adotadas no modelo NLP-BERT resumem, sob a ótica do usuário ou cidadão, três das sete possibilidades de classificação adotadas pelos órgãos do governo federal no Fala.br. Ou seja, verifica-se simplesmente se a pessoa conseguiu obter a informação que queria no todo ou em parte.

Quadro 1. Correspondência entre classificação oficial e classificação NLP-BERT

Classificação Fala.Br	Classificação correspondente (NLP-BERT)
Acesso Concedido	Atendido
Acesso Parcialmente Concedido	Parcialmente Atendido
Acesso Negado	Não Atendido
Não se trata de solicitação de informação	Excluído da amostra
Informação Inexistente	Excluído da amostra
Órgão não tem competência para responder sobre o assunto	Excluído da amostra
Pergunta Duplicada/Repetida	Excluído da amostra

Fonte: elaboração própria

Nesta primeira versão, o modelo só é capaz de realizar a classificação em pedidos cujas respostas: (i) estavam registradas no próprio texto da resposta e não em arquivos anexos (documentos, planilhas, imagens ou outro tipo de arquivo); (ii) não foram objeto de recurso e (iii) foram classificadas como “Acesso Concedido”, “Acesso Parcialmente Concedido”, “Acesso Negado” pelos órgãos federais.

Os pedidos com recursos não integram a amostra porque ainda não foi possível garantir que o modelo analise apenas aqueles cuja tramitação foi de fato encerrada, conforme os prazos definidos na LAI. Foram usados somente os pedidos cujas respostas foram classificadas oficialmente como “Acesso Concedido”, “Acesso Parcialmente Concedido” e “Acesso Negado” porque a base usada para treinar o modelo NLP excluía as demais opções de classificação adotadas por órgãos federais: “Não se trata de solicitação de informação”, “Pergunta duplicada/repetida”, “Órgão não tem competência para responder sobre o assunto” ou “Informação inexistente”

Portanto, a amostra usada é constituída de 48.507 pedidos presentes em bases de dados divulgadas ativamente pela Controladoria-Geral da União (CGU)² que atendem aos critérios descritos acima.

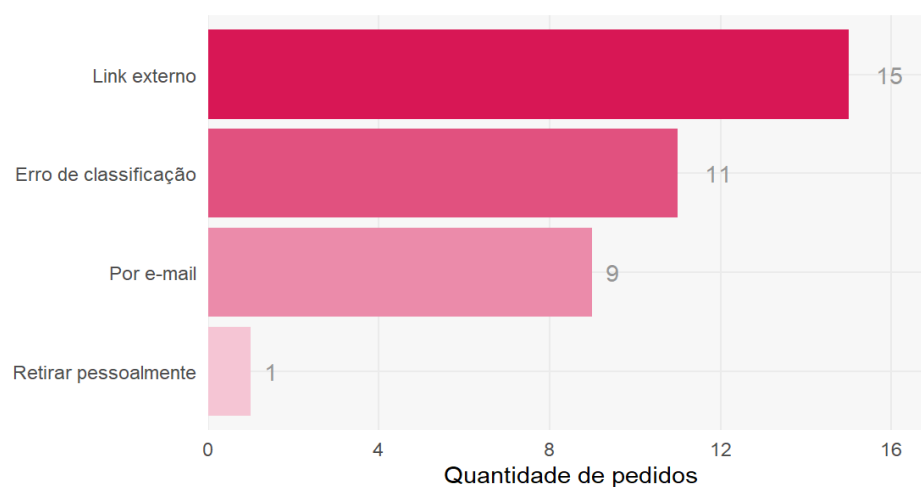
Após a aplicação do modelo sobre os pedidos, uma amostra aleatória de 100 deles foi analisada manualmente pela equipe da Transparência Brasil, a fim de confirmar os resultados obtidos com a ferramenta.

Concluiu-se que a classificação atribuída pelo modelo NLP-BERT a 54 pedidos foi correta e indicou de fato erro na categorização oficial. A classificação de outros 36 estava incorreta e 10 apresentaram ambiguidade (não foi possível determinar com convicção se haviam sido atendidos ou não).

Na maioria dos casos de classificação incorreta pelo modelo, a resposta havia sido fornecida por meio de link externo, por e-mail diretamente ao solicitante ou por indicação de retirada presencial, o que estruturalmente corresponde ao envio da informação dentro de um anexo – hipótese para a qual o modelo não foi treinado, conforme exposto anteriormente.

² Busca de Pedidos e Respostas - Download de dados. Controladoria-Geral da União.
http://www.consultaesic.cgu.gov.br/busca/_layouts/15/DownloadPedidos/DownloadDados.aspx

Gráfico 1. Casos de classificação incorreta pelo modelo NLP-BERT



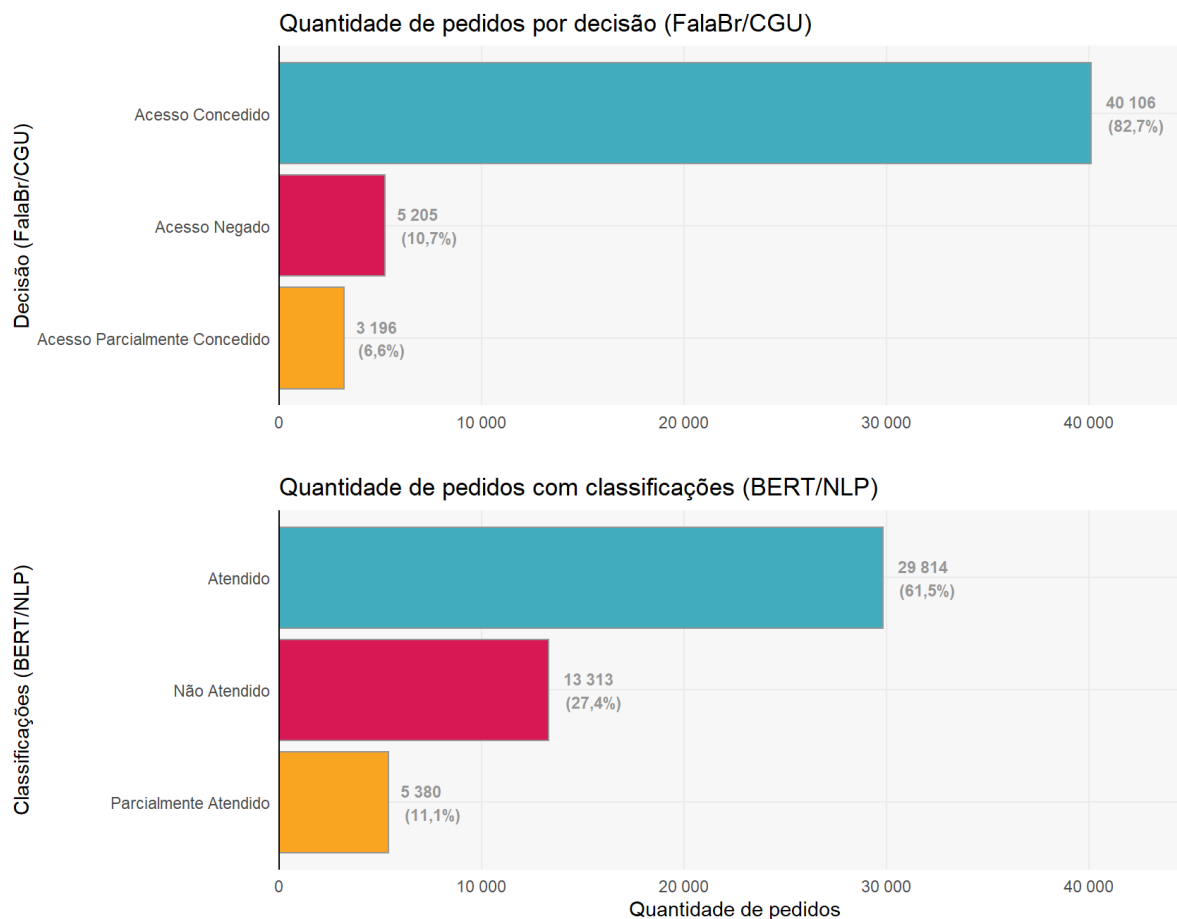
Ou seja, a precisão do modelo é alta.

Mais informações sobre a metodologia podem ser acessadas no seguinte [link](#) e os códigos utilizados podem ser acessados [aqui](#).

A conta não fecha: 32% dos pedidos classificados como “acesso concedido” não foram atendidos

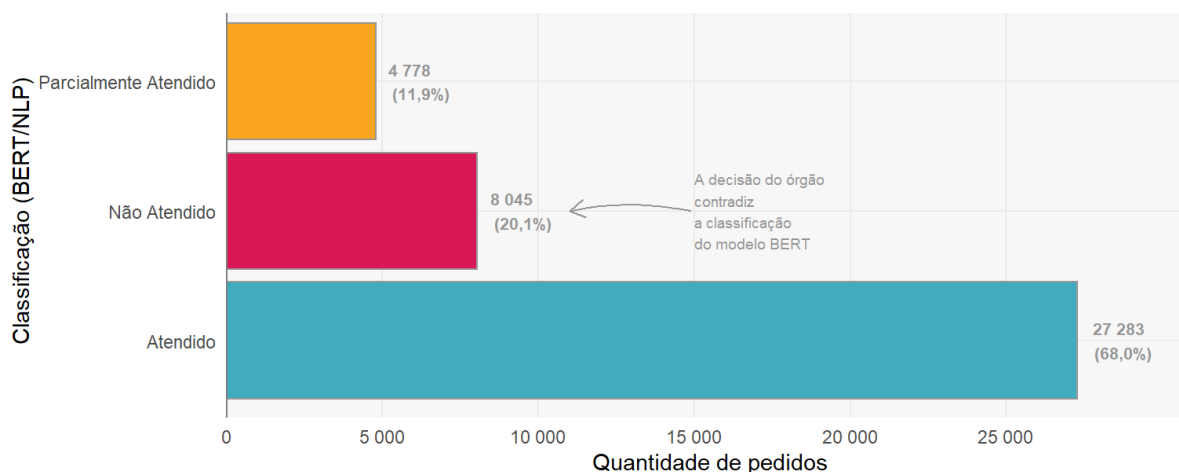
O resultado da comparação entre as classificações mostra que a proporção de fornecimento de informações nos órgãos federais no período analisado está superestimada. Pela categorização oficial, 83% dos pedidos da amostra usada para esta análise foram concedidos. O modelo NLP-BERT, entretanto, indica que apenas 61,5% dos pedidos foram de fato atendidos (Gráficos 2 e 3).

Gráficos 2 e 3: percentuais de pedidos atendidos na classificação oficial e na classificação pelo modelo NLP-BERT



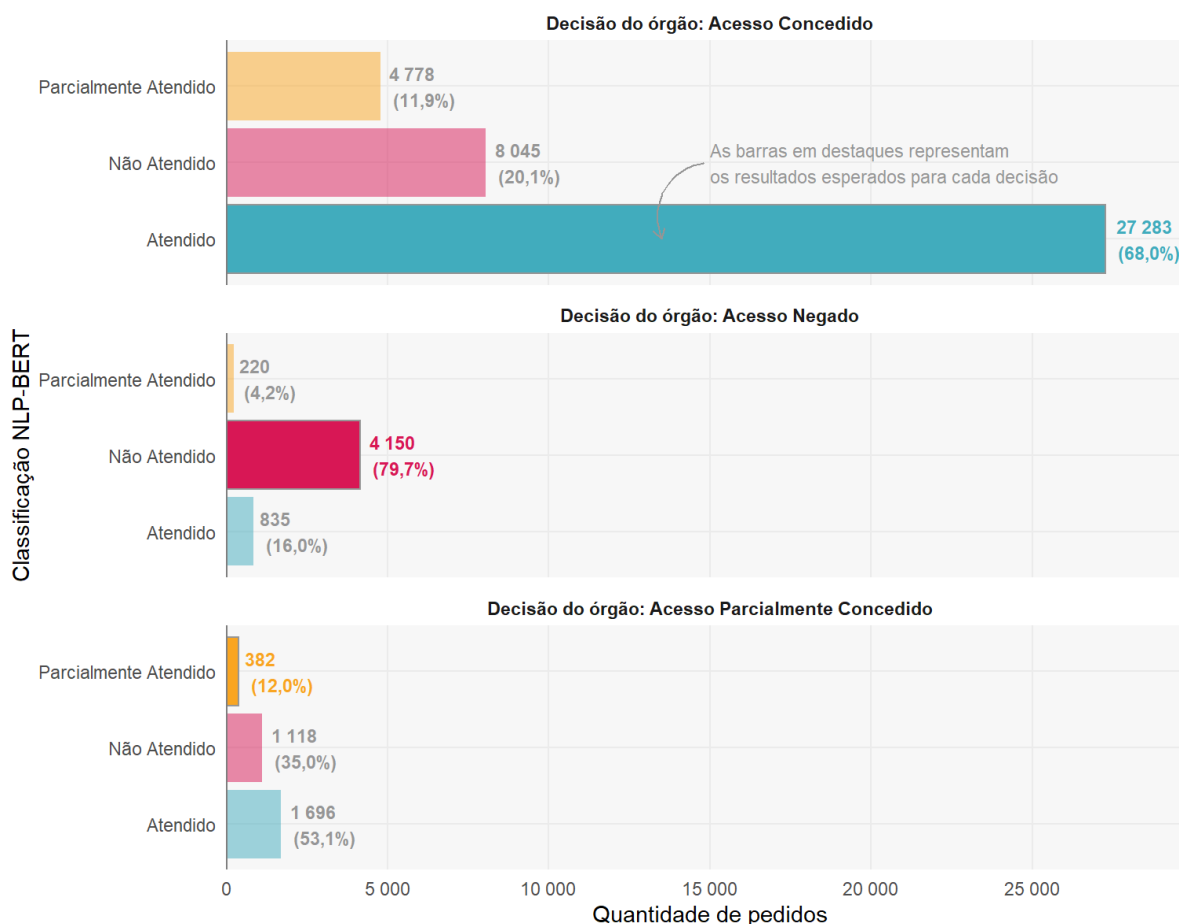
Ou seja, é necessário realizar ajustes na forma como é mensurado o desempenho do atendimento de pedidos de informação no Executivo federal. Entre os pedidos classificados como “acesso concedido” pelos órgãos, o modelo NLP-BERT classificou 20% como “não atendidos”. Outros 12% foram classificados pela ferramenta como “parcialmente atendidos” (Gráfico 3).

Gráfico 3. Classificação independente do atendimento a pedidos classificados como “acesso concedido” pelo órgão



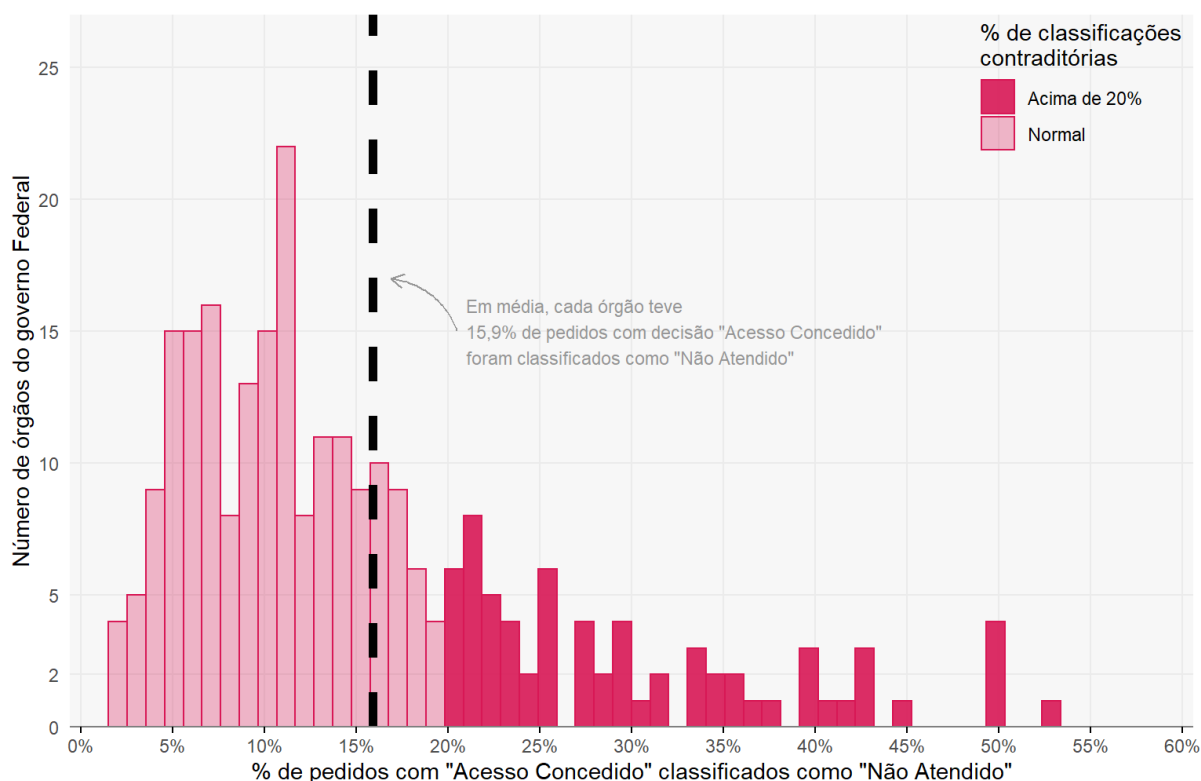
Entre os pedidos classificados pelos órgãos como “acesso parcialmente concedido”, a taxa é um pouco maior: o modelo rotulou 35% deles como “não atendidos” (Gráfico 4).

Gráfico 4. Classificação de atendimento (NLP-BERT) x Decisão do órgão



Em média, cada órgão federal classificou como “acesso concedido” 15,9% dos pedidos considerados não atendidos pelo NLP-BERT (Gráfico 5).

Gráfico 5. Distribuição do % de classificações equivocadas (“acesso concedido” x “não atendido”)



Padrões na classificação equivocada

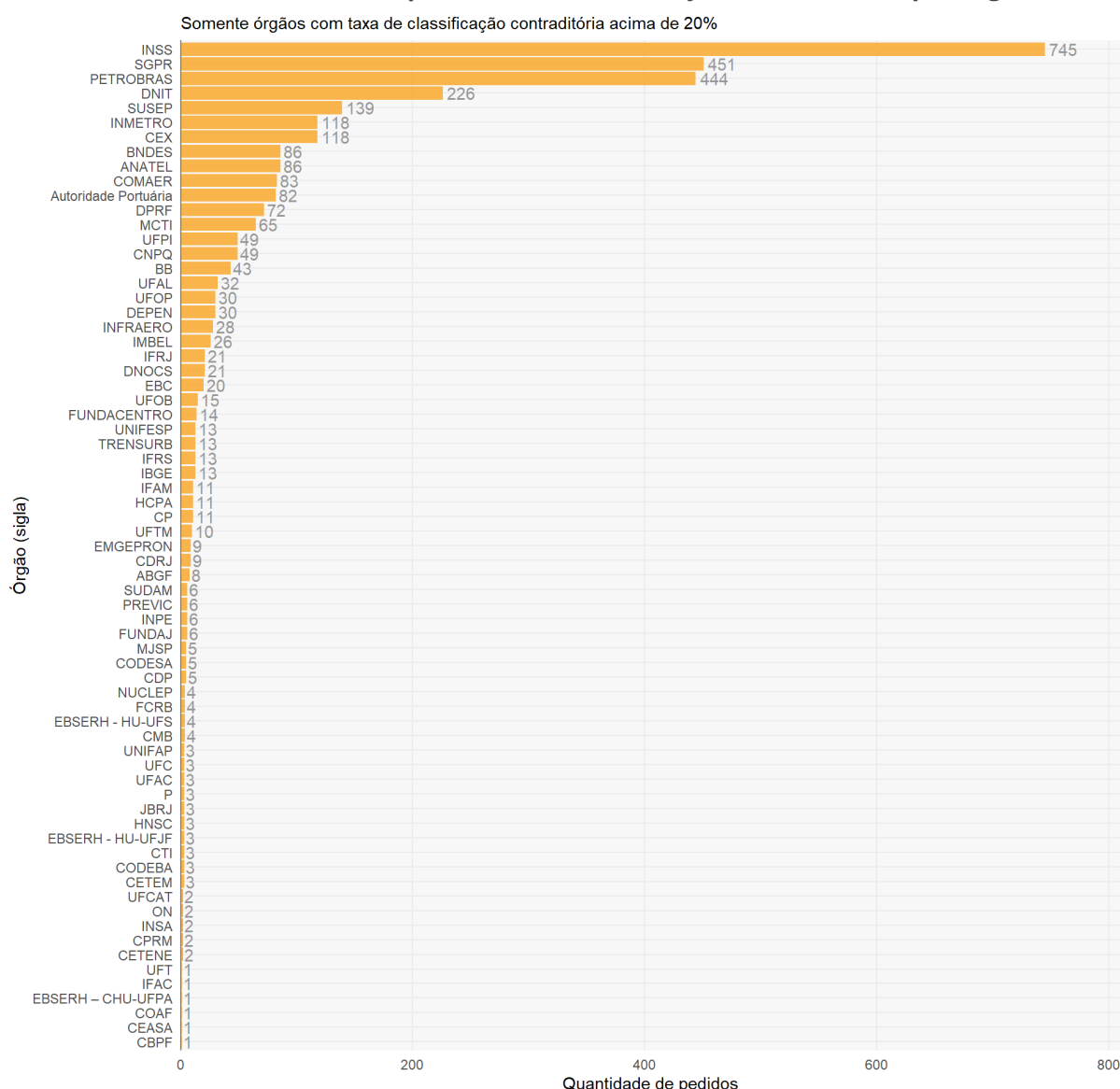
A fim de entender melhor o motivo das classificações equivocadas, foi feita uma avaliação qualitativa de uma amostra de pedidos da ANVISA³, da SGPR, da Petrobrás e do INSS⁴, que apresentaram alguns dos mais altos números desse tipo de ocorrência (Gráfico 6).

³ A ANVISA não aparece no gráfico 6 pois teve 18,1% de classificações contraditórias, mas a porcentagem foi considerada expressiva para ensejar análise qualitativa.

⁴ Foram analisados qualitativamente 20 pedidos da ANVISA, Petrobrás e do INSS e 451 pedidos da SGPR.

O caso da SGPR se destaca: o órgão classificou como “acesso concedido” 362 respostas em que afirmava não dispor das supostas provas sobre fraudes nas eleições de 2018 que o presidente Jair Bolsonaro (PL-RJ) alegou ter, em declaração feita [em março de 2020](#).

Gráfico 6. Quantidade de pedidos com classificação contraditória, por órgão



A ANVISA e o INSS apresentaram padrões semelhantes em seus erros de classificação. Categorizaram como “acesso concedido” respostas em que

apontaram ao requerente que a informação deveria ser obtida em outros meios que não o Fala.br e em que negaram o fornecimento da informação porque era necessária a comprovação da identidade do requerente.

No caso da ANVISA, o órgão considerou como “acesso concedido” respostas em que a informação só poderia ser obtida por meio de consulta às Agências Sanitárias locais. Também foram observados casos em que a informação não pode ser concedida sem a devida comprovação de identidade, como a íntegra de processos administrativos em andamento, cujo acesso é restrito às partes interessadas.

No caso do INSS, a maioria dos casos de erros envolve pedidos por informações pessoais, relativas a tempo de aposentadoria ou concessão de benefícios, por exemplo. O órgão classificou como “acesso concedido” respostas que eram, na verdade, negativas de atendimento sob a justificativa – legalmente válida – de que informações pessoais não podem ser fornecidas sem a devida confirmação de identidade do requerente. Em outros casos, a classificação foi “acesso concedido” apesar de o INSS ter considerado que a demanda não era um pedido de acesso à informação com base na LAI (uma categoria que existe na classificação oficial e, portanto, deveria ter sido aplicada) e orientou ao requerente a busca pela informação por outros meios que não o FalaBr.

Já no caso da Petrobrás, a análise qualitativa apontou que a resposta à maioria dos pedidos classificados equivocadamente como “acesso concedido” afirmava que seria necessário enviar as informações por e-mail para o requerente, mas como o pedido estava anonimizado (a identidade do cidadão e seu meio de contato haviam sido ocultados), seria necessário que o requerente fizesse outra solicitação de informação sem optar pela anonimização.

Cabe notar que a prática de envio da resposta pelo e-mail do solicitante, sob o argumento de que o Fala.br não comportaria o tamanho dos anexos que contêm a informação, é frequente na Petrobrás. Tendo isso em vista, é importante que a CGU elabore formas de verificar se o pedido foi realmente atendido nessas situações.

Ou seja, os órgãos analisados classificaram como concedidos pedidos em que a informação solicitada não foi de fato fornecida. É justamente por isso que ocorre a superestimação dos resultados dos pedidos classificados como concedidos pelos órgãos. Assim, é preciso que esses aprimorem a sua classificação, para que essa reflita da forma mais aproximada possível a percepção do cidadão sobre a resposta.

Na lista dos órgãos com os maiores números de classificações equivocadas (Gráfico 6), aparecem três do Executivo federal: a Secretaria Geral da Presidência da República (SGPR), o Comando da Aeronáutica (COMAER) e o Ministério da Ciência, Tecnologia e Inovação (MCTI).

Entre os 10 primeiros colocados, observa-se que metade são autarquias federais: Instituto Nacional do Seguro Social (INSS), Departamento Nacional de Infraestrutura de Transportes (DNIT), Superintendência de Seguros Privados (SUSEP), Instituto Nacional de Metrologia, Qualidade e Tecnologia (INMETRO) e a Agência Nacional de Telecomunicações (ANATEL). Esse dado indica a necessidade de atuação mais intensiva da Controladoria-Geral da União (CGU) para o aprimoramento das políticas de transparência passiva nesse tipo de instituição pública, enquanto órgão responsável pelas políticas de acesso à informação no governo federal.

Conclusão e recomendações de aprimoramento

Este relatório revelou indícios consistentes de que há problemas na classificação do atendimento a pedidos de informação por órgãos federais: mais de 30% dos que foram categorizados como concedidos não foram de fato atendidos, e com isso a proporção real de pedidos concedidos se encontra superestimada em até 20 pontos percentuais. Isso gera uma falsa percepção de que os dispositivos de transparência passiva da LAI estão sendo satisfatoriamente cumpridos, prejudicando o controle social.

Além disso, com dados imprecisos, perdem-se oportunidades de entender as deficiências da aplicação das políticas de transparência passiva e, conseqüentemente, propor mudanças para a sua melhoria.

Tal constatação aponta para a necessidade de aprimoramento na forma como essa classificação é operada no âmbito da administração pública federal, evitando-se que pedidos de acesso à informação que não tiveram a informação solicitada fornecida – independentemente do motivo – sejam classificados como “acesso concedido”.

Isso permitirá compreender os motivos que impediram o fornecimento da informação e, a partir desse diagnóstico, propor soluções. Essas podem estar direcionadas tanto para a necessidade de melhorar as respostas para atender a uma maior proporção de pedidos, quanto para a necessidade de melhorar sua comunicação externa, a fim de direcionar o cidadão para o canal correto de solicitação de informação ou manifestação. Para que a ação adequada seja empregada é preciso, portanto, que os dados sobre o atendimento dos pedidos de informação via LAI em âmbito federal tenham a sua acurácia aprimorada.

Para que a “conta feche” e as respostas dos pedidos sejam classificadas de forma mais apurada pelo olhar do cidadão, recomendamos à CGU os seguintes pontos de aprimoramento e avaliação:

- **Elaboração de um diagnóstico do sistema de classificação e das orientações repassadas aos órgãos federais para a sua aplicação.** Esta iniciativa é essencial para responder às seguintes perguntas: a alta porcentagem de classificações equivocadas apresentadas neste relatório é decorrente de categorias inadequadas ou insuficientes de classificação ou de sua aplicação incorreta pelos órgãos? Esse diagnóstico conduzirá os esforços de atuação, orientando se deverão se concentrar na alteração das categorias existentes ou em treinamento para servidores, por exemplo.

- Adicione formas de verificar se o pedido foi atendido quando a informação solicitada foi enviada para o e-mail do solicitante.
- **Comunique imediatamente aos órgãos federais a ocorrência dessa discrepância de classificações** no caso dos “acessos concedidos”, sem prejuízo da condução do diagnóstico citado acima, **e os oriente para o correto preenchimento** – em especial nas autarquias, onde foi observado alto número de classificações equivocadas, conforme expresso no gráfico 6.
- **Aprimore a pesquisa de satisfação do usuário**, substituindo a escala de 1 a 5 para classificação do atendimento pelas categorias “atendeu plenamente”, “atendeu parcialmente” e “não atendeu”, para deixar mais claras eventuais discrepâncias entre a classificação do órgão e a percepção do cidadão.
- Inclua os dados coletados por meio da pesquisa de satisfação acima, já presentes na página inicial do Painel Lei de Acesso à Informação, também na aba “Pedidos”.

Para que os dados gerados sobre a aplicação da LAI permitam avaliar com maior precisão se as informações solicitadas via LAI estão sendo fornecidas pelos órgãos, é necessário corrigir a discrepância apontada neste relatório. Tal medida é imprescindível para permitir o controle social e o consequente aperfeiçoamento das políticas públicas de transparência ativa em âmbito federal.