RELATÓRIO METODOLÓGICO

PRIVACIDADE E PROTEÇÃO DE DADOS PESSOAIS

Relatório Metodológico
Privacidade e Proteção de Dados Pessoais 2023

O Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação (Cetic.br), departamento do Núcleo de Informação e Coordenação do Ponto BR (NIC.br), ligado ao Comitê Gestor da Internet no Brasil (CGI.br), apresenta os aspectos metodológicos da pesquisa Privacidade e proteção de dados pessoais 2023: perspectivas de indivíduos, empresas e organizações públicas no Brasil. O objetivo da pesquisa é apurar o cenário atual e compreender os principais desafios para a construção de um ecossistema digital que garanta a privacidade e a proteção de dados pessoais no país. O levantamento de informações teve como base a coleta e o processamento de dados quantitativos por meio de pesquisas conduzidas regularmente pelo Cetic.br|NIC.br.

O projeto possui três objetivos específicos:

- investigar a percepção da população de usuários de Internet sobre o uso e a proteção de seus dados pessoais;
- compreender como pequenas, médias e grandes empresas tratam os dados pessoais de seus clientes/consumidores, bem como questões relevantes associadas à implementação da Lei Geral de Proteção de Dados Pessoais (LGPD) no Brasil;
- traçar um panorama da proteção de dados no contexto dos órgãos governamentais, estabelecimentos públicos de saúde e escolas públicas.

Na sequência, apresentamos os principais aspectos metodológicos das pesquisas utilizadas e as referências para acesso integral ao "Relatório Metodológico" e ao "Relatório de Coleta de Dados" de cada estudo.

Painel TIC – Usuários de Internet (2023)

Realizada por meio de questionários *online*, o Painel TIC foi desenvolvido como uma alternativa à coleta de dados presencial. A sua metodologia vem sendo adotada para o levantamento de dados sobre temas relevantes para o debate sobre a transformação digital desde 2020.

Em 2021, um novo módulo do Painel TIC foi desenvolvido para investigar a percepção da população de usuários de Internet sobre o tratamento e a proteção de seus dados pessoais (CGI.br, 2021). A elaboração de um questionário específico sobre privacidade entre usuários de Internet tomou como ponto de partida diversas pesquisas anteriores com objetivos convergentes. Uma das primeiras coletas de dados identificada foi a pesquisa do Eurobarômetro *Special Eurobarometer 431: Data protection*, de 2015, encomendada pela Comissão Europeia. Outra fonte relevante foi a edição de junho de 2019 da pesquisa *American Trends Panel* do Pew Research Center. Entre levantamentos oficiais produzidos por institutos nacionais de estatística, foi considerada a pesquisa *Survey of Canadians on Privacy-Related Issues*, realizada em 2020 por encomenda do Escritório do Comissariado de Privacidade do Canadá.

Também foi levada em conta a segunda edição da pesquisa Painel TIC COVID-19 do Cetic.br|NIC.br, que incluiu um módulo de privacidade. Esse módulo fazia parte de um esforço regional liderado pelo Banco Interamericano de Desenvolvimento (BID) com o objetivo de medir atitudes e percepções em relação à proteção de dados pessoais considerando o uso das tecnologias de informação e comunicação (TIC) em medidas de contenção da pandemia (CGI.br, 2020).

A pesquisa de 2023 segue os objetivos e referências do estudo de 2021. A população-alvo é composta de indivíduos usuários de Internet com 16 anos ou mais de idade no Brasil, considerando-se tais usuários as pessoas que fizeram uso da rede nos três meses que antecederam a entrevista, segundo recomendação metodológica da União Internacional de Telecomunicações (UIT, 2020).

Para seu desenho amostral, a pesquisa utilizou como base um painel *online* de indivíduos mantido pela Quaest Consultoria e Pesquisa, com aproximadamente 153 mil painelistas. O plano amostral empregado para a obtenção da amostra de respondentes foi do tipo amostragem por cotas, considerando as variáveis sexo, faixa etária, grau de escolaridade, macrorregião e classe. A coleta de dados da pesquisa foi realizada entre os dias 11 e 22 de dezembro de 2023 e, ao todo, foram obtidas 2.618 entrevistas.

Com o objetivo de minimizar os vieses de seleção encontrados em abordagens por cotas, foi construída uma estrutura de pesos para o Painel TIC, tendo como referência uma pesquisa probabilística, a TIC Domicílios 2023[1]. Na etapa inicial, os resultados dessa pesquisa foram recalibrados para a população da Pesquisa Nacional por Amostra de Domicílios Contínua (Pnad Contínua) (Instituto Brasileiro de Geografia e Estatística [IBGE], s.d.), referente ao último trimestre divulgado.

[1] Mais informações disponíveis em: https://www.cetic.br/pt/pesquisa/domicilios

Na sequência, com o intuito de estimar o contingente da população representada pelos respondentes do Painel TIC, adotou-se o procedimento de estimação com base em escores de propensão (*propensity scores*)[2]. Nessa metodologia, são calculados, inicialmente, os escores de propensão de ser usuário de Internet segundo variáveis socioeconômicas, com base na última edição disponível da pesquisa TIC Domicílios[3]. A seguir, esse mesmo modelo é utilizado para estimar os escores de propensão para os respondentes do Painel TIC.

Comparando a distribuição dos escores de propensão do Painel TIC com aquela verificada na última pesquisa TIC Domicílios, é possível determinar a parte da população que, desse último levantamento (ou se toda ela), poderia ser representada pelos respondentes do Painel. Isso equivale a estimar o erro de cobertura do Painel TIC em relação à população-alvo inicialmente considerada para a pesquisa.

Na presente edição do Painel TIC, o público representado equivale a toda a população-alvo da pesquisa TIC Domicílios, o que permite a comparação direta dos resultados da edição com os indicadores equivalentes coletados. Já em relação às edições anteriores do Painel, que não representavam a totalidade da população-alvo, a comparação precisa ser feita por meio dos mesmos recortes populacionais das respectivas edições.

Os resultados completos da pesquisa, bem como a íntegra do "Relatório Metodológico" do estudo, estão disponíveis no *website* do Cetic.br|NIC.br (https://www.cetic.br).

TIC Empresas – Pequenas, médias e grandes empresas (2023)

Realizada desde 2005, a pesquisa TIC Empresas tem como objetivo principal medir a posse e o uso das TIC entre as empresas brasileiras. O levantamento apresenta indicadores que traduzem em números a realidade das empresas brasileiras em relação a diversos temas, tais como acesso às TIC; uso da Internet; comércio eletrônico; habilidades em TIC; *software*; segurança digital; e novas tecnologias.

O universo abordado na pesquisa compreende todas as empresas brasileiras ativas com dez ou mais pessoas ocupadas[4] listadas no Cadastro Central de Empresas (Cempre) do IBGE, pertencentes aos setores da Classificação Nacional de Atividades Econômicas (CNAE) 2.0 de interesse da pesquisa TIC Empresas e à Natureza Jurídica 2 – entidades empresariais, exceto as empresas públicas (Natureza Jurídica 201-1).

[2] Diferentemente da estimativa baseada em um desenho amostral tradicional, as probabilidades de seleção no Painel são desconhecidas e indefinidas, por se tratar de um pseudodesenho amostral. A pseudoprobabilidade é a probabilidade estimada de pertencer à amostra não probabilística usada em vez de uma probabilidade conhecida. Mais informações disponíveis em Baker *et al*. (2013).

[3] Para esta edição do Painel TIC, foi utilizada a TIC Domicílios 2023 (CGI.br, 2024).

[4] A pesquisa TIC Empresas considera pequenas, médias e grandes empresas aquelas com, respectivamente, dez a 49 pessoas ocupadas, 50 a 249 pessoas ocupadas, e 250 pessoas ocupadas ou mais. As microempresas, aquelas com uma a nove pessoas ocupadas, não entram no escopo da pesquisa.

As empresas investigadas correspondem às seguintes seções:

C – Indústria de transformação;

F – Construção;

G – Comércio e reparação de veículos automotores e motocicletas;

H – Transporte, armazenagem e correio;

I – Alojamento e alimentação;

J – Informação e comunicação;

L – Atividades imobiliárias;

M – Atividades profissionais, científicas e técnicas;

N – Atividades administrativas e serviços complementares;

R – Artes, cultura, esporte e recreação;

S – Outras atividades de serviços.

A pesquisa TIC Empresas é desenvolvida com a preocupação de manter a comparabilidade internacional. Para isso, segue os padrões metodológicos propostos no manual da Conferência das Nações Unidas sobre Comércio e Desenvolvimento (UNCTAD, 2020), elaborado pela parceria entre a Organização para a Cooperação e Desenvolvimento Econômico (OCDE), o Instituto de Estatísticas da Comissão Europeia (Eurostat) e a Partnership on Measuring ICT for Development – esta última, uma coalizão formada por diversas organizações internacionais, busca a harmonização de indicadores-chave em pesquisas sobre TIC.

O plano amostral é estratificado em duas etapas, e as empresas são selecionadas aleatoriamente dentro de cada estrato. A primeira etapa compreende a definição de estratos naturais por meio do cruzamento das variáveis região geográfica e mercado de atuação (CNAE 2.0). Com base em cada estrato natural, são definidos os estratos finais, que consideram a divisão dos estratos naturais por porte da empresa[5]. Em 2023, a pesquisa entrevistou um total de 4.457 empresas, sendo que 2.075 responderam às perguntas específicas do módulo sobre privacidade e proteção dos dados pessoais.

As empresas são contatadas por meio da técnica de entrevista telefônica assistida por computador (do inglês, *computer-assisted telephone interviewing* – CATI). Em todas as entidades pesquisadas, busca-se entrevistar o responsável pela área de informática, tecnologia da informação (TI), gerenciamento da rede de computadores ou área equivalente, o que corresponde a cargos como:

[5] As faixas de porte consideradas são: dez a 19 pessoas ocupadas; 20 a 49 pessoas ocupadas; 50 a 249 pessoas ocupadas; e 250 pessoas ocupadas ou mais.

- Diretor da divisão de informática e tecnologia;
- Gerente de negócios (vice-presidente sênior, vice-presidente de linha de negócios, diretor);
- Gerente ou comprador do departamento de tecnologia;
- Influenciador tecnológico (funcionário do departamento comercial ou de operações de TI com influência sobre as decisões a respeito de questões tecnológicas);
- Coordenador de projetos e sistemas;
- Diretor de outros departamentos ou divisões (excluindo informática);
- Gerente de desenvolvimento de sistemas;
- Gerente de informática;
- Gerente de projetos;
- Dono da empresa ou sócio.

Nas empresas que declaram ter, no momento da entrevista, 250 pessoas ocupadas ou mais, a estratégia foi entrevistar um segundo profissional, preferencialmente o gestor da área contábil ou financeira. Quando não encontrado, buscou-se o responsável pela área administrativa, jurídica ou de relações com instituições governamentais, a quem cabem exclusivamente as respostas sobre comércio eletrônico e atividades realizadas na Internet.

Na aplicação do módulo de Privacidade e Proteção de Dados, é entrevistado um respondente adicional, qualificado para responder sobre medidas relativas ao cumprimento da Lei Geral de Proteção de Dados Pessoais (LGPD) na empresa. Para esse módulo, é solicitado que os respondentes da pesquisa TIC Empresas indiquem a pessoa mais familiarizada com o tema na empresa, ou seja, quem poderia responder sobre procedimentos e políticas adotados para coleta, armazenamento e uso de dados pessoais, bem como sobre a adequação da empresa à LGPD. Nos casos em que o tema é liderado pelo respondente da TIC Empresas, a entrevista é realizada com esse profissional. Não é permitido que a organização indique um profissional terceirizado como respondente, buscando-se, alternativamente, identificar o funcionário interno responsável pela contratação desse serviço, de modo a garantir que as entrevistas sejam realizadas com membros da equipe interna da empresa.

Os resultados e as tabelas de proporções, totais e margens de erro da TIC Empresas, bem como as íntegras do "Relatório Metodológico" e do "Relatório de Coleta de Dados" do estudo, estão disponíveis no *website* do Cetic.br|NIC.br (https://www.cetic.br).

TIC Governo Eletrônico – Órgãos públicos federais e estaduais e prefeituras (2023)

Realizada a cada dois anos desde 2013, a pesquisa sobre o uso das tecnologias de informação e comunicação no setor público brasileiro – TIC Governo Eletrônico – investiga a incorporação das tecnologias digitais nos órgãos públicos e o seu uso para a oferta de serviços públicos. O estudo ainda mede a existência de iniciativas relacionadas à promoção do acesso à informação pública e participação da sociedade por meio das novas tecnologias. A partir de 2021, foram incluídos novos módulos relacionados à adoção de novas tecnologias, bem como indicadores sobre privacidade e proteção de dados pessoais.

A pesquisa tem abrangência nacional e inclui duas unidades de análise: órgãos públicos federais e estaduais de todos os poderes (Executivo, Legislativo, Judiciário e Ministério Público) e prefeituras. É realizado um censo em todos os públicos de interesse, excetuando órgãos do Executivo estadual, sendo selecionada uma amostra de aproximadamente 400 entidades públicas. As entrevistas são realizadas por meio de questionário estruturado pela técnica de CATI.

Os indicadores analisados para esta publicação foram coletados entre julho de 2023 e fevereiro de 2024, em 677 órgãos públicos federais e estaduais e 4.265 prefeituras. Os resultados e as tabelas de proporções, totais e margens de erro da TIC Governo Eletrônico estão disponíveis no *website* do Cetic.br|NIC.br (https://www.cetic.br), bem como as íntegras do "Relatório Metodológico" e do "Relatório de Coleta de Dados" do estudo.[6]

TIC Saúde – Estabelecimentos públicos de saúde (2023)

Realizada anualmente desde 2013, a pesquisa TIC Saúde tem o objetivo de compreender o estágio de adoção das TIC nos estabelecimentos de saúde e sua apropriação pelos profissionais da área (médicos e enfermeiros). Para isso, busca identificar a infraestrutura de TIC disponível e investigar o uso de sistemas e aplicações baseados em TIC destinados a apoiar os serviços de assistência e a gestão dos estabelecimentos de saúde. Além disso, mede as atividades realizadas por profissionais de saúde por meio das TIC, bem como as motivações e barreiras para sua adoção e uso.

Em 2021, a pesquisa incluiu um indicador que investigou a adaptação dos estabelecimentos de saúde em relação a algumas medidas indicadas na LGPD. Em 2022, foram inseridos indicadores referentes a treinamento em segurança da informação tanto disponibilizado pelo estabelecimento de saúde quanto realizado pelos profissionais.

[6] Disponíveis em: https://cetic.br/pt/publicacao/pesquisa-sobre-o-uso-das-tecnologias-de-informacao-e-comunicacao-no-setor-publico-brasileiro-tic-governo-eletronico-2023/

A pesquisa TIC Saúde tem abrangência nacional e coleta dados dos estabelecimentos de saúde pública e privada nos três níveis de atenção, selecionando-os com base no Cadastro Nacional de Estabelecimentos de Saúde (CNES), mantido pelo Ministério da Saúde (MS). As entrevistas são realizadas por meio da técnica de CATI e há a possibilidade de autopreenchimento de questionário *web*, por meio de plataforma específica.

Os resultados da edição de 2023 foram coletados entre fevereiro e julho desse mesmo ano com 4.117 gestores, representando um universo de 120.069 estabelecimentos de saúde brasileiros. Os resultados e as tabelas de proporções, totais e margens de erro da TIC Saúde estão disponíveis no *website* do Cetic.br|NIC.br (https://www.cetic.br), bem como as íntegras do "Relatório Metodológico"[7] e do "Relatório de Coleta de Dados" do estudo.[8]

TIC Educação – Escolas públicas (2022 e 2023)

Realizada desde 2010, a pesquisa TIC Educação investiga o acesso, o uso e a apropriação das TIC pela comunidade educacional, principalmente alunos e professores, em atividades de ensino, de aprendizagem e de gestão escolar. Com abrangência nacional, a pesquisa é aplicada anualmente em escolas de Educação Básica, públicas e particulares, localizadas em áreas urbanas e rurais e que oferecem classes de Ensino Fundamental e Médio regular. Na edição 2020, a inclusão de um módulo específico sobre privacidade, com questões acerca da segurança digital e da coleta e proteção de dados pessoais, permitiu um maior contato com as percepções e experiências dos diferentes atores escolares sobre essa temática específica.

Os dados analisados nesta publicação baseiam-se primordialmente nos indicadores coletados nas edições 2022 e 2023 da pesquisa TIC Educação. A edição 2022 foi realizada entre outubro de 2022 e maio de 2023, de forma presencial, por meio da técnica de entrevista pessoal assistida por computador (do inglês *computer-assisted personal interviewing* [CAPI]), em 1.394 escolas. Ao todo, foram entrevistados 7.192 estudantes do 4º ano do Ensino Fundamental ao 3º ano do Ensino Médio, 1.424 professores, 873 coordenadores pedagógicos e 959 gestores escolares. Os dados da edição 2023 foram coletados entre os meses de agosto de 2023 e abril de 2024, por meio de entrevistas telefônicas (CATI), com 3.004 gestores escolares.

Assim como para as demais pesquisas, os resultados e as tabelas de proporções, totais e margens de erro da TIC Educação estão disponíveis no *website* do Cetic.br|NIC.br (https://www.cetic.br), bem como as íntegras do "Relatório Metodológico"[9] e do "Relatório de Coleta de Dados"[10] dos estudos.

[7] Disponível em: https://cetic.br/media/microdados/773/tic_saude_2023_relatorio_metodologico_v1.0.pdf

[8] Disponível em: https://cetic.br/media/microdados/771/tic_saude_2023_relatorio_coleta_de_dados_v1.0.pdf

[9] Disponível em: https://cetic.br/media/microdados/785/tic_educacao_2023_relatorio_metodologico_v1.0.pdf

[10] Disponível em: https://cetic.br/media/microdados/784/tic_educacao_2023_relatorio_coleta_de_dados_v1.0.pdf

Disseminação dos dados

Os resultados das pesquisas mencionadas anteriormente são apresentados de acordo com as variáveis descritas no "Relatório Metodológico" de cada estudo, no item "Domínios de interesse para análise e divulgação".

Arredondamentos fazem com que, em alguns resultados, a soma das categorias parciais difira de 100% em questões de resposta única. O somatório de frequências em questões de respostas múltiplas usualmente é diferente de 100%. Vale ressaltar que, nas tabelas de resultados, o hífen (-) é utilizado para representar a não resposta ao item. Por outro lado, como os resultados são apresentados sem casa decimal, as células com valor zero indicam que houve resposta ao item, mas ele é explicitamente maior do que zero e menor do que um.

Os resultados das pesquisas são publicados em formato *online* e disponibilizados no *website* do Cetic.br|NIC.br (https://www.cetic.br). As tabelas de proporções, totais e margens de erros calculadas para cada indicador estão disponíveis para *download* em português, inglês e espanhol. Mais informações sobre a documentação, os metadados e as bases de microdados estão disponíveis na página de microdados (https://www.cetic.br/microdados/).

Referências

Baker, R., Brick, J. M., Bates, N. A., Battaglia, M., Couper, M. P., Dever, J. A., Gile, K. J., & Tourangeau, R. (2013). *Report of the AAPOR Task Force on non-probability sampling*. https://aapor.org/wp-content/uploads/2022/11/NPS_TF_Report_Final_7_revised_FNL_6_22_13-2.pdf

Comitê Gestor da Internet no Brasil. (2020). *Painel TIC COVID-19: Pesquisa sobre o uso da Internet no Brasil durante a pandemia do novo coronavírus – 2ª edição: Serviços públicos on-line, telessaúde e privacidade*. https://cetic.br/media/docs/publicacoes/1/20201001085713/painel_tic_covid19_2edicao_livro%20eletr%C3%B4nico.pdf

Comitê Gestor da Internet no Brasil. (2021). *Painel TIC COVID-19: Pesquisa web sobre o uso da Internet no Brasil durante a pandemia do novo coronavírus*. https://www.cetic.br/pt/publicacao/painel-tic-covid-19/

Comitê Gestor da Internet no Brasil. (2024). *Pesquisa sobre o uso das tecnologias de informação e comunicação nos domicílios brasileiros: TIC Domicílios 2023*. https://cetic.br/pt/publicacao/pesquisa-sobre-o-uso-das-tecnologias-de-informacao-e-comunicacao-nos-domicilios-brasileiros-tic-domicilios-2023/

Conferência das Nações Unidas sobre Comércio e Desenvolvimento. (2020). *Manual for the production of statistics on the digital economy 2020*. https://unctad.org/publication/manual-production-statistics-digital-economy-2020

Instituto Brasileiro de Geografia e Estatística. (s.d.). *Pesquisa nacional por amostra de domicílios contínua (Pnad Contínua)*. https://www.ibge.gov.br/estatisticas/sociais/trabalho/9173-pesquisa-nacional-por-amostra-de-domicilios-continua-trimestral.html

Lei Geral de Proteção de Dados Pessoais – LGPD. 13.709, de 14 de agosto de 2018. (2018). Lei Geral de Proteção de Dados Pessoais (LGPD). https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm

União Internacional de Telecomunicações. (2020). *Manual for measuring ICT access and use by households and individuals, 2020 edition*. https://www.itu.int/en/ITU-D/Statistics/Pages/publications/manual.aspx