



RELATÓRIO METODOLÓGICO

PESQUISA TIC PROVEDORES 2020

Relatório Metodológico TIC Provedores

O Comitê Gestor da Internet no Brasil (CGI.br), por meio do Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação (Cetic.br), do Núcleo de Informação e Coordenação do Ponto BR (NIC.br), apresenta o “Relatório Metodológico” da pesquisa TIC Provedores.

Realizado desde 2011, o estudo tem o objetivo de gerar informações que proporcionem uma visão ampla sobre a atuação do setor de provimento de acesso à Internet no Brasil. Abarcando todo o território nacional, a investigação busca dimensionar as principais características da atuação dos provedores no Brasil, bem como identificar as necessidades e potencialidades dessas empresas ao abordar questões que estão divididas nos seguintes módulos:

- **Módulo A:** Dados gerais da empresa;
- **Módulo B:** Serviços ofertados e mercado de atuação;
- **Módulo C:** Modelo de atuação;
- **Módulo D:** Infraestrutura – tecnologia e velocidades de acesso;
- **Módulo E:** Pontos de Troca de Tráfego;
- **Módulo F:** Ativação do IPv6;
- **Módulo G:** Segurança.

O processo de realização da pesquisa baseia-se na coleta de informações em todas as unidades presentes no cadastro (censo) com a intenção de reunir o maior número de provedores de serviços de acesso à Internet possível. Para a criação de uma listagem inicial, parte-se da base de informações do cadastro de empresas que possuem licença da Agência Nacional de Telecomunicações (Anatel) e também da base de empresas cadastradas como possuidoras de Sistema Autônomo do NIC.br. O apoio de associações e outras organizações públicas e privadas, ligadas aos provedores de serviços de acesso à Internet, é fundamental para que a pesquisa seja realizada.

Objetivos da pesquisa

O objetivo da TIC Provedores é oferecer um mapeamento do setor de provimento de acesso à Internet no Brasil. Para tanto, o estudo busca caracterizar as empresas provedoras de Internet em termos de serviços oferecidos, atuação no mercado e adoção de tecnologias.

Conceitos e definições

Para compor o cadastro inicial da pesquisa são usadas duas fontes de informação: o cadastro de empresas que possuem licença para a distribuição de Serviços de Comunicação Multimídia (SCM) e da base de informações do Sistema de Coleta de Informações (SICI). A essas duas bases são incorporadas informações da base de Sistemas Autônomos registrados no NIC.br.

Empresa provedora de Internet e serviço comunicação multimídia

No Brasil, para se tornar provedora de Internet, uma empresa primeiro precisa ser constituída formalmente, isto é, possuir um número do Cadastro Nacional de Pessoa Jurídica (CNPJ) e pleitear junto à Agência Nacional de Telecomunicações (Anatel) a sua licença para a prestação de Serviços de Comunicação Multimídia (SCM). De acordo com definição no *website* da Anatel, SCM é:

(...) um serviço fixo de telecomunicações de interesse coletivo, prestado em âmbito nacional e internacional, no regime privado, que possibilita a oferta de capacidade de transmissão, emissão e recepção de informações multimídia, permitindo inclusive o provimento de conexão à Internet, utilizando quaisquer meios, a Assinantes dentro de uma Área de Prestação de Serviço.¹

Dessa forma, licenciada como SCM, uma empresa pode distribuir informações multimídia, ou melhor, dados. Contudo, não tem permissão para oferecer telefonia fixa (definida pelo regulador como Serviços de Telefonia Fixa Comutada – STFC) ou televisão paga (definida pelo regulador como Serviços de Acesso Condicionado – SEAC). Portanto, segundo a regulação do setor, ao provedor que possui a licença SCM, é permitido distribuir pela sua infraestrutura o conteúdo de uma página da Internet, mas ele necessita de outras licenças para transmitir, por exemplo, conteúdos exclusivos de canais de televisão ou permitir ligações telefônicas.

¹Mais informações no *website* da Anatel. Recuperado em 10 março, 2021, de <https://www.gov.br/anatel/pt-br/regulador/outorga/comunicacao-multimidia>

Sistemas autônomos

Um Sistema Autônomo (do inglês, *Autonomous System – AS*) é definido – segundo o Grupo de Trabalho (GT) Marco Civil e as Responsabilidades do CGI.br, um documento que oferece recomendações sobre a aplicação das leis sobre a Internet no Brasil – como “uma rede ou um grupo de redes IP sob uma única administração, a qual determina como trafegar e distribuir os pacotes de dados em seu interior” (Comitê Gestor da Internet no Brasil [CGI.br], 2018, p. 6).

POPULAÇÃO-ALVO

O universo abordado na pesquisa compreende as empresas que possuem licença concedida pela Anatel para prestação de Serviços de Comunicação Multimídia (SCM) e são provedoras de acesso à Internet no Brasil.

UNIDADE DE ANÁLISE E REFERÊNCIA

A unidade de análise é a empresa provedora de serviço de acesso à Internet.

DOMÍNIOS DE INTERESSE PARA ANÁLISE E DIVULGAÇÃO

Para as unidades de análise e referência, os resultados são divulgados para domínios definidos com base nas variáveis e níveis descritos a seguir:

- Região: corresponde à divisão regional do Brasil, segundo critérios do Instituto Brasileiro de Geografia e Estatística (IBGE), nas macrorregiões Centro-Oeste, Nordeste, Norte, Sudeste e Sul;
- Número de acessos: considera-se o número máximo de acessos mensais – que são os meios pelos quais a conexão à Internet é estabelecida – declarados nos últimos 12 meses na base SICI da Anatel.

Instrumento de coleta

INFORMAÇÕES SOBRE OS INSTRUMENTOS DE COLETA

A pesquisa é realizada por meio de entrevista telefônica com questionário estruturado preenchido em computador (do inglês, *Computer Assisted Telephone Interviewing – CATI*), ou por questionário *web* (do inglês, *Computer Assisted Web Interviewing – CAWI*), caso o respondente solicite esse meio de resposta. Os questionários são aplicados na íntegra para todos os respondentes da pesquisa. Para mais informações a respeito do questionário, ver item “Instrumento de coleta” no “Relatório de Coleta de Dados” da pesquisa TIC Provedores.

PLANO AMOSTRAL

Cadastro e fontes de informação

A pesquisa TIC Provedores conta com o apoio de diversas associações nacionais de provedores de acesso à Internet para a divulgação da iniciativa. Para a construção do cadastro, são utilizados os dados existentes nas bases do SICI e SCM, da Anatel, e do NIC.br – nesta última, estão informações referentes ao cadastro de empresas que possuem Sistema Autônomo. O cadastro de provedores é fornecido pela Anatel – base de SCM e SICI. Essa base do SICI é dividida para o planejamento da pesquisa em duas partes distintas: empresas com informação de acessos entre um período determinado e empresas do cadastro sem informação de acessos nesse período (para resultados da presente edição ver “Relatório de Coleta de Dados”). A essas duas partes são agregados os dados da base de Sistemas Autônomos do NIC.br.

Esse conjunto total é dividido em estratos, como indicado na Tabela 1. As frequências de registros de empresas em cada um dos estratos formados constam no “Relatório de Coleta de Dados”.

TABELA 1

ESTRATOS PARA ORGANIZAÇÃO DA COLETA

Região	AS	Porte
Norte	AS	Nenhum
	AS	De 1 a menos de 5 000
	AS	De 5 000 a menos de 50 000
	AS	De 50 000 e mais
	Não é AS	Nenhum
	Não é AS	De 1 a menos de 5 000
Nordeste	AS	Nenhum
	AS	De 1 a menos de 5 000
	AS	De 5 000 a menos de 50 000
	AS	De 50 000 e mais
	Não é AS	Nenhum
	Não é AS	De 1 a menos de 5 000
	Não é AS	De 5 000 a menos de 50 000
Sudeste	AS	Nenhum
	AS	De 1 a menos de 5 000
	AS	De 5 000 a menos de 50 000
	AS	De 50 000 e mais
	Não é AS	Nenhum
	Não é AS	De 1 a menos de 5 000
	Não é AS	De 5 000 a menos de 50 000
	Não é AS	De 50 000 e mais

CONTINUA ►

► CONCLUSÃO

Região	AS	Porte
Sul	AS	Nenhum
	AS	De 1 a menos de 5 000
	AS	De 5 000 a menos de 50 000
	AS	De 50 000 e mais
	Não é AS	Nenhum
	Não é AS	De 1 a menos de 5 000
	Não é AS	De 5 000 a menos de 50 000
Centro-Oeste	AS	Nenhum
	AS	De 1 a menos de 5 000
	AS	De 5 000 a menos de 50 000
	Não é AS	Nenhum
	Não é AS	De 1 a menos de 5 000
	Não é AS	De 5 000 a menos de 50 000

A pesquisa leva em consideração separadamente os estratos, para que esforços de coleta distintos sejam adotados para obtenção de respostas para o estudo. A motivação para isso é que não se pode realizar esforço de coleta idêntico para todas as empresas, e acredita-se que aquelas alocadas no estrato sem informação de acessos não têm a mesma probabilidade de serem provedoras ativas como as incluídas nos demais estratos.

Os maiores provedores de atuação nacional, segundo classificação da Anatel, são Claro, Vivo, TIM, Oi e Sky/AT&T, e não foram incluídos na pesquisa. Ainda que respondam pela maior parte do mercado de provimento, o correspondente a 62%, de acordo com os dados de dezembro de 2020², considera-se que tais empresas apresentam características muito distintas do conjunto da população de provedores. O instrumento de coleta adotado seria insuficiente para refletir as operações desses maiores provedores em toda a sua complexidade. Por serem poucas unidades, a exclusão das grandes operadoras de telecomunicações não compromete as estimativas produzidas pela TIC Provedores quanto à quantidade de empresas e as características do setor.

Coleta de dados em campo

MÉTODO DE COLETA

Todas as empresas são contatadas por meio da técnica de entrevista telefônica assistida por computador (CATI). Há a possibilidade de autopreenchimento de questionário *web*, por meio de plataforma específica. Essa opção é oferecida para aqueles respondentes que solicitem espontaneamente responder via Internet ou para aqueles que prontamente se recusam a responder a pesquisa pelo telefone.

² Dados disponíveis no *website* da Anatel. Recuperado em 8 março, 2021, de <https://www.anatel.gov.br/paineis/acessos/banda-larga-fixa>

A estes provedores é enviado um *link* específico para o seu questionário, permitindo alterações na resposta, bem como há o acompanhamento e a sensibilização, via telefone, daqueles respondentes que ainda se mostram hesitantes em iniciar ou em concluir o questionário.

Em todas as empresas pesquisadas, busca-se entrevistar o responsável pela área de informática, tecnologia da informação, gerenciamento da rede de computadores ou área equivalente, o que corresponde a cargos como:

- Proprietário;
- Diretor da divisão de informática e tecnologia;
- Gerente de negócios (vice-presidente sênior, vice-presidente de linha de negócios, diretor).

Processamento dos dados

Para fins do tratamento de não respostas e obtenção de pesos para os dados das empresas pesquisadas, as ocorrências de coleta foram classificadas em três grupos:

- **Grupo 1:** empresas provedoras de acesso à Internet – correspondem às empresas que responderam “Sim” à questão sobre provimento de Internet;
- **Grupo 2:** empresas que não existem ou que não são provedoras de acesso à Internet – correspondem às empresas que responderam “Não” à questão sobre provimento de Internet ou que se declararam inativas;
- **Grupo 3:** empresas em que não foi possível identificar a sua condição de provimento e de atividade – correspondem às empresas que não foram possíveis de se contatar, recusaram-se a participar da pesquisa ou que abandonaram a entrevista antes de responder à questão sobre provimento de Internet (ver mais sobre resultado do campo no “Relatório de Coleta de Dados”).

Diante das informações obtidas em campo, o total de empresas provedoras ativas no país foi estimado por meio de um processo de pós-estratificação, que considerou informações obtidas no cadastro de pessoas jurídicas no *site* da Receita Federal do Brasil³. O número de empresas provedoras ativas é dado pelo total de empresas do cadastro inicial da pesquisa segundo a classificação obtida após coleta dos dados, em cada estrato de atividade econômica (CNAE) listado na base da Receita Federal. A estimação foi realizada considerando três estratos de CNAE Fiscal:

- Empresas ativas cuja CNAE principal é prestação de Serviços de Comunicação e Multimídia;

³ A base contém informações de todas as empresas brasileiras e sua situação de atividade. Para a base de dados disponibilizada em novembro de 2020 foram consideradas apenas as empresas em situação ativa. Recuperado em 21 janeiro, 2021, de <https://receita.economia.gov.br/orientacao/tributaria/cadastros/cadastro-nacional-de-pessoas-juridicas-cnpj/dados-publicos-cnpj>

- Empresas ativas cuja CNAE principal é prestação de provimento de acesso a redes de comunicação;
- Empresas ativas com outras CNAEs principais.

O total de empresas provedoras foi estimado em cada pós-estrato de forma a considerar a classificação de acordo com a condição de resposta à pesquisa (grupos de classificação 1 e 2). Para o grupo 3, em que não há informação sobre a condição de provimento de acesso à Internet, a pós-estratificação foi realizada a partir das respostas dos grupos 1 e 2. Além disso, cada pós-estrato foi dividido segundo estratos definidos por:

- Localização da sede da empresa segundo a Grande Região a qual pertence (Norte, Nordeste, Sudeste, Sul e Centro-Oeste);
- Ser ou não um Sistema Autônomo;
- Porte segundo número máximo de acessos mensais declarados no SICI nos 12 meses anteriores ao início da pesquisa (nenhum acesso, de 1 a menos de 5 mil acessos, de 5 mil a menos de 50 mil acessos, de 50 mil acessos ou mais).

Em cada um dos estratos construídos com esses cruzamentos, para as empresas ativas na base de CNPJs, foi estimado o total de provedores segundo a Fórmula 1.

FÓRMULA 1

$$P = \sum_i N_i \times \frac{p_i}{(p_i + q_i)}$$

P é o total de empresas provedoras no país

N_i é o total de empresas ativas da base da pesquisa do pós-estrato i

p_i é o total de empresas ativas da base da pesquisa do pós-estrato i respondentes da pesquisa que são provedoras de Internet (grupo 1)

q_i é o total de empresas ativas da base da pesquisa do pós-estrato i respondentes da pesquisa que não são provedoras de Internet (grupo 2)

A partir dessa definição, estima-se o total de empresas provedoras de acesso à Internet no Brasil, conforme indicado no “Relatório de Coleta de Dados”.

PROCEDIMENTOS DE PONDERAÇÃO

Para obter os resultados para o universo de provedores com base nos dados coletados, é feita correção de não resposta para os respondentes em cada pós-estrato. Esse procedimento considera que aqueles provedores que não responderam à pesquisa (recusas diversas, desistências, etc.) são homogêneos em relação às informações prestadas pelos respondentes, dentro do pós-estrato. A correção de não resposta consiste em dar pesos às empresas respondentes para compensar as não respostas. Os pesos de cada informante da pesquisa são obtidos mediante a razão do total de empresas estimadas como provedoras no pós-estrato pelo total de empresas respondentes no pós-estrato.

FÓRMULA 2

$$w_{ji} = \frac{P_i}{n_i}$$

w_{ji} é o peso do informante j do pós-estrato i

P_i é o total de provedores estimados no pós-estrato i

n_i é o total de provedores respondentes no pós-estrato i

ERROS AMOSTRAIS

As medidas ou estimativas dos erros amostrais dos indicadores da TIC Provedores levam em consideração em seus cálculos o plano amostral por estratos empregado na pesquisa.

Assim, a divulgação dos erros amostrais, expressos pela margem de erro, é feita a partir das variâncias estimadas. As margens de erro são calculadas para um nível de confiança de 95%. Isso indica que os resultados, baseados nessa amostra, são considerados precisos, dentro do intervalo definido pelas margens de erro. Se a pesquisa for repetida várias vezes, em 95% delas o intervalo poderá conter o verdadeiro valor populacional. Outras medidas derivadas dessa estimativa de variabilidade são comumente apresentadas, tais como erro padrão, coeficiente de variação ou intervalo de confiança.

O cálculo da margem de erro considera o produto do erro padrão (raiz quadrada da variância estimada) pelo valor 1,96 (valor da distribuição normal que corresponde ao nível de confiança escolhido de 95%). Esses cálculos são feitos para cada estimativa de cada uma das tabelas, o que significa que todas as tabelas de indicadores possuem margens de erro relacionadas às suas estimativas apresentadas em cada célula.

Disseminação dos dados

Os resultados desta pesquisa são divulgados de acordo com as seguintes variáveis de cruzamento: região e classe de número de acessos, conforme os dados SCM do período que precede a realização do estudo.

Arredondamentos fazem com que, em alguns resultados, a soma das estimativas das categorias parciais supere 100% em questões de resposta única. O somatório de frequências em questões de resposta múltipla usualmente ultrapassa 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 significam que houve resposta ao item, mas ele é explicitamente maior do que zero e menor do que um.

Os resultados desta pesquisa são publicados em formato *on-line* e disponibilizados no *website* (www.cetic.br) e no portal de visualização de dados do Cetic.br (<https://data.cetic.br>). As tabelas de proporções, totais e margens de erro 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 da pesquisa estão disponíveis na página de microdados (<https://www.cetic.br/microdados/>).

Referências

Bolfarine, H., & Bussab, W. O. (2005). *Elementos de amostragem*. São Paulo: Blucher. Cochran, W. G. (1977). *Sampling techniques* (3ª ed.). Nova Iorque: John Wiley & Sons.

Comitê Gestor da Internet no Brasil – CGI.br. (2018). *GT Marco Civil e as responsabilidades do CGI.br*. São Paulo: CGI.br Recuperado em 10 fevereiro, 2019, de <https://www.CGI.br/media/docs/publicacoes/4/GT%20Marco%20Civil%20e%20as%20responsabilidades%20do%20CGI.br.pdf>

Hansen, M. H., Hurwitz, W. N., & Madow, W. G. (1953). *Sample survey methods and theory*. Nova Iorque: Wiley.

Instituto Brasileiro de Geografia e Estatística – IBGE. (s.d.). *Pesquisa nacional por amostra de domicílios (Pnad)*. Recuperado em 9 setembro, 2016, de http://downloads.ibge.gov.br/downloads_estatisticas.htm

Kish, L. (1965). *Survey Sampling*. Nova Iorque: Wiley.

Lumley, T. (2010). *Complex surveys: a guide to analysis using R*. Nova Jersey: John Wiley & Sons.

Särndal, C., Swensson, B., & Wretman, J. (1992). *Model assisted survey sampling*. Nova Iorque: Springer Verlag.
