

# RELATÓRIO METODOLÓGICO TIC PROVEDORES

## INTRODUÇÃO

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), departamento do Núcleo de Informação e Coordenação do Ponto BR (NIC.br), apresenta a metodologia da pesquisa TIC Provedores.

Realizado pela primeira vez em 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 e identificar as necessidades e potencialidades das empresas de provimento de acesso à Internet no país, abordando questões 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 por meio de um censo, com a intenção de reunir o maior número de provedores de serviços de Internet (PSI) possível. Para a criação de uma listagem inicial, parte-se da base de informações do Sistema de Coleta de Informações (SICI), da Agência Nacional de Telecomunicações (Anatel), e da base de empresas cadastradas como possuidoras de Sistema Autônomo (do inglês, *Autonomous System* – AS) do NIC.br. O apoio de associações e outras organizações públicas e privadas, ligadas aos provedores de serviços de 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 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 outorga para a distribuição de Serviços de Comunicação Multimídia (SCM) e a 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 outorga 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.<sup>1</sup>

Dessa forma, outorgada como SCM, uma empresa pode distribuir informações multimídia, isto é, dados, mas 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 básica do setor, ao provedor que possui a licença SCM, é permitido distribuir pela sua infraestrutura o conteúdo de uma página da Internet, mas ele não está autorizado a transmitir, por exemplo, conteúdos exclusivos de canais de televisão ou permitir ligações telefônicas.

### SISTEMA AUTÔNOMOS

Um Sistema Autônomo 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).

<sup>1</sup> Mais informações no *website* da Anatel. Recuperado em 10 fevereiro, 2019, de <http://www.anatel.gov.br/setorregulado/comunicacao-multimidia-outorga>

## 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 médio de acessos SCM/Anatel:** é o conjunto de meios pelos quais a conexão à Internet é estabelecida. Vale lembrar que não se trata de assinantes, mas sim de pontos de conexão, os quais podem agregar vários assinantes.

## INSTRUMENTOS DE COLETA

### INFORMAÇÕES SOBRE OS INSTRUMENTOS DE COLETA

São aplicados dois questionários distintos. Para as empresas que possuem registros de acessos no período é aplicado o questionário completo. Já as que não possuem acessos declarados no período são contatadas e apenas respondem se são ou não provedoras de acesso à Internet, a fim de identificar aquelas que efetivamente são provedoras. Isso corresponde ao preenchimento de um questionário simplificado, com uma única pergunta sobre a elegibilidade da empresa. Para mais informações a respeito do questionário, ver item “Instrumento de Coleta” no “Relatório de Coleta de Dados” da pesquisa TIC Provedores.

## PROTOCOLO DE PESQUISA

### CADASTRO E FONTES DE INFORMAÇÃO

O Cetic.br conta com o apoio de diversas associações nacionais de provedores de acesso à Internet para a obtenção de informações a respeito dessas empresas. Para a construção do cadastro, são utilizados os dados existentes nas bases do SICI, da Anatel e do NIC.br – nesta última, estão informações referentes ao cadastro de empresas que possuem Sistema Autônomo.

O cadastro de provedores obtido no *website* da Anatel – base do SICI<sup>2</sup> – é dividido em duas partes distintas: empresas com informação de acessos entre um período determinado, conforme descrito no “Relatório de Coleta de Dados”, e empresas do cadastro sem informação de acessos nesse período. A essas duas partes são agregados os dados da base do NIC.br.

Esse conjunto total é dividido em três 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  
PARTIÇÃO DA POPULAÇÃO DE PESQUISA EM ESTRATOS PARA ORGANIZAÇÃO DA COLETA

Estrato	Descrição
1	Grandes provedores (aqueles com mais de 100 mil acessos)
2	Provedores médios e pequenos (aqueles com menos de 100 mil acessos)
3	Empresas sem declaração de acessos no período

A pesquisa é realizada separadamente nos três estratos, para que esforços de coleta e questionários 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 3 não têm a mesma probabilidade de serem provedoras ativas como as incluídas nos estratos 1 e 2.

## COLETA DE DADOS EM CAMPO

### MÉTODO DE COLETA

Os provedores com acessos declarados são classificados em dois grupos: grandes ou pequenos e médios. Para os grandes provedores, é feito um contato mais especializado e com um número maior de tentativas para que sejam efetivamente efetuadas as entrevistas, dado o peso inerente a suas respostas para a realidade do setor no país. Todas as empresas são contatadas por meio da técnica de Entrevista Telefônica Assistida por Computador (do inglês, *Computer Assisted Telephone Interviewing* – 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 recusem a responder a pesquisa pelo telefone. 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.

<sup>2</sup> Mais informações no *website* do SICI. Recuperado em 10 janeiro, 2019, de <https://sistemas.anatel.gov.br/sici/>

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 são classificadas em quatro grupos:

- Grupo 1: empresas provedoras de acesso à Internet com dados coletados;
- Grupo 2: empresas provedoras não respondentes;
- Grupo 3: empresas que não existem ou não são provedoras de acesso à Internet;
- Grupo 4: empresas para as quais não é possível identificar com segurança a sua existência ou se são provedoras de acesso à Internet.

O número de empresas provedoras da população-alvo é dado pelo total de empresas do cadastro inicial da pesquisa segundo a situação de coleta nas duas bases de investigação. Para cada uma delas, as empresas são divididas segundo a situação de contato:

- Grupos 1, 2 e 3: empresas contatadas;
- Grupo 4: empresas não contatadas.

O total de empresas provedoras é estimado para cada base somando-se o total de empresas classificadas nos grupos 1 e 2 ao total daquelas classificadas no grupo 4, multiplicado pela razão entre a soma das empresas consideradas provedoras (soma dos grupos 1 e 2) e empresas contatadas (soma dos grupos 1, 2 e 3).

$$\text{Total de empresas provedoras por base de investigação} = (N_1 + N_2) + \left[ N_4 \times \left( \frac{N_1 + N_2}{N_1 + N_2 + N_3} \right) \right],$$

onde:

$N_i$  é o total de empresas do grupo  $i$ .

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 declaração de acessos no período de referência, é feita correção de não resposta para os respondentes em cada estrato. Esse procedimento considera que aqueles provedores que não responderam à pesquisa (recusas diversas e desistências) são homogêneos em relação às informações prestadas pelos respondentes, dentro do 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 identificadas no estrato pelo total de empresas respondentes no estrato:

$$w_{ig} = \frac{N_g}{n_g},$$

onde:

$w_i$  é o peso do informante  $i$  do estrato  $g$ ;

$N_g$  é o total de provedores no estrato  $g$ ; e

$n_g$  é o total de provedores respondentes no estrato  $g$ .

### ERROS DE ESTIMAÇÃO

As medidas dos erros de estimação dos indicadores da TIC Provedores levam em consideração a correção empregada de não resposta por estratos. A divulgação dos erros de estimação é feita mediante apresentação das margens de erro calculadas para um nível de confiança de 95%. A ideia é que os valores das margens de erro podem ser usados para construir intervalos com limites definidos pela estimativa pontual – mais ou menos a margem de erro. Esses intervalos são tais que, se a pesquisa for repetida várias vezes em iguais condições, em cerca de 95% delas os intervalos de confiança assim definidos irão conter o valor populacional do parâmetro estimado. Outras medidas derivadas dessa estimativa de variabilidade são comumente apresentadas, tais como erro padrão ou coeficiente de variação.

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 médio 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* do Cetic.br (<http://www.cetic.br>) e no portal de visualização de dados da entidade (<http://data.cetic.br/cetic>). As tabelas de totais e margens de erros calculadas para cada indicador estão apenas disponíveis no *website* do Cetic.br.

## 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://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](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.