

RELATÓRIO DE METODOLOGIA

ESTIMATIVA DOS FLUXOS DE PASSAGEIROS
DO TRANSPORTE AÉREO



UNIVERSIDADE FEDERAL DE SANTA CATARINA - UFSC
LABORATÓRIO DE TRANSPORTES E LOGÍSTICA - LABTRANS
MINISTÉRIO DOS TRANSPORTES, PORTOS E AVIAÇÃO CIVIL - MTPA

APOIO TÉCNICO NO PLANEJAMENTO DO SETOR DE AVIAÇÃO CIVIL

**FASE 3 - DESENVOLVIMENTO DE UMA MATRIZ
ORIGEM/DESTINO DE PASSAGEIROS EM ÂMBITO NACIONAL
PARA IDENTIFICAÇÃO DOS REAIS DESEJOS DE VIAGEM
CONTEMPLADOS OU NÃO PELO TRANSPORTE AÉREO**

**Relatório de Metodologia - Estimativa dos fluxos de
passageiros do transporte aéreo**

FLORIANÓPOLIS, DEZEMBRO/2018

Versão 1.0

HISTÓRICO DE VERSÕES

Data	Versão	Descrição	Autor
14/12/2018	1.0	Entrega da primeira versão do Relatório de Metodologia para a estimativa dos fluxos de passageiros do transporte aéreo envolvida nas atividades da Fase 3	LabTrans/UFSC

Sumário

Introdução	5
1. O planejamento de sistemas de transporte	7
2. Coleta de dados.....	9
2.1. Delimitação da área de estudo e zoneamento	9
2.2. Padrões de viagens	10
2.3. Formas de obtenção dos dados	11
2.3.1. Entrevistas	11
2.3.2. Contagens volumétricas	12
2.3.3. Dados produzidos por aparelhos com registro de localização	12
2.4. Matriz Origem/Destino	13
3. Big Data	15
3.1. O que é Big Data.....	15
3.2. Vantagens e aplicações	17
3.3. Privacidade dos dados e transparência quanto à utilização	18
4. O uso do Big Data no planejamento de transportes	19
4.1. Estudos realizados.....	19
4.1.1. Projeto BigData4ATM.....	20
4.1.2. Secretaria de Turismo do Estado do Espírito Santo	20
4.1.3. Metro Corredor Kansas City	21
4.1.4. Boston	22
4.1.5. Sudoeste da Alemanha.....	23
5. Metodologia	25
5.1. Coleta dos dados de localização de usuários de telefonia móvel.....	30
5.2. Validação dos dados.....	32
5.3. Processamento.....	32
5.3.1. Dos dados à matriz O/D	32
5.3.2. Da matriz O/D às análises.....	33
5.4. Expansão	33
5.5. Análise.....	34
5.5.1. Áreas de influência dos aeródromos.....	34
5.5.2. Linhas de desejo de deslocamentos.....	36

5.5.3. Potencial de migração para o modo aéreo	36
Considerações finais	39
Referências	41
Lista de abreviaturas e siglas.....	45
Lista de figuras	47
Lista de tabelas.....	47
Anexo – Termo de Referência	49

Introdução

A Secretaria Nacional de Aviação Civil do Ministério dos Transportes, Portos e Aviação Civil (SAC/MTPA) está desenvolvendo, em cooperação com o Laboratório de Transportes e Logística da Universidade Federal de Santa Catarina (LabTrans/UFSC), estudos e pesquisas para apoio técnico no planejamento do setor de aviação civil brasileiro.

Em atendimento ao escopo definido para a Fase 3 do Plano de Trabalho do Projeto de Extensão intitulado “Apoio Técnico no Planejamento de Investimentos do Setor de Aviação Civil” – previsto no Termo de Execução Descentralizada nº 01/2018 celebrado entre a UFSC e a SAC/MTPA, relativa à elaboração de uma matriz Origem/Destino (O/D) de passageiros em âmbito nacional para identificação dos reais desejos de viagens contemplados ou não pelo transporte aéreo, este relatório visa apresentar a metodologia abarcada nas atividades de coleta de dados necessários, de elaboração da matriz O/D e das análises a serem realizadas.

As informações requeridas para a estimativa da matriz O/D serão obtidas por meio de dados de Big Data, que possibilitarão o entendimento do comportamento dos indivíduos nos deslocamentos entre as zonas de interesse da SAC/MTPA, as Unidades Territoriais de Planejamento (UTPs). A formação dessa matriz apresentará os principais fluxos de viagem, ponto a ponto, dos indivíduos que utilizam a malha aérea.

Este relatório se divide em tópicos sobre: o planejamento de sistemas de transporte, coleta de dados, Big Data, uso de Big Data no planejamento de transportes e, por fim, a metodologia abrangendo dos dados de entrada (provenientes de Big Data) aos produtos (matriz O/D e análises).

1. O planejamento de sistemas de transporte

O planejamento de sistemas de transporte molda as políticas, estratégias e programas de transporte, buscando um sistema integrado que movimenta pessoas e bens de maneira eficiente (BRUTON, 1979; CAMPOS, 2013).

Em geral, costuma-se organizar o processo de planejamento em três níveis: estratégico, tático e operacional. Tais níveis são descritos a seguir de acordo com Pereira (2005).

- Nível estratégico: o planejamento no nível estratégico diz respeito às ações (estratégias) de longo prazo e à execução de programas para o desenvolvimento do sistema de transportes de forma integrada, englobando o planejamento das principais redes de transporte do sistema, bem como de sua infraestrutura e terminais. Neste nível, deve-se desenvolver as metas da região e as políticas do estado e do país (CARTER; HOMBURGER, 1978 apud PEREIRA, 2005).
- Nível tático: no planejamento a nível tático, as análises são de médio a longo prazo, e as tarefas podem ser inerentes às diretrizes do plano de estratégia ou provenientes de decisões baseadas em falhas operacionais.
- Nível operacional: por fim, a nível operacional, estão as ações de curto prazo, para as quais são necessárias informações detalhadas acerca do objeto analisado.

A Tabela 1 apresenta exemplos de ações para os níveis de planejamento citados, considerando-se sua aplicação ao transporte aéreo.

Tabela 1 – Atividades relativas aos diferentes níveis de planejamento

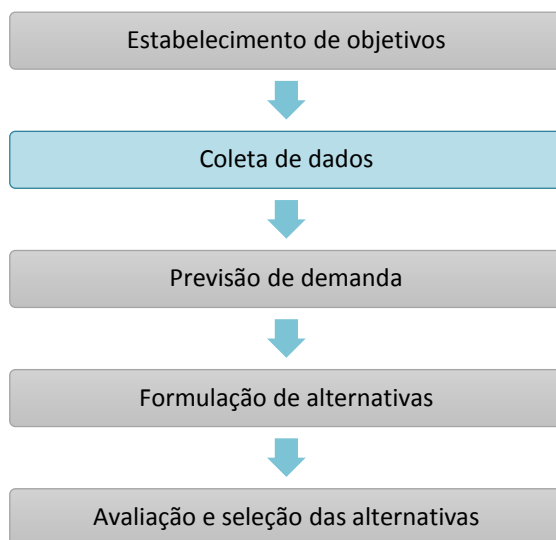
Nível de planejamento		
Estratégico	Tático	Operacional
Macrossimulação	Macro ou microssimulação	Microssimulação
Análise da demanda reprimida	Projeto geométrico da infraestrutura do Lado Ar	Otimização do controle de tráfego aéreo
Análise e escolha de investimentos em aeródromos específicos para atração de novas linhas aéreas	Projetos de melhoria da infraestrutura do Lado Terra	Investimentos de melhoria no atendimento ao usuário pela aquisição de novos ou melhores equipamentos e/ou treinamento ou contratação de novos funcionários
Modificações na rede de transporte aéreo.	Projetos de implantação de auxílios à navegação aérea	
	Elaboração de planos diretores aeroportuários.	Otimização de movimentação interna do aeródromo

Elaboração: LabTrans/UFSC (2018)

No processo do planejamento, deve-se identificar a demanda atual por transportes, a previsão de demanda para horizontes de planejamento, a formulação de alternativas para o atendimento a essa demanda, a verificação da viabilidade técnica e econômica das várias alternativas formuladas e, por fim, a seleção das melhores alternativas.

Para isso, o processo de planejamento de sistemas de transportes preconiza as etapas expostas na Figura 1 – Processo de planejamento de transportes, na qual se destaca a etapa de coleta de dados, que consiste no objetivo da metodologia para estimativa de fluxos de passageiros do transporte aéreo proposta neste relatório.

Figura 1 – Processo de planejamento de transportes



Elaboração: LabTrans/UFSC (2018)

- Estabelecimento de objetivos: os objetivos podem ser econômicos ou não econômicos e também variam conforme o nível de planejamento.
- Coleta de dados: coleta de dados sobre as características e padrões das viagens atuais na área em estudo.
- Previsão de demanda: por meio de modelos específicos, a demanda atual é projetada para horizontes de planejamento e alocada no modo de transporte em foco.
- Formulação de alternativas: são elaboradas propostas de modificações, principalmente em infraestrutura, para aumento de capacidade.
- Avaliação e seleção das alternativas: as alternativas são avaliadas e selecionadas considerando-se os objetivos elencados.

Visto que a atividade proposta está relacionada à coleta de dados, essa etapa é abordada de forma detalhada no capítulo que segue.

2. Coleta de dados

O processo de planejamento de sistemas de transporte envolve a utilização de uma quantidade considerável de dados. Embora a quantidade dos dados a coletar e os detalhes em que estes se apresentam variem com o nível de planejamento, propósito e o tamanho da área em estudo, os princípios básicos que devem ser seguidos na fase de coleta de dados são semelhantes nos vários estudos (BRUTON, 1979).

As diferentes pesquisas necessárias para a coleta dos dados básicos consomem tempo e exigem expressivos recursos humanos e financeiros; desse modo, é necessária uma cuidadosa programação na fase de coleta de dados, sendo prática corrente analisar com detalhes cada elemento no processo (BRUTON, 1979).

Segundo Bruton (1979), os objetivos da fase de coleta e análise de dados no processo de planejamento de sistemas de transporte são:

- Determinar onde começam e terminam as viagens;
- Determinar os fatores que influenciam a geração de viagens;
- Estabelecer os principais corredores de movimento.

Para cumprir com os objetivos, as etapas de delimitação da área de estudos, o zoneamento, a definição dos padrões de viagem estudados e a forma de obtenção dos dados devem ser definidos.

2.1. Delimitação da área de estudo e zoneamento

A área de estudo deve englobar todos os movimentos importantes da região e incluir as áreas que serão desenvolvidas no futuro, dentro do período para o qual se planeja. A linha externa que limita a área de estudo é conhecida como cordão externo ou *cordon line* (BRUTON, 1979).

Para possibilitar a análise dos movimentos internos na área de estudo, faz-se uma subdivisão da área em zonas de tráfego. As zonas representam regiões com características homogêneas ou atributos específicos que podem ser utilizados como variáveis explicativas do crescimento do tráfego, tais como PIB, população, empregos, entre outros. Assim, a zona de tráfego é a unidade-base de análise, e as características da área de estudo são pesquisadas e analisadas ao nível individual dessas unidades (BRUTON, 1979; CAMPOS, 2013).

No geral, a divisão da área de estudo em zonas de tráfego visa a atender às seguintes finalidades (BRASIL, 2006):

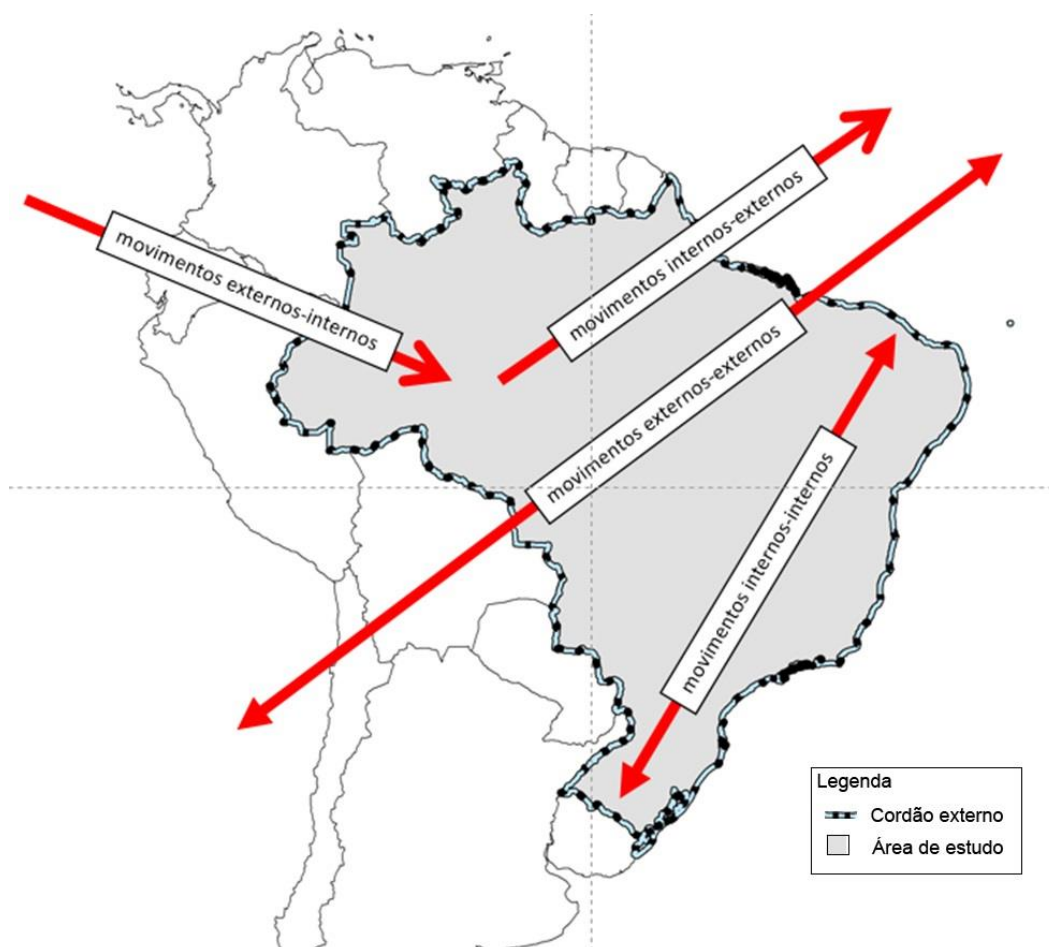
- Agrupar os dados de viagens de modo a reduzir os números de origens e destinos a serem considerados, simplificando desta forma a distribuição da movimentação.
- Fornecer a base para a determinação das viagens atuais e futuras, necessária à estimativa da projeção da demanda.
- Permitir o tratamento estatístico dos fatores de geração de tráfego em termos de regiões homogêneas.

2.2. Padrões de viagens

Os estudos para planejamento de sistemas de transporte têm como objeto fundamental a viagem. Define-se viagem como o percurso entre dois pontos com objetivo específico. Estes pontos são chamados de extremos de viagem, sendo que o extremo de saída chamamos “origem” e o de chegada “destino” (GERMANI et al., 1973).

É comum dividir as viagens em quatro movimentos relacionados à área de estudo (BRUTON, 1979), conforme se observa na Figura 2, que considera o Brasil como área de estudo.

Figura 2 – Padrões de viagem



Fonte: Bruton (1979). Elaboração: LabTrans/UFSC (2018)

- Movimentos externos-externos: com uma origem e um destino fora da área definida pelo cordão externo. Por exemplo, viagens com destino e origem internacionais. Devido à característica do escopo do estudo, essas movimentações podem não ser objeto de análise. São conhecidas também como fluxos de passagem.
- Movimentos externos-internos: com uma origem fora do cordão externo e um destino dentro dele. Pode-se citar como exemplo, no modo de transporte aéreo, passageiros/voos com origem internacional, mas que desembarcam no Brasil.

- Movimentos internos-externos: com origem dentro da área definida pelo cordão externo e destino fora desta área. Constituem um exemplo disso os passageiros/voos com origem nacional e destino a algum país estrangeiro.
- Movimentos internos-internos: com origem e destino dentro da área limitada pelo cordão externo, por exemplo, passageiros/voos com origem e destino dentro do território nacional.

2.3. Formas de obtenção dos dados

Para a elaboração da matriz O/D, faz-se necessário obter dados que informem a origem e o destino das viagens. Deve-se objetivar saber o padrão atual dos deslocamentos realizados, fornecendo dados essenciais sobre os atuais desejos e hábitos de viagem, os quais, em conjunto com estudos econômicos, formam a base para a projeção dos padrões de viagem futuros (BRUTON, 1979).

Além da quantidade de viagens, outras informações podem se fazer necessárias a depender do foco do estudo, tais como:

- Característica/perfil do passageiro;
- Identificação da viagem;
- Local de real origem e real destino da viagem;
- Horário e duração da viagem;
- Motivos de viagem;
- Modos da viagem.

Outro aspecto importante na coleta de dados é a escolha do método utilizado para a sua realização. Entre as principais formas de obtenção dos dados estão as entrevistas, contagem de veículos e, mais recentemente, o uso de dados produzidos por aparelhos com registro de localização (CAMPOS, 2013). Salienta-se que os dados podem ser obtidos de mais de uma fonte.

Após o processamento dos dados das diversas fontes e a expansão da amostra, gera-se como resultado a matriz O/D da área de estudo.

2.3.1. Entrevistas

O planejamento da coleta de dados por entrevista se inicia pelo estudo amostral. Devem ser aplicadas técnicas para selecionar amostras estatisticamente válidas. Uma vez completa a seleção da amostra, inicia-se a coleta de dados relativos aos movimentos gerados. Para isso, são usados entrevistadores de campo especialmente treinados, atuando sob estreita supervisão de técnicos qualificados. Na entrevista, são feitas perguntas sobre todos os movimentos realizados, relativos à origem, ao destino, ao propósito e à forma de realização da viagem (BRUTON, 1979; ORTÚZAR e WILLUMSEN, 2011).

Além dos questionamentos quanto aos movimentos realizados, há também as pesquisas de preferência revelada ou declarada. Na revelada, os dados são obtidos a partir de situações reais ou observadas pelo indivíduo. Na declarada, são obtidas informações sobre deslocamentos a partir de situações hipotéticas. Esse tipo de pesquisa permite estabelecer curvas de demanda para sistemas ainda não implementados ou avaliar a sensibilidade da população diante de alguma estratégia hipotética (CAMPOS, 2013).

As entrevistas podem ser feitas no domicílio, na própria via, dentro do veículo e em locais estratégicos como pontos de parada do sistema de transporte, rodovias de maior fluxo, terminais rodoviários ou aeroportuários. Para as viagens que utilizam o modo aéreo, por exemplo, a Empresa de Planejamento e Logística (EPL) (2014) realizou entrevistas presenciais individuais nas áreas de embarque nacionais e internacionais em 65 aeroportos.

Para obter uma amostra representativa da utilização do transporte aéreo em todas as épocas do ano, os trabalhos podem ser divididos em quatro períodos, sendo a amostra calculada e distribuída de forma proporcional ao número de passageiros embarcados por rota. Esse tipo de pesquisa é válido quando se buscam dados apenas de um modo de transporte, possuindo, assim, a desvantagem de não coletar dados sobre as viagens entre zonas de tráfego, realizadas por outro modo de transporte.

2.3.2. Contagens volumétricas

As contagens volumétricas visam determinar a quantidade, o sentido e a composição do fluxo de veículos/pessoas que passam por um ou vários pontos selecionados da rede de transporte, numa determinada unidade de tempo.

Os dados de contagens volumétricas podem ser usados para a expansão da matriz O/D ou para estimar uma matriz sintética. O método consiste na obtenção de matrizes O/D a partir dos fluxos observados nos *links* selecionados que conectam diferentes zonas (BERA; RAO, 2011).

2.3.3. Dados produzidos por aparelhos com registro de localização

Segundo Ortúzar e Willumsen (2011), os avanços principalmente nos aparelhos de telecomunicação e GPS estão mudando a forma de coleta de dados de viagens. Esse tipo de coleta de dados oferece vantagens específicas no rastreamento de movimentos por longos períodos de tempo.

Conforme Alexander et al. (2015), a onipresença dos telefones móveis, juntamente com o avanço da tecnologia, facilitou o armazenamento dos registros de uso que contêm coordenadas com data e hora de clientes anônimos, fornecendo informações espaço-temporais sobre os padrões de deslocamento dos usuários. Além disso, os dados de localização oferecem “pegadas digitais” em uma escala que dificilmente pode ser obtida por meio das pesquisas por entrevista.

2.4. Matriz Origem/Destino

Dados de O/D das viagens representam uma das principais fontes de informação para o planejamento estratégico das redes de transporte. Com base nesses dados, pode-se organizar a informação sob a forma de matrizes O/D, que indicam a quantidade de viagens entre cada par de zonas de tráfego.

De modo geral, os totais de viagens são apresentados conforme mostra a Tabela 2. Cada elemento da matriz indica o número de viagens para um par de zonas, em um dos sentidos, isto é, há um elemento indicando o número de viagens com origem na zona *i* e destinadas à zona *j*, e outro elemento indicando as viagens originadas em *j* e destinadas a *i*.

Tabela 2 – Modelo de matriz O/D

Destino Origem	1	. .	j	. .	n	$\sum_{j=1}^n V_{i,j}$
1	$V_{1,1}$. .	$V_{1,j}$. .	$V_{1,n}$	G1
.
i	$V_{i,1}$.	$V_{i,j}$. .	.	Gi
.
n	$V_{n,1}$. .	$V_{n,j}$. .	$V_{n,n}$	Gi
$\sum_{i=1}^n V_{i,j}$	Ai	. .	Aj	. .	An	$\sum_{j=1}^n \sum_{i=1}^n V_{i,j}$

Fonte: Germani et al. (1973). Elaboração: LabTrans/UFSC (2018)

Onde:

$V_{i,j}$ = número de viagens originadas na zona i e destinadas à zona j

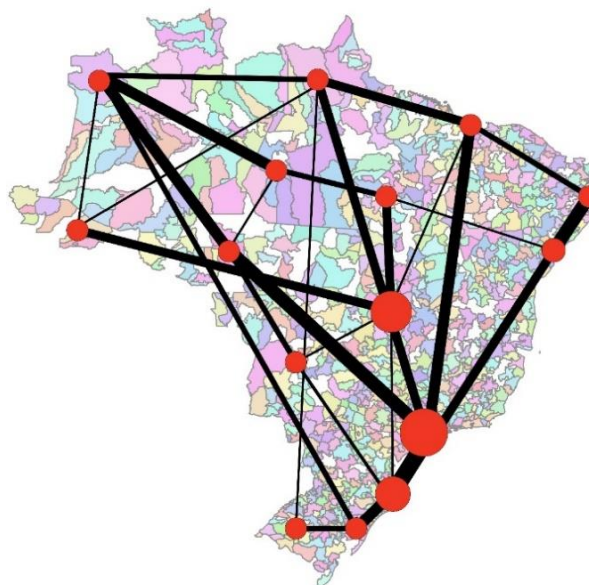
G_i = número total de viagens originadas na zona i

A_j = número total de viagens destinadas à zona j

n = número de zonas.

Outra representação da matriz O/D são os mapas de “linhas de desejo”, que mostram de maneira gráfica os volumes de viagens, ou seja, as necessidades de transportes entre todas as zonas de tráfego, permitindo identificar os eixos preferenciais de transportes da área de estudo, conforme a Figura 3.

Figura 3 – Exemplo fictício de mapa das linhas de desejo entre as zonas de tráfego



Elaboração: LabTrans/UFSC (2018)

3. Big Data

Neste capítulo, são apresentados o conceito de Big Data, as características, as vantagens, as aplicações e uma questão pertinente para sua utilização, a necessidade de privacidade dos dados e de transparência quanto à finalidade.

Vale salientar que a coleta de dados para elaboração de matrizes O/D originados de registros de localização de aparelhos de telecomunicação e Global Positioning System (GPS), abordados no item 2.3.3, envolve Big Data.

3.1. O que é Big Data

De acordo com Chen e Zhang (2014), Big Data é uma coleção de volumosos conjuntos de dados com uma diversidade de tipos tão grande que torna difícil seu processamento utilizando abordagens do estado da arte ou plataformas tradicionais de processamento de dados. O Big Data derivou do conceito de gerenciamento de dados 3-D formulado por Laney (2001), que consiste na exploração de dados nas dimensões volume, velocidade e variedade (3 Vs).

O volume diz respeito à vasta diversidade de fontes de dados e à consequente dificuldade de armazenamento. Segundo Sagioglu e Sinanc (2013), esses dados são gerados de transações *on-line*, *e-mails*, vídeos, áudios, imagens, fluxos de cliques, registros, publicações, consultas de pesquisas, registros de saúde, interações em redes sociais, dados científicos, sensores e telefones celulares e suas aplicações.

A velocidade, por sua vez, refere-se à coleta dos dados, muitas vezes se dando em tempo real e, assim, necessitando de um processamento de alto desempenho.

Por fim, quanto à variedade, há dados de três tipos: estruturados, semiestruturados e não estruturados, sendo que os primeiros se encontram armazenados com identificadores e são facilmente classificados, ao passo que os não-estruturados são aleatórios e de difícil análise (SAGIROGLU; SINANC, 2013). Os dados estruturados estão em bancos de dados tradicionais, para os quais é necessária a definição precisa de onde cada informação estará. Nos dados não estruturados, por sua vez, é necessária a intervenção humana para sua preparação. Por fim, os dados semiestruturados apresentam algum nível de estruturação intermediário. A variedade se refere, além do tipo dos dados, às fontes, dimensões e níveis de agregação.

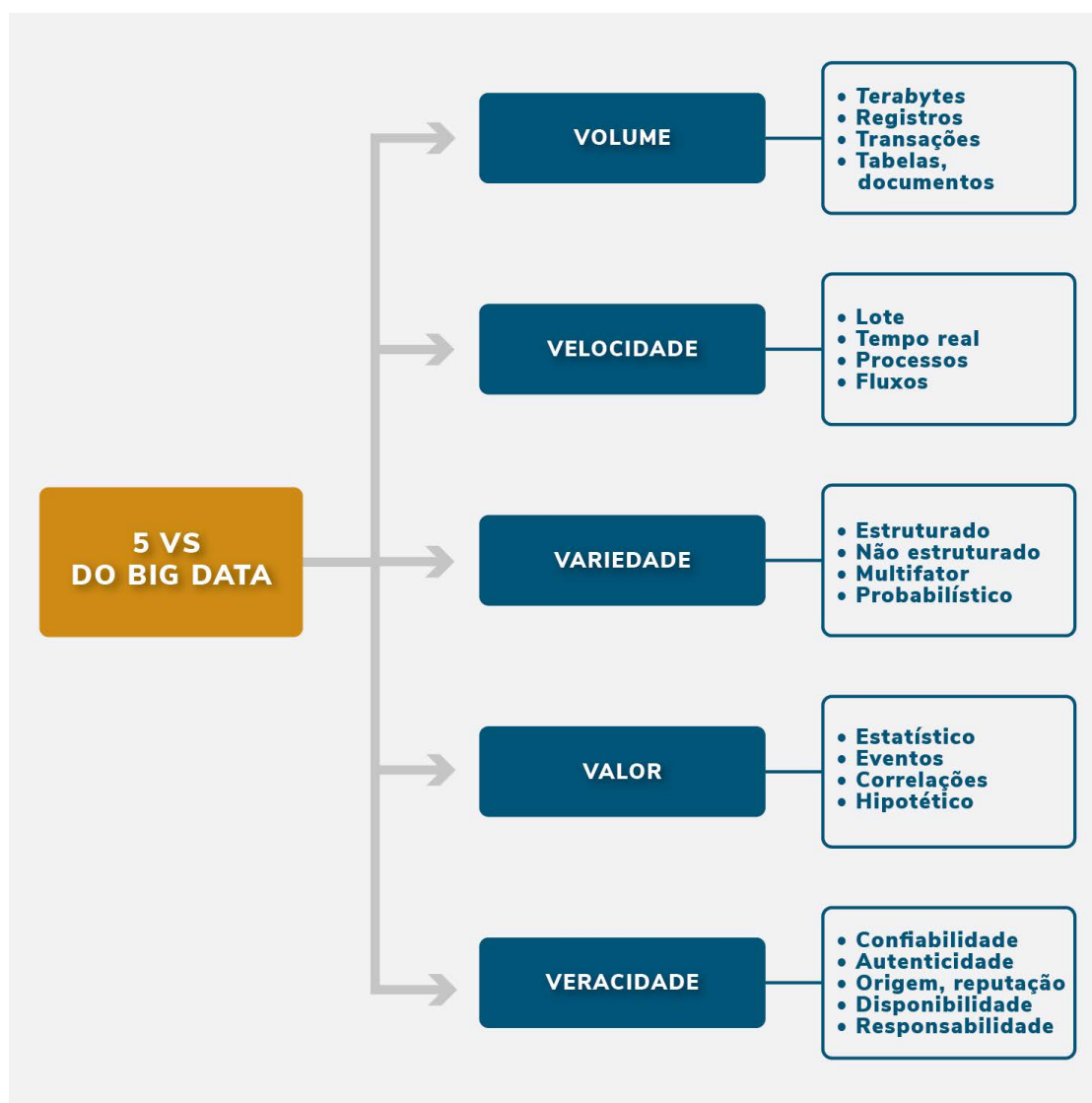
Atualmente, a comunidade científica e empresas de tecnologia podem considerar dimensões adicionais. Segundo Demchenko, Laat e Membrey (2014), além das três dimensões que constituem as propriedades nativas/originais do Big Data, há as dimensões valor e veracidade (surgindo, assim, a definição de propriedades dos 5 Vs), que resultaram da classificação inicial dos dados e do seu processamento no contexto de um processo ou modelo específico.

De acordo com Demchenko et al. (2013), o valor é definido pelo valor agregado que os dados coletados podem fornecer ao pretendido processo, atividade ou análise/hipótese de previsão, ao passo que, quanto à veracidade, deve-se assegurar que os dados utilizados sejam confiáveis, autênticos e protegidos de acesso e modificação não autorizada.

Assim, levando-se em conta a conceituação mais atual, o Big Data pode ser visto como um imenso **volume** de dados que apresentam **veracidade** quanto à sua origem e **variedade** quanto à sua estrutura e que surgem em torno dos negócios diariamente, permitindo-se agregar **valor** (*insights* produzidos na análise) quando existir **velocidade** para obtenção dos resultados.

A Figura 4 apresenta as características relacionadas aos 5 Vs, de acordo com Demchenko et al. (2013).

Figura 4 – 5 Vs do Big Data



Fonte: Demchenko et al. (2013). Elaboração: LabTrans/UFSC (2018)

A validade da utilização de Big Data, no entanto, não é intrínseca a essas características dos dados. Ela depende, principalmente, das perguntas feitas, das quais surgirão as definições das melhores fontes de dados e das melhores soluções para tratamento e análise. Em outras palavras, não é a quantidade de dados disponíveis (entre outras características) que importa, mas sim o seu uso.

Nesse sentido, a Figura 5 apresenta a estrutura de Big Data e a orquestração do processo.

Figura 5 – Estrutura de Big Data e orquestração do processo



Elaboração: LabTrans/UFSC (2018)

3.2. Vantagens e aplicações

Dados provenientes de Big Data são considerados dados de alto valor agregado aos negócios e organizações, uma vez que, através de um processamento eficiente e de inteligência analítica, os planejadores podem tomar decisões baseadas em evidências concretas. Chen e Zhang (2014) citam aplicações de Big Data em três principais domínios: comércio e administração, administração pública e campos de pesquisa científica.

O McKinsey Global Institute (MGI), em seu relatório *Big Data: the next frontier for innovation, competition, and productivity* (MANYIKA et al., 2011), apontou os potenciais impactos – em termos de benefícios financeiros para empresas e de valor para o usuário – do uso de Big Data para cinco domínios: dados baseados em localização, varejo americano, indústria, setor público europeu e assistência de saúde americana. Cinco anos depois, o MGI (HENKE et al., 2016) identificou que dados baseados em localização foram os que capturaram maior quantidade dos valores previstos em 2011, uma captura de 50 a 60%.

Uma das grandes possibilidades do Big Data é fornecer um melhor conhecimento de clientes/usuários e compreensão de seu comportamento e de suas preferências. Nos sistemas de transporte, os usuários são os indivíduos que realizam deslocamento através do(s) modo(s) de transporte de interesse. Informações de demanda provenientes de Big Data, portanto, podem oferecer maior assertividade das ações estratégicas no planejamento da infraestrutura de transportes, otimizando investimentos. Para tanto, os dados têm de ser tratados e analisados criteriosamente, conhecendo-se bem as questões a que devem responder.

3.3. Privacidade dos dados e transparência quanto à utilização

Uma grande preocupação a respeito do uso de Big Data é a privacidade dos usuários. Ainda que estes possam se beneficiar quando a empresa utiliza dados de Big Data com o objetivo de melhorar seus produtos, processos e serviços, é necessária precaução para garantir a privacidade dos usuários. Nesse sentido, as empresas que coletam e armazenam dados de interesse podem fornecê-los não de forma individual, contendo informações de identificação pessoal (PII – do inglês, *Personally Identifiable Information*), mas agregados e anonimizados. A anonimização se refere a um procedimento que faz com que os registros não possam ser associados, de forma direta ou indireta, a um indivíduo.

Além disso, é importante que haja transparência sobre a finalidade da utilização dos dados e resultados obtidos, a fim de que os usuários possam compreender o valor do compartilhamento de seus dados.

4. O uso do Big Data no planejamento de transportes

No contexto da coleta de dados no planejamento de transportes com o Big Data, avaliou-se a existência das diversas fontes de dados passíveis de serem utilizadas para a elaboração de matrizes O/D. Uma fonte de obtenção de informações relevantes do deslocamento de pessoas que vem sendo utilizada é a análise da pegada digital, por meio de registros dos usuários de telefonia móvel na rede de telecomunicações, de registros de GPS, entre outros dados gerados durante o seu deslocamento com informações de espaço e tempo. No que se refere à telefonia móvel, segundo a Agência Nacional de Telecomunicações (Anatel) (2018), no Brasil, para cada 100 habitantes, há 113,54 acessos em operação na telefonia móvel. De acordo com García et al. (2016), são vantagens da utilização de dados de telefonia móvel para entender o deslocamento dos usuários em comparação às pesquisas tradicionais:

- possibilidade de amostras grandes;
- os dados são coletados passivamente, eliminando a interação com os entrevistados, o que frequentemente resulta em respostas incorretas ou imprecisas;
- economia de custo, visto que os dados já são registrados pelas operadoras de telefonia móvel.

Em relação à economia de custo, Hemmings e Goves (2016) citam que pesquisas presenciais têm apresentado elevados custos e consumo de tempo para condução e autorização, e que há um grande potencial de redução desses custos (financeiro e de tempo) com a utilização de dados de telefonia móvel.

Tendo em vista a importância da matriz O/D para o planejamento de transportes e a necessidade de refletir a realidade com a abrangência territorial desejada, a utilização de Big Data é uma alternativa a ser explorada. A seleção das fontes de dados, por sua vez, depende da qualidade das informações que podem fornecer a respeito do deslocamento dos usuários.

4.1. Estudos realizados

Estudos já realizados, nacional e internacionalmente, foram pesquisados no intuito de se conhecer melhor como foi realizada a coleta de dados, utilizando Big Data, sobre deslocamentos de indivíduos. Percebeu-se que muitos deles tiveram uma abrangência geográfica reduzida, contemplando o deslocamento entre poucos pares O/D.

Por sua vez, a matriz O/D desejada para o planejamento do setor aéreo deve contemplar deslocamentos entre as 780 UTPs, distribuídas por todo o território brasileiro, totalizando uma grande quantidade de deslocamentos observados. No entanto, através de pesquisas e debates acerca dos estudos existentes, foi possível conhecer os critérios técnicos envolvidos que podem ser previstos também para o estudo pretendido junto à SAC/MTPA.

A seguir, são descritos alguns desses estudos e seus aspectos mais relevantes.

4.1.1. Projeto BigData4ATM

Hemmings e Goves (2016) descrevem o estudo realizado no projeto europeu BigData4ATM, que teve como objetivo obter informações de deslocamento (porta a porta) entre as cidades de Madrid e Barcelona pelos modos de transporte rodoviário, ferroviário – através do *High Speed Train* (HST) – e aéreo.

Devido à localização do terminal do HST, mais central em relação à localização do aeroporto, esse modo de transporte se tornou bastante competitivo com o aéreo, apesar da velocidade de deslocamento deste último ser mais atrativa. O tempo de deslocamento até o aeroporto, somado ao tempo do deslocamento aéreo, para muitos usuários, é próximo ao tempo de deslocamento por meio do HST.

Quanto à metodologia, foram coletados dados durante vários meses, com identificadores anônimos dos usuários, que incluem dados sociodemográficos (idade e sexo) e dados espaço-temporais de registros para atividades como chamadas, SMS e transferência de dados. O modo de transporte utilizado foi identificado pela localização dos registros, com base em uma análise das antenas existentes na rota preferencial de cada modo de transporte considerado.

A amostra foi reduzida de acordo com o critério de frequência de registros, para levar em conta a granularidade mínima necessária dos dados. Foram descartados, também, os usuários que não realizaram atividades em ambas as cidades de Madrid e Barcelona e, portanto, que não realizaram deslocamento entre elas.

O estudo prevê a implementação de algumas melhorias, a saber: introdução de critérios de tempo de viagem e/ou análise conjunta dos dados de telefone com escalas de voo, trem e ônibus.

4.1.2. Secretaria de Turismo do Estado do Espírito Santo

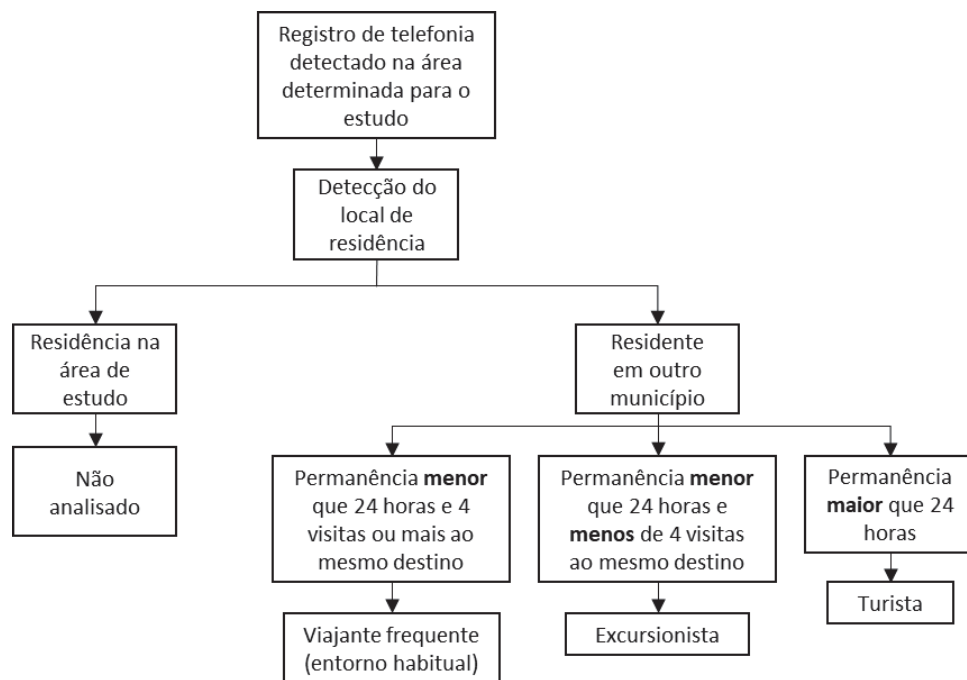
Uma experiência nacional da utilização de Big Data em mobilidade foi realizada pela Secretaria de Turismo do Espírito Santo (2017), em sua *Pesquisa de demanda e fluxo turístico por meio de sinal de telefonia móvel no Estado do Espírito Santo*.

Essa pesquisa buscou levantar uma matriz O/D turística com base no rastreamento (*tracking*) dos indivíduos por meio de registros na rede de telecomunicações. Foram analisados dados de origem e destino, local de estadia e tempo de permanência.

De acordo com recomendações da Organização Mundial do Turismo (OMT) (PAKMAN, 2014), definiu-se nessa pesquisa que os sinais de telefonia móvel teriam de ser detectados por no mínimo 40 minutos a no máximo uma hora dentro da área delimitada pelo raio de ação de uma única antena, sem que este seja o local de residência do indivíduo. Com isso, evitou-se que os indivíduos em deslocamento (de passagem) fossem considerados na pesquisa. Quando o tempo de permanência fosse superior ao período estipulado, classificar-se-ia o usuário como viajante.

A Figura 6 apresenta o processo de filtragem do perfil dos usuários utilizado.

Figura 6 – Processo de filtragem do perfil dos usuários



Fonte: Secretaria de Turismo do Espírito Santo (2017)

Foi efetuada, também, a extrapolação dos dados para representar toda a população da região analisada, correspondente ao local de residência do viajante, ou seja, sua origem.

4.1.3. Metro Corredor Kansas City

De acordo com o estudo de Wang et al. (2013), entender o fluxo de tráfego de origem/destino e a demanda de viagem entre cidades é uma tarefa essencial e primordial para o planejamento e o gerenciamento estratégico.

Observou-se que os métodos tradicionais de coleta de dados de origem/destino, geralmente, dependem de pesquisas telefônicas domésticas, dados do censo ou pesquisas nas vias. Com os avanços recentes em redes sem fio e dispositivos móveis, surgiu o interesse em se utilizar o celular, que tem a maior taxa de penetração e é um dispositivo em potencial para ser usado em rastreamento de veículo e viajante, para o fornecimento de informações sobre redes veiculares.

Neste estudo, as etapas da utilização de algoritmos para rastreamento de celulares se dividem em sete módulos, a saber:

- Aquisição de dados e localização: estabelece um banco de dados inicial para processos de rastreamento e filtragem.
- Acompanhamento de dados e filtragem: são averiguados os dados de trajetória e outros parâmetros operacionais para cada telefone celular e eliminados quaisquer dados de localização errôneos.
- Identificação do estado de movimento: identifica se os telefones celulares estão em condição de movimento ou não, tendo como critério de classificação o tempo de permanência.

- Determinação da viagem e origem/destino: é baseada nos estados de movimento dos telefones celulares, a partir de um valor limite para identificar se os telefones celulares que não estão se movendo estão temporariamente parados ou se os usuários estão permanecendo na região determinada. Esse valor varia conforme as áreas das regiões dadas podendo ser configurado para ser mais de 6 horas, a fim de identificar se os usuários dos dispositivos são residentes ou funcionários.
- Mapeamento: pode ser processado em qualquer módulo acima, desde que os pontos de dados de localização estejam disponíveis para inspecionar possíveis pontos inválidos e, também, demonstrar a distribuição espacial dos pontos de dados.
- Caracterização do telefone celular: todos os dados são criptografados. No entanto, a classificação e o tipo de viajante podem ser identificados pela observação da trajetória, origem e destino dos usuários do dispositivo.
- Derivar O/D dinâmico e demanda de viagem: quando a origem e destino são determinados, o tempo base do tráfego de dados O/D pode ser obtido somando-se o número de usuários de telefones celulares que viajam da região de origem para a região de destino por intervalos de tempo.

Neste estudo, aborda-se um caso de estudo em Kansas City, onde trabalham muitas pessoas que residem nas cidades adjacentes. Além disso, alguns funcionários públicos trabalham em Topeka, mas moram perto de Lawrence ou de Kansas City, e sabe-se que essas três cidades formam um corredor de tráfego conhecido como Metro Corredor Kansas. Por conta disso, considerou-se a implantação de um sistema de transporte intercidades para reduzir o impacto desse tráfego.

O Metro Corredor Kansas City contém 262 locais equipados com dispositivos de sinal de telefonia móvel. A maioria dos pontos de sinal estão localizados dentro ou perto das áreas mais urbanas das maiores cidades da região e foram ligados através de torres de sinais ou antenas.

Em relação à metodologia, a área de estudo foi dividida em três regiões (A, B e C) e 870.042 telefones celulares foram rastreados em seis semanas de análise, sendo delimitado um tempo fixo para que houvesse uma avaliação adequada da transmissão do tráfego de cada região e, assim, fosse possível um estudo detalhado através dos dados móveis da população.

Neste estudo, o intuito foi analisar a viabilidade do uso de dados de telefone celular para rastrear atividades de tráfego nessas cidades, derivar o tráfego O/D e observar a demanda de viagens por horário e dados de tráfego pendulares ao longo do corredor, com base em uma observação de seis semanas. Após essa análise, foi constatado o potencial da utilização dos dados móveis na pesquisa desse estudo de caso.

4.1.4. Boston

O estudo de Alexander et al. (2015) tem como objetivo analisar, através de dados móveis, a origem/destino e hora da partida de viagens nas principais cidades dos Estados Unidos.

Quanto à metodologia, a pesquisa busca levantar uma matriz O/D mapeando sinais de celulares. São inferidas localizações, tais como: casa, trabalho e outros, dependendo da observação da frequência de deslocamento, dia da semana e momento do dia, a fim de representar a matriz O/D.

Essas viagens são multiplicadas por expansão de fatores como população censitária dividido pelo número de dias do usuário observado, distinguindo uma rota geral comum. Agregando viagens diárias dos usuários e interferências se consegue modelar deslocamentos comuns que podem servir de informação para o planejamento de sistemas de transporte e de investimentos. Possibilita, ainda, a exploração dos conjuntos de dados para ver onde poderão ajudar a melhorar os modelos de viagens urbanas existentes e os aplicativos relacionados ao congestionamento.

Neste estudo, foi realizado o mapeamento de usuários de telefonia móvel na área metropolitana de Boston, confrontando-se os registros na rede com pesquisas locais e nacionais para fins de validação.

O estudo destaca, ainda, a importância do processamento dos dados, com grande parte dos métodos utilizados tendo por objetivo descartar o ruído e extrair informações acuradas de padrões de viagem representativos da área de estudo.

4.1.5. Sudoeste da Alemanha

Friedrich et al. (2010) apresentam uma metodologia para a geração de matrizes O/D, do transporte ferroviário e rodoviário, por meio de dados da telefonia móvel. Pelo deslocamento dos dispositivos é identificado o ponto de registro inicial e final de cada trajetória, determinando, assim, as zonas de origem e de destino.

Os conjuntos de dados utilizados no artigo foram registrados na rede da operadora T-Mobile, em uma área de estudo de 100 km x 100 km, no sudoeste da Alemanha, e abrange um período de, aproximadamente, 70 dias. O método utiliza principalmente dados de atualizações da área de localização, que são gravados a partir de qualquer telefone celular em modo de espera e, portanto, não requer chamadas telefônicas. Assim, é possível monitorar viagens integralmente.

Diante disso, são usadas três fontes de dados principais como entrada na metodologia utilizada para gerar a matriz O/D através de dados de telefone celular, a saber:

- um modelo de rede contendo a infraestrutura rodoviária e ferroviária relevantes e a posição dos locais de contagem;
- curvas de volumes de tráfego em intervalos de 15 minutos de locais de contagem na rede rodoviária (automóveis e caminhões) e na ferroviária (passageiros);
- trajetórias dos dados flutuantes de telefone (FPD – do inglês, *Floating Phone Data*).

O estudo indica que as matrizes O/D da FPD são ideais para calibrar modelos nacionais de demandas de viagens com zonas maiores. Elas são úteis, ainda, para prever o Nível de Serviço (LOS – do inglês, *Level of Service*) em redes rodoviárias de longa distância e para monitorar o desenvolvimento da demanda de uma nova infraestrutura de transporte implementada.

5. Metodologia

Diante da importância do conhecimento da demanda pelo transporte aéreo no planejamento estratégico do setor, buscou-se explorar o potencial do Big Data na coleta de dados para estimativa de fluxos de passageiros do transporte aéreo com o detalhamento necessário.

Quando o foco do planejamento se restringe à infraestrutura aeroportuária de forma individual, é fundamental se conhecer a demanda observada no aeroporto, expressa em número de passageiros embarcados e desembarcados, não havendo a necessidade de conhecimento das localidades de origem e de destino desses passageiros.

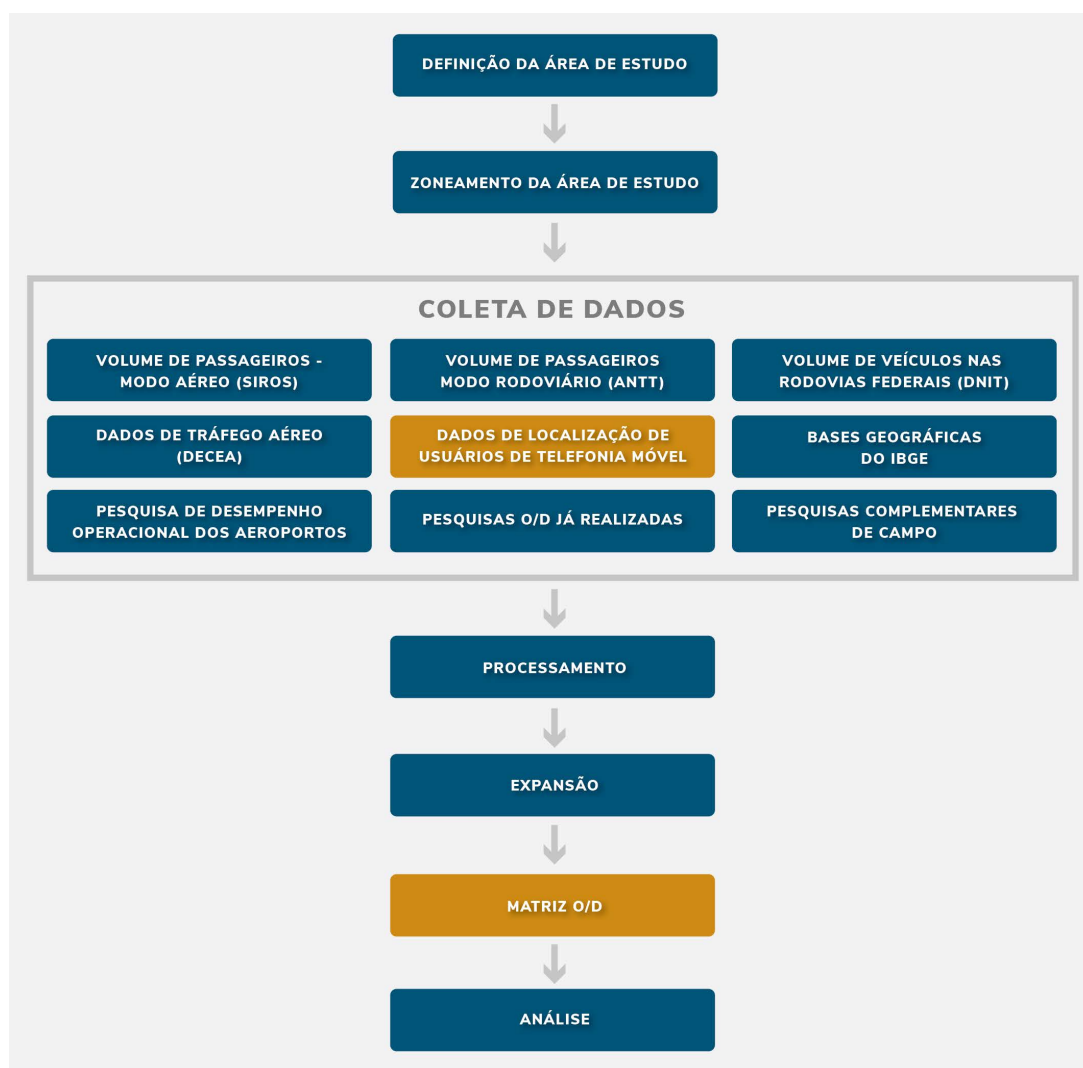
No entanto, quando o planejamento deve levar em consideração os impactos que os investimentos na infraestrutura aeroportuária podem ter na rede de transportes como um todo, é necessário se conhecer a demanda não somente em termos de passageiros processados nos aeroportos – informação que é facilmente obtida através dos dados abertos da Agência Nacional de Aviação Civil (ANAC), apresentados de forma interativa no Sistema Hórus, mas também em termos das características do deslocamento desejado ponto a ponto.

Sabendo as localidades de origem e de destino dos passageiros, juntamente com volumes de deslocamentos não aéreos, é possível identificar demandas reprimidas devido à ausência de aeroportos e/ou voos nessas localidades ou pela distância entre estas e aeroportos existentes, fatores que são determinantes na geração de demanda pelo transporte aéreo.

Dados provenientes de Big Data de registros de usuários de telefonia móvel podem fornecer, então, informações essenciais para a assertividade das ações estratégicas da SAC/MTPA no setor de transporte aéreo, com vistas a contribuir com a acessibilidade e conectividade da rede de transportes brasileira, objetivos estratégicos do Plano Aeroviário Nacional (PAN) 2018-2038 (BRASIL, 2018).

A Figura 7 apresenta o fluxograma da metodologia que será utilizada para a elaboração da matriz O/D da área de estudo, utilizando-se os dados de registros de usuários de telefonia móvel para estimativa dos fluxos dos passageiros.

Figura 7 – Fluxograma da metodologia



Elaboração: LabTrans/UFSC (2018)

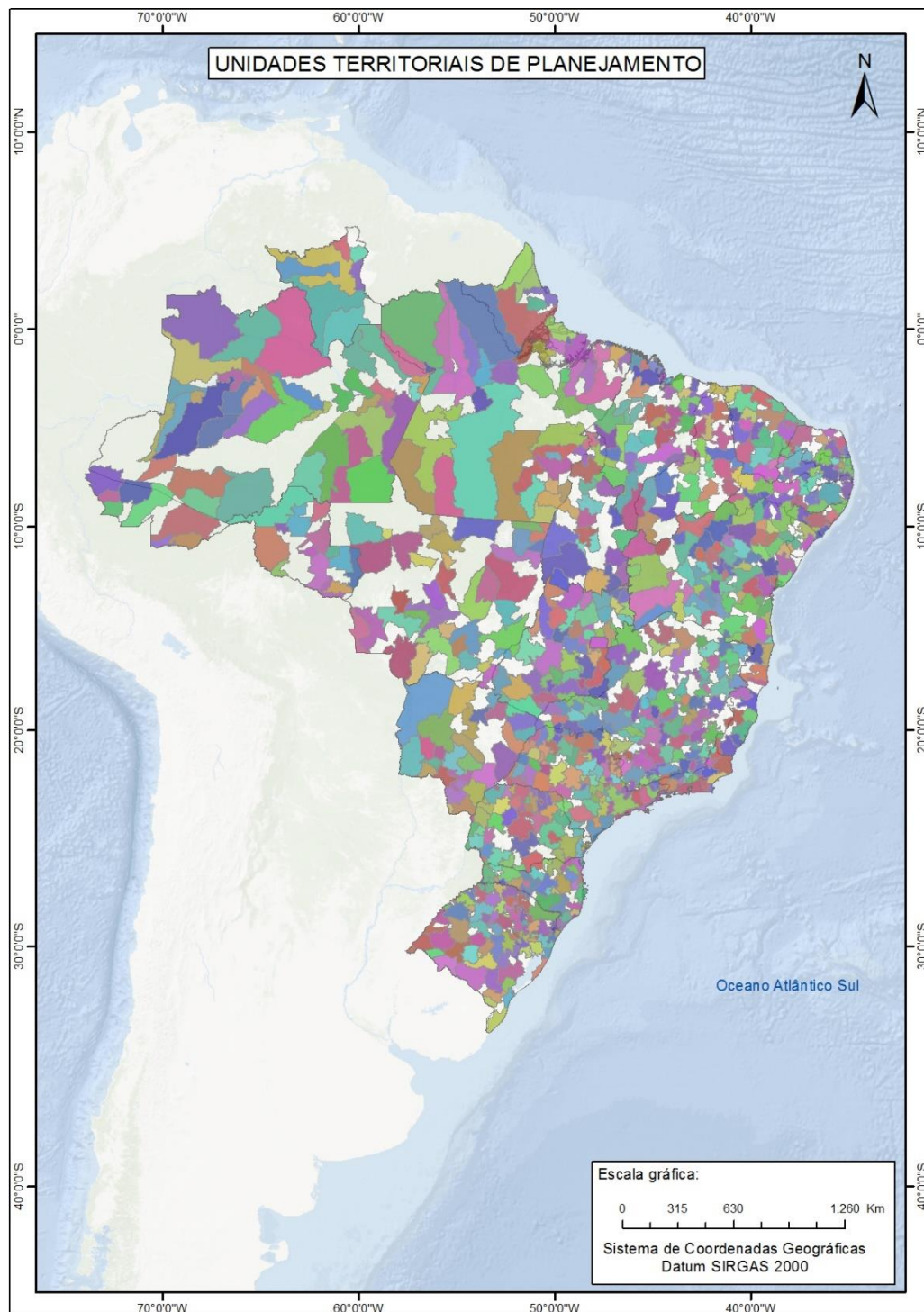
Levando-se em consideração as características dos estudos objeto deste Relatório de Metodologia, pode-se classificá-los como sendo um esforço de coleta de dados da demanda atual por transporte aéreo, que servirá como base para um planejamento de nível estratégico do setor aéreo nacional.

A área de estudo compreende, portanto, todo o território nacional. No que se refere às zonas de tráfego, na metodologia utilizada para as projeções de demanda para os aeroportos brasileiros 2017-2037, realizada pelo MTPA (BRASIL, 2017, p. 16), foram definidas as Unidades Territoriais de Planejamento (UTP), diante da necessidade de:

“(…) predefinir regiões onde se concentram a maior parte da demanda potencial e real pelo transporte aéreo, de modo que suas variáveis possam ser estudadas como explicativas no procedimento de modelagem, ou mesmo delimitar regiões onde já existem infraestruturas aeroportuárias instaladas, operantes ou não, que podem ser alvo de planejamento futuro e que representam regiões de concentração da população e, assim, do maior potencial de demanda pelo transporte aéreo”.

As UTPs foram definidas não só para as projeções de demanda, mas também para o PAN (BRASIL, 2018). De acordo com o Plano de Trabalho do projeto de cooperação entre a SAC/MTPA e o LabTrans/UFSC, elas devem ser utilizadas também como as unidades territoriais de referência para a elaboração da matriz O/D. As UTPs são apresentadas na Figura 8.

Figura 8 – Mapa das UTPs



Elaboração: LabTrans/UFSC (2018)

A definição das UTPs teve o objetivo de representar regiões de concentração da população e, assim, do maior potencial de demanda pelo transporte aéreo. De acordo com BRASIL (2017), os resultados da pesquisa origem/destino realizada pela SAC com a EPL permitiram identificar que, cerca de, 84% da demanda dos aeroportos faz parte da população das aglomerações urbanas imediatamente próximas a eles.

Assim, a partir da hipótese de que a maior parte da demanda pelo transporte aéreo está presente nas aglomerações urbanas em que há infraestrutura aeroportuária instalada, a definição das UTPs partiu das informações do estudo do Instituto Brasileiro de Geografia e Estatística (IBGE) intitulado *Arranjos populacionais e aglomerações urbanas* (IBGE, [2016]).

Somadas a esse primeiro recorte a partir dos arranjos populacionais e aglomerações urbanas, foram criadas novas UTPs, a fim de contemplar:

- municípios com aeródromos instalados segundo o cadastro de aeródromos públicos, sendo cerca de 640 à época em que foi consolidado o banco de dados (ANAC, 2016), e aeródromos constantes nos Planos Aeroviários Estaduais (PAE);
- municípios indutores do turismo de classes A ou B, segundo o Programa de Regionalização do Turismo (BRASIL, 2015).

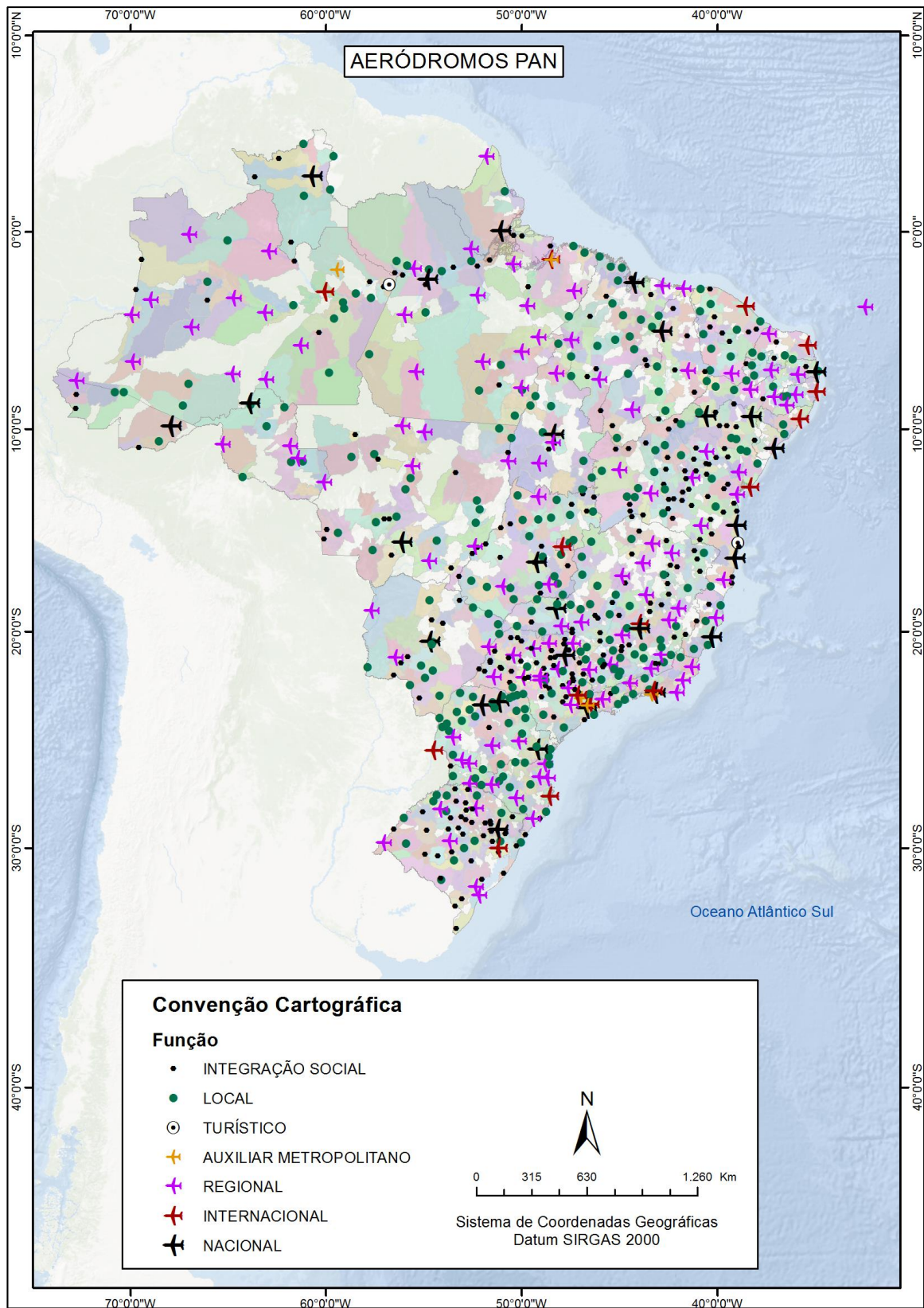
Além disso, foram realizados os seguintes ajustes:

- adequação de UTPs da Região Norte, de modo a considerar o transporte hidroviário, conforme base de dados cedida pela Agência Nacional de Transportes Aquaviários (ANTAQ) de 2013 referentes a uma pesquisa realizada sobre as ligações desse modo de transporte, a qual possibilitou a identificação de ligações pendulares;
- ampliação da abrangência territorial de determinadas UTPs considerando-se na definição da área de captação dos aeródromos a presença de municípios em isócronas de 30, 60, 90 e 120 minutos de deslocamento rodoviário aos aeródromos;
- ajustes para eliminar citações duplicadas de municípios em diferentes UTPs (sobreposições), tendo prevalência a influência de municípios-sede das UTPs cuja categoria de centralidade nas Regiões de Influência das Cidades (REGIC) é maior;
- ajustes realizados em ambiente SIG para eliminar as inconsistências nas formas das UTPs, a saber: municípios isolados de suas sedes e UTPs muito próximas, totalmente ou parcialmente inscritas em outras UTPs.

Desse processamento, resultaram 780 UTPs. Elas serão objeto do estudo de deslocamento dos indivíduos devido à sua relevância em termos de aglomerações urbanas existentes no país, de aeródromos instalados, indução turística e participação do modo hidroviário no transporte de pessoas na Região Norte do País.

Os aeródromos, por sua vez, são os definidos pela SAC/MPTA na base do PAN. Na Figura 9, são apresentados esses aeródromos no mapa do território brasileiro com as respectivas funções.

Figura 9 – Aeródromos da área de estudo



Elaboração: LabTrans/UFSC (2018)

Como parte do processo de desenvolvimento da matriz O/D para estimativa de fluxos de passageiros do transporte aéreo, será executada a coleta de dados de localização de usuários de telefonia móvel. As etapas de coleta e validação desses dados seguem descritas nos itens subsequentes, bem como o documento de Termo de Referência (TR) (Anexo 1) elaborado para a aquisição dos dados.

Essa aquisição será realizada através de subcontratação do fornecedor da base de dados, por meio de processo de licitação, prevista no Plano de Trabalho do projeto de cooperação entre a SAC/MTPA e o LabTrans/UFSC. O TR é um documento anexo ao edital dessa subcontratação e apresenta os parâmetros e as características necessárias para o desenvolvimento da base de dados de localização de usuários de telefonia móvel, bem como as qualificações técnicas e critérios de aceites da base, sendo considerado, assim, um documento complementar a este relatório de metodologia.

5.1. Coleta dos dados de localização de usuários de telefonia móvel

A coleta de dados consiste em mapear o deslocamento da população entre as UTPs, a fim de suportar a tomada de decisões de planejamento da SAC/MTPA para investimentos em infraestrutura aeroportuária.

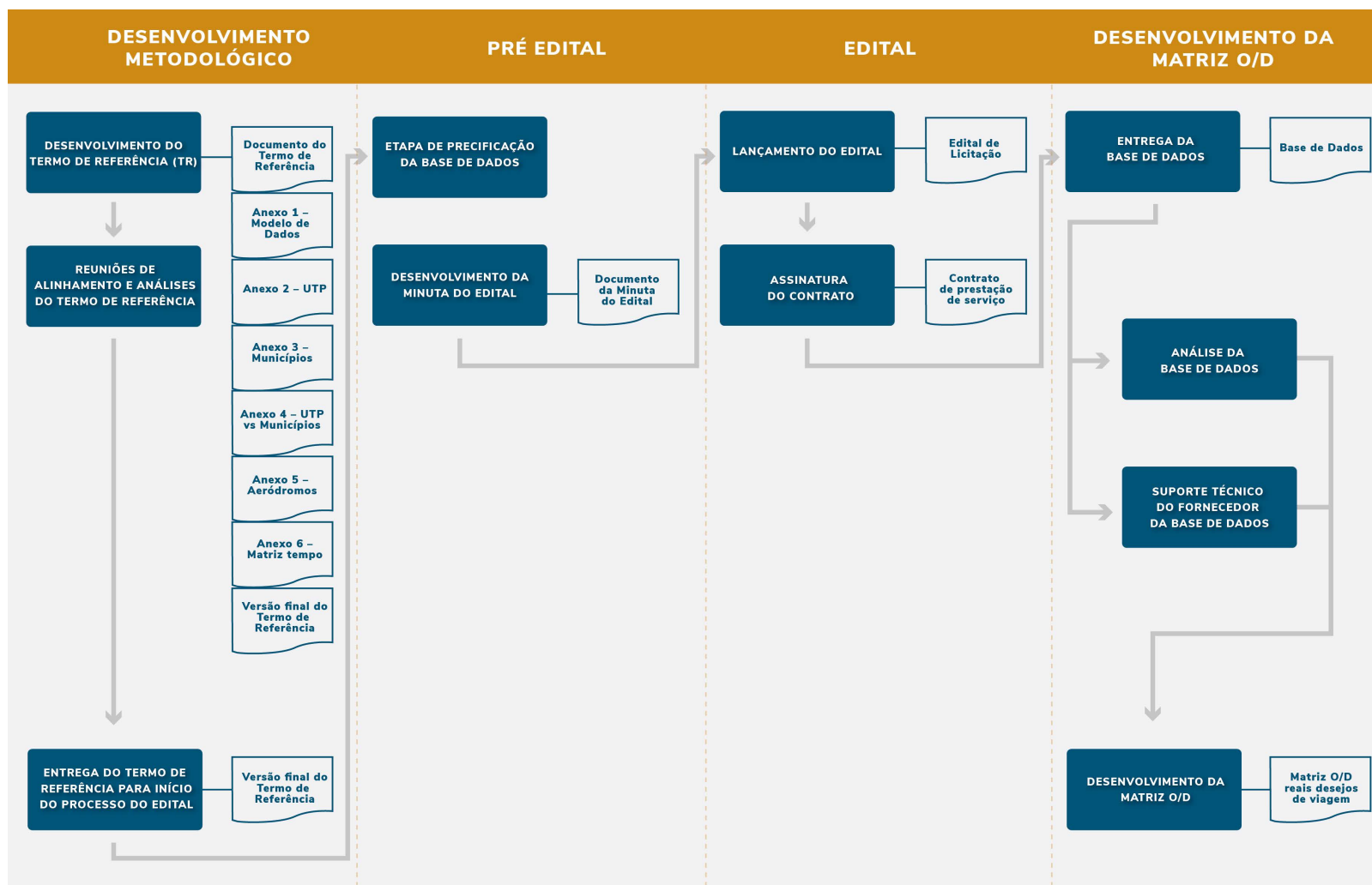
Estão elencados no TR (Anexo 1) os critérios técnicos a respeito da coleta de dados de localização dos usuários da telefonia móvel por parte da empresa a ser contratada por meio de licitação.

Os dados consistem em volumes de deslocamento, classificados pelo modo de transporte utilizado (aéreo ou não aéreo). O critério definido em conjunto com a SAC/MTPA para essa classificação foi o tempo de deslocamento, de acordo com uma velocidade de referência para a divisão entre aéreo e não aéreo, aliado à proximidade a aeródromos em um dos extremos da viagem. A identificação de uma localidade de presença do usuário como sua origem ou seu destino – UTPs ou municípios – se dará quando o registro permanecer estacionado por período superior a 6 horas, com base na metodologia de Wang et al. (2013). Os dados para obtenção dos volumes de deslocamento por parte da contratada podem ser provenientes de:

- 1) registros dos usuários de telefonia móvel;
- 2) registros de GPS;
- 3) outras potenciais fontes de dados acerca do deslocamento de indivíduos entre as UTPs.

Na Figura 10, estão mapeadas as etapas para a atividade de coleta de dados para desenvolvimento da matriz O/D.

Figura 10 – Fluxograma das etapas de coleta de dados e elaboração da matriz O/D



Fonte: LabTrans/UFSC (2018)

5.2. Validação dos dados

Está prevista no TR a validação dos dados por meio da contratante, contando com Suporte Técnico Especializado da contratada para esclarecimento de dúvidas e, no caso de verificação de alguma inconsistência, a resolução do problema com eventual atualização dos dados.

Para validação dos dados fornecidos e para as análises posteriores à sua elaboração, as seguintes fontes de dados podem ser levadas em consideração:

- Bases geográficas do IBGE;
- Volume de passageiros do modo aéreo, através de consulta ao Sistema de Registro de Operações (SIROS);
- Dados de controle de tráfego aéreo divulgados no “Anuário Estatístico de Tráfego Aéreo” do ano de 2017, publicado pelo CGNA/DECEA, que apresenta o total de movimentos por mês por tipo de voo para cada um dos 32 aeródromos pertencentes ao Sistema de Controle do Espaço Aéreo Brasileiro (SISCEAB);
- Resultados da “Pesquisa de Satisfação do Passageiro e de Desempenho Aeroportuário”, fornecidos pela SAC/MTPA;
- Volume de passageiros do modo rodoviário, proveniente das “Estatísticas e Estudos Rodoviários – Dados Operacionais” da ANTT, para transportadores regulares e fretados;
- Volume de veículos nas rodovias federais do “Plano Nacional de Contagem de Tráfego – PNCT”, proveniente do Departamento Nacional de Infraestrutura de Transportes (DNIT);
- Pesquisas O/D já realizadas, como da matriz O/D elaborada na Pesquisa de Origem/Destino do Transporte Aéreo de Passageiros realizada pela EPL (2014);
- Pesquisas de campo complementares.

5.3. Processamento

Neste item, são detalhadas as etapas de processamento dos dados fornecidos à elaboração da matriz O/D e, a partir dela, às análises relacionadas.

5.3.1. Dos dados à matriz O/D

Conforme estipulado no Termo de Referência (Anexo 1), os dados de localização geográfica a serem fornecidos deverão ser processados e expressos em número de viagens/deslocamentos por mês e por fluxo. Assim, o produto será uma planilha em que cada linha será correspondente a um fluxo de deslocamento observado e as colunas corresponderão aos seguintes atributos do dado:

- Identificador numérico do dado
- Ano de referência do dado
- Mês de referência do dado
- Identificador numérico da UTP de origem

- Identificador numérico da UTP de destino
- Modo de transporte
- Identificador numérico do município de origem
- Identificador numérico do aeródromo de embarque
- Identificador numérico do município de destino
- Identificador numérico do aeródromo de desembarque
- Quantidade de viagens observadas

Os identificadores numéricos dos municípios de origem e de destino e dos aeródromos de embarque e de desembarque são atributos que apresentarão valores somente se o modo de transporte for aéreo, conforme critério de classificação descrito no TR.

Com os dados fornecidos, será elaborada, considerando-se deslocamentos aéreos e não aéreos, uma matriz de dimensão 780 x 780. Ressalta-se que, para certos pares O/D, conforme processo de anonimização dos dados, não haverá um número mínimo de deslocamentos para que esse fluxo seja informado nos resultados fornecidos.

5.3.2. Da matriz O/D às análises

Os valores resultantes do processamento que serão apresentados nas matrizes O/D representam os fluxos com os maiores volumes de deslocamentos, ou seja, os fluxos de maior relevância em termos de desejos de deslocamento da população a partir das UTPs.

Para análise e validação dos dados produzidos para geração da matriz O/D e da própria matriz, será realizada a comparação com as fontes de dados citadas no item 5.2.

5.4. Expansão

Nas pesquisas tradicionais, os dados coletados em diferentes tipos de pesquisa de campo, por meio de procedimentos amostrais, devem ser expandidos de modo a representar a população, compreendendo, também, as áreas não cobertas pelas pesquisas realizadas. Isto é feito através do uso de fatores de expansão que são calculados para cada zona de interesse considerada na pesquisa.

Encontra-se estipulada no TR a cobertura geográfica mínima da empresa fornecedora dos dados e a necessidade de comprovação por parte da contratada de fornecimento de amostra mínima e nível de confiabilidade. Partindo-se dessa amostra, serão realizadas expansões dos dados para as áreas não abrangidas nesta. Para tal, pretende-se utilizar os dados da pesquisa da EPL, entre outras fontes (item 5.2) com potenciais dados complementares.

Pode-se levar em consideração, em uma expansão local, a participação de mercado do fornecedor dos dados na área, entre outros aspectos que possam interferir na representatividade.

5.5. Análise

Com a matriz O/D resultante, serão feitas análises que apresentam aspectos relevantes acerca dos deslocamentos dos indivíduos entre as UTPs, pelo modo de transporte aéreo e não aéreo.

Uma primeira análise tratará da identificação dos fluxos com os maiores volumes de deslocamento, independentemente do modo de transporte, os quais representam os principais desejos de deslocamento da população.

Para os deslocamentos aéreos, serão analisados os municípios de origem e de destino. Quando o município de origem não é onde está localizado o aeródromo de embarque e/ou o destino não é onde está localizado o aeródromo de desembarque, sabe-se que não foi possível o deslocamento de forma direta pelo modo aéreo. Portanto, teremos aqui o resultado dos reais desejos de viagens contemplados ou não pelo modal aéreo.

A partir dos valores na matriz O/D, poderão ser realizadas as seguintes análises:

- Áreas de influência dos aeródromos
- Linhas de desejo de deslocamentos
- Potencial de migração para o modo aéreo.

5.5.1. Áreas de influência dos aeródromos

Uma área de influência se refere à área geográfica com potencial de utilização da instalação; sendo, neste caso, o aeródromo. Sua delimitação varia de acordo com os critérios utilizados, que, de modo geral, se relacionam ao uso do solo no entorno da instalação em análise. Assim, dependendo da complexidade da metodologia para delimitação dessa área, ela pode variar de um círculo, quando utilizado o critério de um determinado raio de distância, a um polígono irregular. De acordo com Silva (2006), são exemplos de fatores para delimitação de áreas de influência: “natureza do próprio centro, acessibilidade, barreiras físicas, limitações de tempo e distância de viagem, poder de atração e competição, distância do centro da cidade e principais competidores e concorrência externa”.

Para cada aeroporto, serão analisados os municípios de origem dos passageiros embarcados e os volumes de deslocamentos gerados a partir deles. Analogamente, serão analisados os municípios de destino dos passageiros desembarcados e os volumes destinados a estes. Com isso, será possível delimitar as áreas de influência dos aeródromos através dos deslocamentos observados.

A Figura 11 apresenta uma alternativa de representação dessas informações, por meio de flechas indicando o fluxo dos centroides¹ dos municípios ao aeroporto de embarque, com a espessura proporcional ao volume de viagens. Outra representação possível é uma poligonal com vértices nos centroides dos municípios de origem, no entanto, esta considera como área de influência toda a área delimitada, incluindo municípios que não foram identificados como municípios de origem e que estiverem inseridos nessa área.

¹ Denomina-se centroide o marco do IBGE dos municípios brasileiros, geralmente a parte do município que é considerada o centro de massa, acumulando grande parte dos habitantes e das atividades econômicas locais.

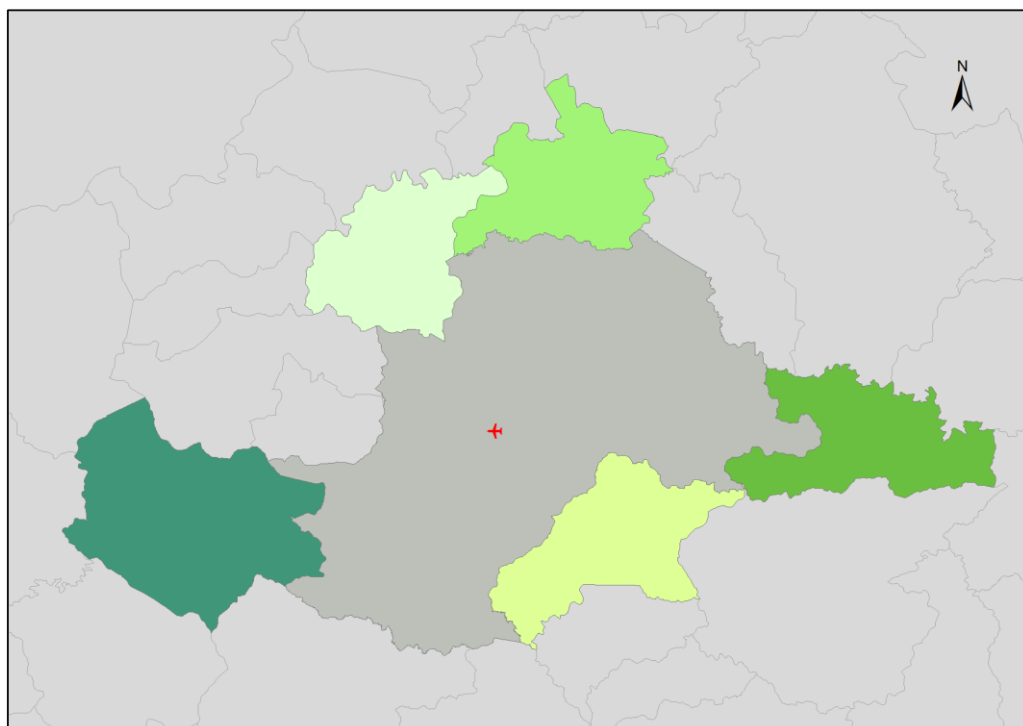
Figura 11 – Exemplo fictício de representação da área de influência do aeródromo (1)



Elaboração: LabTrans/UFSC (2018)

Pode-se, ainda, representar a área de influência através de uma escala de cores para os municípios de acordo com intervalos de volumes de viagens, como mostra a Figura 12.

Figura 12 – Exemplo fictício de representação da área de influência do aeródromo (2)



Elaboração: LabTrans/UFSC (2018)

5.5.2. Linhas de desejo de deslocamentos

Conhecendo os municípios de origem e de destino dos passageiros, será possível, também, analisar as linhas de desejo de deslocamentos entre as UTPs, que serão representadas no mapa do território brasileiro. Essas linhas terão uma escala de cores ou de espessura baseada no volume total de deslocamentos, conforme mostrado na Figura 3 do item 2.4. Além disso, os nós (UTPs de origem/destino) podem ter uma escala de cores de acordo com o volume de viagens geradas/destinadas.

Por meio dessas linhas, é possível verificar se fluxos com volumes de deslocamentos relevantes têm como origem e/ou destino municípios onde estão localizados os aeroportos. Esses aeroportos requerem especial preocupação quanto à manutenção ou investimentos em infraestrutura, a fim de atenderem à demanda com níveis de serviço adequados.

Nos fluxos para os quais não há aeroporto instalado no município de origem e/ou de destino, podem ser identificadas localizações estratégicas para instalação de aeroportos.

5.5.3. Potencial de migração para o modo aéreo

Por fim, ainda com o uso das informações levantadas na matriz O/D, será realizada a análise do potencial de migração de viagens para o modo aéreo. Para os deslocamentos entre UTPs que apresentam aeródromos que operam voos da aviação comercial regular, confrontando-se os volumes dos deslocamentos realizados, independentemente do modo de transporte, com os volumes de deslocamentos aéreos, será possível conhecer a demanda reprimida, ou seja, o número de indivíduos que não utilizam o modo aéreo para o deslocamento. Para estes, será analisado o potencial de migração de outros modos de transporte para o aéreo.

Considerações finais

Com o objetivo de atender às expectativas da SAC/MTPA em relação à elaboração da matriz O/D e suas análises, e tendo em vista a necessidade de adquirir os dados para as atividades previstas na Fase 3 do presente Plano de Trabalho, este relatório foi elaborado para apresentação dos processos a serem realizados. Apresentou-se, também, o conhecimento e material desenvolvido em temas específicos, principalmente Big Data e assuntos relacionados.

O conteúdo apresentado agrega as informações obtidas e definições resultantes de discussões da equipe técnica durante o período de elaboração da minuta do Termo de Referência (Anexo 1) para contratação do fornecimento dos dados necessários para elaboração da matriz O/D. Além disso, mostraram-se as fontes de dados complementares aos dados de Big Data a serem adquiridos que foram levantadas até o momento de entrega do relatório de metodologia.

Conforme exposto no desenvolvimento do tema, contextualizando a matriz O/D no planejamento do transporte aéreo, os produtos descritos visam apoiar a SAC/MTPA no conhecimento do panorama atual do deslocamento da população brasileira em seus reais desejos de viagem para o setor de aviação civil dentro das zonas de interesse, as UTPs. O conteúdo apresentado no presente relatório será detalhado com maior profundidade conforme as atividades da Fase 3 avançarem.

Referências

AGÊNCIA NACIONAL DE TELECOMUNICAÇÕES (ANATEL). **Brasil registra redução de 2,88% no número de acessos em operação na telefonia móvel em 12 meses**. Última atualização: 30 abr. 2018. Disponível em: <<http://www.anatel.gov.br/dados/destaque-1/283-brasil-tem-236-2-milhoes-de-linhas-moveis-em-janeiro-de-2018>>. Acesso em: 03 ago. 2018.

AGÊNCIA NACIONAL DE AVIAÇÃO CIVIL (ANAC). **Lista de aeródromos públicos registrados**. 2016. Disponível em: <<http://www.anac.gov.br/assuntos/dados-e-estatisticas/aeroportos>>. Acesso em: 2016.

ALEXANDER, L. et al. **Origin-destination trips by purpose and time of day inferred from mobile phone data**. Transportation Research Part C: Emerging Technologies, [S.l.], v. 58, p. 240-250, set. 2015. Disponível em: <<http://dx.doi.org/10.1016/j.trc.2015.02.018>>. Acesso em: 27 jul. 2018.

BERA, S.; RAO, K. V. **Estimation of origin-destination matrix from traffic counts: the state of the art**. European Transport\Trasporti Europei, n. 49, p. 3-23, 2011. Disponível em: <https://www.openstarts.units.it/bitstream/10077/6182/1/49D_BeraRao.pdf>. Acesso em: 1 nov. 2018.

BRASIL. DEPARTAMENTO NACIONAL DE INFRAESTRUTURA DE TRANSPORTES (DNIT). **Manual de estudos de tráfego**. Rio de Janeiro, 2006. 384 p. Disponível em: <http://www1.dnit.gov.br/arquivos_internet/ipr/ipr_new/manuais/manual_estudos_trafego.pdf>. Acesso em: 1º nov. 2018.

BRASIL. Ministério do Turismo. **Programa de Regionalização do Turismo**. 27 maio 2015. Disponível em: <<http://www.turismo.gov.br/acesso-a-informacao/63-acoes-e-programas/4882-programa-de-regionalizacao-do-turismo.html>>. Acesso em: 26 nov. 2018.

BRASIL. Ministério dos Transportes, Portos e Aviação Civil. Departamento de Planejamento e Gestão Portuária. **Plano Aeroviário Nacional 2018-2038: objetivos, estratégias e investimentos para desenvolvimento do transporte aéreo brasileiro**. Brasília, nov. 2018. Disponível em: <https://www.transportes.gov.br/images/AVIACAO_CIVIL/PAN/PAN2018_ebook.pdf>. Acesso em: 4 dez. 2018.

BRASIL. Ministério dos Transportes, Portos e Aviação Civil. Departamento de Planejamento e Gestão Portuária. **Projeções de demanda para os aeroportos brasileiros 2017-2037: metodologia e resultados**. Brasília, set. 2017. Disponível em: <http://www.transportes.gov.br/images/AEROPORTOS/Relatorio_Projecoes_2017_2037_Versao_final.pdf>. Acesso em: 26 nov. 2018.

BRUTON, M. J. **Introdução ao planejamento dos transportes**. Rio de Janeiro: Intercedência, 1979.

CAMPOS, V. B. G. **Planejamento de transportes – conceitos e modelos**. Rio de Janeiro: Intercedência, 2013.

CARTER, E. C.; HOMBURGER, W. S. **Introduction to Transportation Engineering**. Virginia, USA: Arlington, 1978.

CHEN, C. L. P.; ZHANG, C. **Data-intensive applications, challenges, techniques and technologies: a survey on Big Data**. Information Sciences, Macau, ago. 2014. Disponível em: <https://www.researchgate.net/publication/262305146_Data-intensive_applications_challenges_techniques_and_technologies_A_survey_on_Big_Data>. Acesso em 23 nov. 2018.

DEMCHENKO, Y. et al. **Addressing big data issues in Scientific Data Infrastructure**. 2013 In: INTERNATIONAL CONFERENCE ON COLLABORATION TECHNOLOGIES AND SYSTEMS (CTS), 2013, San Diego. Anais... San Diego: IEEE, 2013. p. 48-55.

DEMCHENKO, Y.; LAAT, C.; MEMBREY, P. **Defining architecture components of the Big Data Ecosystem**. In: INTERNATIONAL CONFERENCE ON COLLABORATION TECHNOLOGIES AND SYSTEMS (CTS), 2014, Minneapolis. Anais... Minneapolis: IEEE, 2014. p. 104-112.

EMPRESA DE PLANEJAMENTO E LOGÍSTICA S.A (EPL). **Pesquisa de origem/destino do transporte aéreo de passageiros – relatório técnico**. Brasília, 2014. Disponível em: <<https://www.epl.gov.br/pesquisa-de-origem-e-destino-nos-aeroportos>>. Acesso em: 18 set. 2018.

FRIEDRICH, M. et al. **Generating Origin–Destination Matrices from Mobile Phone Trajectories**. Transportation Research Record: Journal of the Transportation Research Board, No. 2196, p. 93-101, 2010. Disponível em: <<https://trrjournalonline.trb.org/doi/10.3141/2196-10>>. Acesso em: 27 jul. 2018.

GARCÍA, P. et al. **Big Data Analytics for a Passenger-Centric Air Traffic Management System: A Case Study of Door-to-Door Intermodal Passenger Journey Inferred from Mobile Phone Data**. In: SESAR INNOVATION DAYS, 6th, 2016, Netherlands. Anais... Netherlands: Technical University of Delft, 2016. p. 1-8.

GERMANI, E.; MORENO NETO, F.; SCATENA J. C.; KAYAL M.; BELDA R.; DOS SANTOS, S. **Planejamentos de transportes**. Departamento de engenharia de transportes escola politécnica – USP. São Paulo. 1973.

HEMMINGS, T.; GOVES, C. **Utilising Mobile Network Data for Transport Modelling: Recommendations Paper**. Reino Unido, 2016.

HENKE, N. et al. **The age of analytics: competing in a data-driven world**. 2016. Disponível em: <<https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/the-age-of-analytics-competing-in-a-data-driven-world>>. Acesso em: 18 set. 2018.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Arranjos populacionais e concentrações urbanas**. [2016]. Disponível em: <http://www.ibge.gov.br/home/geociencias/geografia/geografia_urbana/arranjos_populacionais/default.shtm?c=9>. Acesso em: 26 nov. 2018.

LANEY, D. **3-D Data Management: Controlling Data Volume, Velocity, and Variety**. Gartner Blog Network, 2001. Disponível em: <<http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>>. Acesso em: 23 nov. 2018.

MANYIKA, J. et al. **Big data: The next frontier for innovation, competition, and productivity**. 2011. Disponível em: <<https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/big-data-the-next-frontier-for-innovation>>. Acesso em: 18 set. 2018.

ORTÚZAR, J. D. D.; WILLUMSEN, L. G. **Modelling Transport**. 4. ed. [S.l.]: Wiley, 2011.

PAKMAN, E. T. **Sobre as definições de turismo da OMT: uma contribuição à História do Pensamento Turístico**. In: XI Seminário da Associação Nacional Pesquisa e Pós Graduação em Turismo. Universidade do Estado do Ceará. UECE set. 2014. Fortaleza. Anais... Anptur Associação Nacional de Pesquisa e Pós Graduação em Turismo. Disponível em: <<https://www.anptur.org.br/anais/anais/sumario.php?versao=11>> . Acesso em: 23 nov. 2018.

PEREIRA, L. F. **Um procedimento de apoio a decisão para escolha de sistemas de controle de tráfego considerando a coleta automatizada de dados**. 2005. Dissertação (Mestrado em Ciências e Engenharia de Transportes) – Instituto Militar de Engenharia, Rio de Janeiro, 2005.

SAGIROGLU, S.; SINANC, D. **Big Data: a review**. In: INTERNATIONAL CONFERENCE ON COLLABORATION TECHNOLOGIES AND SYSTEMS (CTS), 2013, San Diego. Anais... San Diego: IEEE, 2013. p. 42-47.

SECRETARIA DE TURISMO DO ESTADO DO ESPÍRITO SANTO. **Pesquisa de demanda e fluxo turístico por meio de sinal de telefonia móvel no Estado do Espírito Santo**. 2017. Disponível em: <<https://observatoriodoturismo.es.gov.br/descritivo-metodologico>>. Acesso em: 23 nov. 2018.

SILVA, L. R. **Metodologia de delimitação da Área de Influência dos Pólos Geradores de Viagens para estudos de Geração de Viagens – Um estudo de caso nos supermercados e hipermercados**. 2006. 184 f. Dissertação (Mestrado em Transportes)-Departamento de Engenharia Civil e Ambiental, Universidade de Brasília, Brasília. Disponível em: <<http://repositorio.unb.br/bitstream/10482/5545/1/2006-Leandro%20Rodrigues%20e%20Silva.pdf>>. Acesso em: 12 dez. 2018.

WANG, M. H. et al. Estimating Dynamic Origin-Destination Data and Travel Demand Using Cell Phone Network Data. **Springer Science+Business Media**, New York, 30 apr. 2013. Disponível em: <https://www.researchgate.net/publication/257797926_Estimating_Dynamic-Origin-Destination_Data_and_Travel_Demand_Using_Cell_Phone_Network_Data>. Acesso: 14 set. 2018.

Lista de abreviaturas e siglas

ANAC	Agência Nacional de Aviação Civil
Anatel	Agência Nacional de Telecomunicações
ANTAQ	Agência Nacional de Transportes Aquaviários
ANTT	Agência Nacional de Transportes Terrestres
CGNA	Centro de Gerenciamento da Navegação Aérea
DECEA	Departamento de Controle do Espaço Aéreo
DNIT	Departamento Nacional de Infraestrutura de Transportes
EPL	Empresa de Planejamento e Logística
FPD	<i>Floating Phone Data</i>
GPS	Global Positioning System
HST	<i>High Speed Train</i>
IBGE	Instituto Brasileiro de Geografia e Estatística
LabTrans	Laboratório de Transportes e Logística
LOS	<i>Level of Service</i>
MGI	<i>McKinsey Global Institute</i>
MTPA	Ministério dos Transportes, Portos e Aviação Civil
MTUR	Ministério do Turismo
O/D	Origem/Destino
OMT	Organização Mundial do Turismo
PAE	Plano Aeroviário Estadual
PAN	Plano Aeroviário Nacional
PII	<i>Personally identifiable information</i>
SAC	Secretaria Nacional de Aviação Civil
SIROS	Sistema de Registro de Operações
SISCEAB	Sistema de Controle do Espaço Aéreo Brasileiro
TR	Termo de Referência
UFSC	Universidade Federal de Santa Catarina
UTP	Unidade Territorial de Planejamento

Lista de figuras

Figura 1 – Processo de planejamento de transportes.....	8
Figura 2 – Padrões de viagem.....	10
Figura 3 – Exemplo fictício de mapa das linhas de desejo entre as zonas de tráfego	13
Figura 4 – 5 Vs do Big Data	16
Figura 5 – Estrutura de Big Data e orquestração do processo	17
Figura 6 – Processo de filtragem do perfil dos usuários	21
Figura 7 – Fluxograma da metodologia	26
Figura 8 – Mapa das UTPs	27
Figura 9 – Aeródromos da área de estudo	29
Figura 10 – Fluxograma das etapas de coleta de dados e elaboração da matriz O/D	31
Figura 11 – Exemplo fictício de representação da área de influência do aeródromo (1) ...	35
Figura 12 – Exemplo fictício de representação da área de influência do aeródromo (2) ...	35

Lista de tabelas

Tabela 1 – Atividades relativas aos diferentes níveis de planejamento	7
Tabela 2 – Modelo de matriz O/D	13

Anexo 1 – Termo de Referência

ANEXO I

TERMO DE REFERÊNCIA

SELEÇÃO PÚBLICA DE FORNECEDORES Nº. 105/2018

PROCESSO Nº. 331271729

1. OBJETIVO

Este Termo de Referência tem por objetivo estabelecer os termos a serem observados no processo de Seleção Pública de Fornecedores para a contratação de serviço de fornecimento de **Base de Dados**, que deverá vir acompanhada de **Relatório de Metodologia**, **Metadados**, **Tabela de Dados Auxiliares**, **Relatório Descritivo do Método de Geração da Base de Dados**, extraídos de registros de usuários de telefonia móvel na rede de telecomunicações e/ou, registros de GPS, ou ainda de outros dados gerados durante o seu deslocamento com informações de espaço e tempo ocorridas pelo modo aéreo e não aéreo na área de abrangência de 780 Unidades Territoriais de Planejamento (UTPs), e **Suporte Técnico** necessários ao desenvolvimento de uma matriz origem/destino (O/D) de passageiros em âmbito nacional para fins de identificação dos reais desejos de viagem contemplados ou não pelo modo aéreo.

2. DISPOSIÇÕES INICIAIS

2.1. PARTICIPAÇÃO EM CONSÓRCIO E/OU JOIN VENTURE

Devido às características do objeto a ser contratado – inerentes à origem, à natureza, ao escopo temporal, à multidisciplinaridade do projeto, bem como à análise dos dados a serem disponibilizados –, a possibilidade de formação de consórcios e/ou join venture, além de reforçar a capacidade técnica e financeira da licitante, proporcionará maior oportunidade à participação de um maior número de empresas interessadas.

2.2. SUBCONTRATAÇÃO

Diante das especificidades do objeto a ser contratado, não será permitida a figura da subcontratação.

2.2. FONTE DE RECURSOS

Os recursos orçamentários necessários à contratação correrão por conta do Projeto nº 046/2018 – Apoio Técnico no Planejamento de Investimentos do Setor de Aviação Civil, administrado pela FAPEU.

3. JUSTIFICATIVA ADMINISTRATIVA

O fornecimento dos dados atenderá às necessidades da Fundação de Amparo à Pesquisa e Extensão Universitária (FAPEU) no que se refere às informações requeridas para o desenvolvimento de uma matriz O/D de passageiros em âmbito nacional para identificação dos reais desejos de viagem contemplados ou não pelo transporte aéreo. Essa necessidade está presente na atividade 3.1 da Fase 3 do Plano de Trabalho do Projeto de Extensão intitulado “Apoio Técnico no Planejamento de

Investimentos do Setor de Aviação Civil”, previsto no Termo de Execução Descentralizada nº 01/2018 celebrado entre a UFSC e a SAC/MTPA, publicado no Diário Oficial da União (DOU) de 27 de abril de 2018, estando sob gestão administrativa e financeira da FAPEU, conforme contrato nº 262/2018 celebrado entre a UFSC e a FAPEU.

4. DESCRIÇÃO DOS SERVIÇOS E AVALIAÇÃO DO CUSTO

Quadro quantitativo dos serviços a serem prestados e valor máximo:

ITEM	DESCRIÇÃO	QUANT.	VALOR R\$
01	Contratação de serviço de fornecimento de Base de Dados, que deverá vir acompanhada de Relatório de Metodologia, Metadados, Tabela de Dados Auxiliares, Relatório Descritivo do Método de Geração da Base de Dados, extraídos de registros de usuários de telefonia móvel na rede de telecomunicações e/ou, registros de GPS, ou ainda de outros dados gerados durante o seu deslocamento com informações de espaço e tempo ocorridas pelo modo aéreo e não aéreo na área de abrangência de 780 Unidades Territoriais de Planejamento (UTPs), e Suporte Técnico necessários ao desenvolvimento de uma matriz origem/destino (O/D) de passageiros em âmbito nacional para fins de identificação dos reais desejos de viagem contemplados ou não pelo modo aéreo.	01 Un.	1.737.960,00

O valor máximo da contratação será de R\$ 1.737.960,00 (um milhão, setecentos e trinta e sete mil, novecentos e sessenta reais).

5. DO OBJETO

O objeto se designa como a contratação de serviço de fornecimento de **Base de Dados**, que deverá vir acompanhada de **Relatório de Metodologia, Metadados, Tabela de Dados Auxiliares, Relatório Descritivo do Método de Geração da Base de Dados**, extraídos de registros de usuários de telefonia móvel na rede de telecomunicações e/ou, de registros de GPS, ou ainda de outros dados gerados durante o seu deslocamento com informações de espaço e tempo ocorridas pelo modo

aéreo ou não aéreo na área de abrangência de 780 UTPs, e **Suporte Técnico** necessários ao desenvolvimento de uma matriz O/D de passageiros em âmbito nacional para fins de identificação dos reais desejos de viagem contemplados ou não pelo modo aéreo.

6. JUSTIFICATIVA TÉCNICA

Tradicionalmente, uma matriz O/D é elaborada utilizando-se informações coletadas por meio de pesquisa presencial com os usuários do(s) modo(s) de transporte considerado(s). Com a evolução tecnológica nas últimas décadas e com o surgimento do conceito de *Big Data*, que diz respeito à possibilidade atual de armazenamento de grandes coleções de dados, tornou-se possível criar soluções e estratégias orientadas por dados (*data driven*) em muitos segmentos da economia com base na utilização de dados de diversas fontes e seu correto tratamento, visando extrair informações, como tendências no comportamento dos usuários.

Nesse contexto, no planejamento de transportes, percebeu-se a existência de novas fontes de dados, passíveis de serem utilizadas para a elaboração de matrizes O/D. **Uma fonte de obtenção de informações relevantes do deslocamento de pessoas que vem sendo utilizada atualmente é a análise de sua pegada digital, por meio de registros dos usuários de telefonia móvel na rede de telecomunicações, de registros de GPS, entre outros dados gerados durante o seu deslocamento com informações de espaço e tempo.** No que se refere à telefonia móvel, segundo a Agência Nacional de Telecomunicações (ANATEL, 2018), no Brasil há 113,54 acessos dos serviços móveis pessoais para cada 100 habitantes. De acordo com García et al. (2016), são vantagens da utilização de dados desses serviços para entender o deslocamento dos usuários em comparação com as pesquisas tradicionais:

1. possibilidade de amostras grandes;
2. coleta dos dados passivamente, eliminando a interação com os entrevistados, o que frequentemente resulta em respostas incorretas ou imprecisas;
3. economia de custo, visto que os dados já são registrados pelas operadoras de telefonia móvel.

Em relação à economia de custo, Hemmings e Goves (2016) citam que pesquisas presenciais têm apresentado elevados custos e consumo de tempo para condução e autorização, e que há um grande potencial de redução desses custos (financeiro e de tempo) com a utilização de dados de telefonia móvel.

Tendo em vista a importância da matriz O/D para o planejamento de transportes e a necessidade de refletir a realidade com a abrangência territorial desejada, a utilização de dados provenientes de *Big Data* é uma alternativa a ser explorada.

Nesse sentido, visando a obtenção de dados provenientes de *Big Data*, necessários ao desenvolvimento de uma matriz O/D de passageiros em âmbito nacional para fins de identificação dos reais desejos de viagem contemplados ou não pelo modo aéreo, se faz mister a contratação de serviço de fornecimento de Base de Dados, que deverá vir acompanhada do Relatório de Metodologia, dos Metadados, das Tabelas de Dados Auxiliares, do Relatório Descritivo do Método de Geração da Base de Dados, extraídos de registros de usuários de telefonia móvel na rede de telecomunicações e/ou, registros de GPS, ou ainda de outros dados gerados durante o seu deslocamento com

informações de espaço e tempo ocorridas pelo modo aéreo e não aéreo na área de abrangência de 780 UTPs, e do Suporte Técnico.

7. DAS ESPECIFICAÇÕES DOS SERVIÇOS

- **RELATÓRIO DE METODOLOGIA**

Deverá ser elaborado, após reunião inicial para fins de alinhamento a ser realizada entre contratada e contratante no prazo máximo de até dez dias após a celebração do contrato, um relatório discorrendo sobre as etapas, as análises e as inferências a serem utilizadas no desenvolvimento da Base de Dados e dos métodos de geração dos dados.

- **BASE DE DADOS**

A Base de Dados deverá ser desenvolvida conforme atributos mínimos dos dados a serem entregues, previstos no Anexo 1 deste Termo de Referência, e contemplará as seguintes especificações:

Origem e natureza dos dados

Os dados devem ser agregados, expressos em número de viagens por fluxo mensal entre UTPs, contendo os atributos mínimos apresentados no Anexo 1 deste Termo de Referência e serem provenientes de:

- 1) registros dos usuários de telefonia móvel; e/ou
- 2) registros de GPS; e/ou
- 3) outras potenciais fontes de dados acerca do deslocamento de indivíduos entre as UTPs.

Formato dos dados

Os dados devem ser fornecidos em Base de Dados contendo os atributos mínimos definidos no Anexo 1 deste Termo de Referência.

Os dados devem ser anonimizados, e os fluxos devem contemplar os deslocamentos entre as 780 UTPs, previstas no Anexo 2 deste Termo de Referência, independentemente do modo de transporte utilizado.

Para os deslocamentos entre UTPs realizados pelo modo aéreo, devem ser informados os municípios de origem e/ou destino reais do passageiro, conforme Anexos 3 e 4 deste Termo de Referência, além dos aeródromos de origem e/ou de destino, previstos no Anexo 5, mesmo se tais municípios situarem-se fora das UTPs de origem ou destino.

Durante o processamento dos dados, quando identificado um novo aeródromo que não estiver presente na lista de aeródromos de referência constante no Anexo 5 deste Termo de Referência, aquele deverá ser incluído na listagem de aeródromos disponibilizada. Para a realização da inclusão do novo aeródromo, deverão ser seguidas duas observações: a criação de um identificador próprio para aquele ponto no município, seguida da inclusão da sua coordenada geográfica.

A classificação do modo de transporte utilizado (**aéreo** ou **não aéreo**) será feita com base no tempo de deslocamento do usuário no par O/D entre as UTPs (Anexo 6). Deverá ser considerada uma velocidade de 200 km/h como valor de referência entre

cada ligação O/D para a divisão entre aéreo e não aéreo. Quando na aferição do último local de um registro na origem e do primeiro local desse registro no destino verificar-se que, em razão do tempo e da distância de deslocamento, a velocidade resultar superior ao valor de referência da ligação O/D, aliado à proximidade a aeródromos em um dos extremos da viagem, o modo de deslocamento desse registro será classificado como aéreo.

Serão caracterizados como origem ou destino os locais (UTPs ou municípios) em que o registro permanecer estacionado por período superior a 6 (seis) horas, conforme Wang (2013).

Escopo temporal

Os dados devem abranger o período de 12 (doze) meses e devem ser coletados de forma ininterrupta a partir do ano de 2017. O período específico de doze meses será estabelecido na reunião inicial para fins de alinhamento a ser realizada entre contratada e contratante no prazo máximo de até dez dias após a celebração do contrato.

Aspectos relacionados à qualidade dos dados

A contratada deve comprovar significância estatística de acordo com metodologia renomada sobre o tema. Da mesma forma, deve apresentar amostra mínima e nível de confiabilidade em relação ao universo e a abrangência geográfica condizente ao escopo do trabalho, sendo necessária a cobertura mínima de 70% dos municípios que compõem as UTPs listadas no Anexo 4. Na comprovação de amostra mínima e nível de confiabilidade, deve ser respeitada a abrangência geográfica em todos os Estados e Regiões.

De acordo com a Anatel (2016), o serviço de telefonia móvel deve ter cobertura de, no mínimo, 80% da área urbana das sedes dos municípios, sendo seu cumprimento acompanhado pela fiscalização da Anatel periodicamente. Levando este parâmetro em consideração, a contratada deve comprovar ter dados referentes à cobertura, no mínimo, do padrão 3G em pelo menos 80% de cada uma das 137 mesorregiões brasileiras como definidas pelo Instituto de Pesquisa Econômica Aplicada (IPEA, 2014).

• METADADOS

Deverá ser disponibilizada juntamente com a Base de Dados, a documentação dos Metadados.

Entende-se por metadado os dados que fornecem informações sobre outros dados. De acordo com a Secretaria de Tecnologia da Informação, órgão federal vinculado ao Ministério do Planejamento, Orçamento e Gestão “[...] metadados são utilizados para registrar atributos sobre um recurso informacional visando facilitar a recuperação, a gestão, a interoperabilidade, dar suporte à identificação digital e dar suporte ao arquivamento e preservação” (BRASIL, 2017). Ademais, através dos Metadados é possível obter um conhecimento preliminar das propriedades e das características de um dado. Uma Base de Dados que possua uma documentação através dos seus Metadados terá seu conteúdo contextualizado, e também será possível, por exemplo, a própria identificação da autenticidade e da qualidade dos dados.

Além dos dados e informações que contextualizem a Base de Dados, deverão ser abordadas na documentação dos Metadados as seguintes informações relativas a aspectos qualitativos e à manipulação da Base de Dados:

- nível de tecnologia de cobertura da rede (2G, 3G, 4G, 5G) e sua partição;
- outras tecnologias consideradas no levantamento, entre elas, Wi-Fi, Bluetooth, GPS;
- especificação da utilização de Registros de Dados de Chamadas (CDRs) somente ou outros (citando os registros);
- especificação da utilização de dados ativos somente ou de dados passivos também;
- especificação das fontes dos dados utilizados (mais de um operador de telefonia móvel/outras fontes) para a composição da base final;
- métodos de validação aplicados;
- se os dados correspondem a dados já coletados/manipulados ou se foram coletados/manipulados para atender ao objeto da contratação.

• **TABELA DE DADOS AUXILIARES**

Deverá ser disponibilizado juntamente com a Base de Dados, a Tabela de Dados Auxiliares.

As tabelas de dados auxiliares contêm o detalhamento de um item participante de uma determinada relação com a tabela principal. Estas tabelas são comumente utilizadas em modelos entidade-relacionamento em banco de dados. O relacionamento ocorre através da relação das chaves primárias das tabelas envolvidas.

• **RELATÓRIO DESCRITIVO DO MÉTODO DE GERAÇÃO DA BASE DE DADOS;**

Deverá ser disponibilizado juntamente com a Base de Dados um arquivo contendo relatório que descreva, de forma sucinta, as informações agregadas à Base de Dados, a fim de validar os passos e as inferências apresentadas no Relatório de Metodologia, além de citar outros eventos relevantes que possam vir a ocorrer.

O Relatório Descritivo do Método de Geração da Base de Dados deverá delinear:

- tipos de eventos considerados válidos para geração da Base de Dados;
- cobertura espacial, período de abrangência e justificativas de quaisquer eventos desprezados;
- se foram assinaladas tendências de deslocamento e de que modo foram levantadas;
- se dados secundários foram utilizados e de que forma;
- detalhamento de qualquer outra manipulação que tenha sido realizada nos dados;
- se há restrição de licença em torno do uso do produto (dados) para uma finalidade específica e qual o potencial de reutilização de dados para outros fins, estudos ou pesquisas futuras.

• **SUPORTE TÉCNICO ESPECIALIZADO**

O Suporte Técnico especializado que deverá ser disponibilizado pelo período de 6 (seis) meses após a disponibilização da Base de Dados tem por escopo o esclarecimento de dúvidas que venham a surgir durante o uso dos dados pela contratante relativas à transferência e à importação dos dados, o suporte na validação dos dados, bem como a garantia da qualidade dos dados, ou seja, caso seja verificada alguma inconsistência, é necessário que a contratada dê todo o suporte necessário para que este problema seja resolvido, até mesmo via atualização dos dados, caso necessário.

8. CRITÉRIOS DE QUALIFICAÇÃO TÉCNICA

Para demonstrar a capacidade técnica, a proponente deverá apresentar, juntamente com a documentação de habilitação, a comprovação de cobertura mínima em 70% dos municípios que compõem as UTPs listadas no Anexo 4 deste Termo de Referência, assim como possuir dados referente à cobertura, no mínimo, do padrão 3G em pelo menos 80% de cada uma das 137 mesorregiões brasileiras definidas pelo IPEA (2014). Na comprovação de amostra mínima e nível de confiabilidade, deve ser respeitada a abrangência geográfica em todos os Estados e Regiões.

A proponente deverá apresentar, ainda, uma amostra preliminar da Base de Dados a ser entregue, com o formato e as especificações detalhadas no Anexo 1, e com as características elencadas para a mesma previstas no Item 7 deste Termo de Referência. Essa amostra preliminar da Base de Dados poderá ter até cinco levantamentos de deslocamentos pares O/D entre as 780 UTPs, e que pelo menos um desses deslocamentos pares O/D tenha característica de utilização do modal aéreo, informando os municípios de origem e/ou destino real do passageiro.

A aceitação da amostra preliminar da Base de Dados estará sujeita à verificação de conformidade com as seguintes características:

- conter as informações e especificações detalhadas conforme Anexo 1.
- possuir aceitação estatística e representatividade da amostra, conforme especificações elencadas para a base de dados previstas no Item 7 deste Termo de Referência.

Em substituição a amostra preliminar da Base de Dados a ser entregue para fins de comprovação de qualificação técnica, a proponente poderá apresentar certificados ou resultados de trabalhos realizados que comprovem a utilização de dados de telefonia móvel para estudos voltados à mobilidade urbana ou interurbana.

NOTA:

CONSÓRCIO E OU JOIN VENTURE – Os atestados de comprovação da capacidade técnica exigidos poderão ser efetuados, no todo ou parte, por qualquer uma das consorciadas, ou por todas, por meio do somatório de seus respectivos atestados.

9. CRITÉRIOS DE ACEITAÇÃO DO OBJETO

Deverá ser efetuada, entre a contratada e a contratante, em até dez dias após a emissão do Pedido de Compra/Serviço, uma reunião de alinhamento para fins de detalhamento da proposta metodológica a ser utilizada no Relatório de Metodologia que deverá ser disponibilizado conforme previsto no Item 10 deste Termo de

Referência. Esta reunião de alinhamento deverá ocorrer em Florianópolis (SC) em local, data e hora a serem definidos pela contratante.

A entrega do Relatório de Metodologia, da Base de Dados, dos Metadados, da Tabela de Dados auxiliares e do Relatório Descritivo do Método de Geração da Base de Dados, *apriori*, deverá ser por meio de um *File Transfer Protocol* (FTP) privado, a ser disponibilizado pela contratada, cujos acessos deverão ser controlados por intermédio de *login* e senha. Todos os acessos ao FTP devem ser disponibilizados juntamente com a entrega do Relatório de Metodologia, devendo permanecer disponíveis durante toda a vigência do contrato. O meio de disponibilização poderá ser alterado caso as partes acordem nova alternativa na reunião inicial de alinhamento.

O sistema de arquivos a ser utilizado para disponibilização dos dados poderá ser, por exemplo, .csv, *backup* de SGBD, *webservices*. Todavia, deverá ser estabelecida extensão final entre as partes durante a reunião de alinhamento a ser realizada em Florianópolis (SC).

A Base de Dados poderá ser particionada em diferentes arquivos, em eventual restrição de tamanho físico, desde que acordado previamente com a contratante. Entretanto, quando disponibilizado de forma particionada, esta deverá estar acompanhada de uma documentação para orientar a junção e a identificação da sua ordem.

A contratada deverá prover Suporte Técnico Especializado a ser prestado por meio de *e-mail*, telefone 0800 ou videoconferência durante o período de 6 (seis) meses após a disponibilização da Base de Dados. O Suporte Técnico Especializado engloba o auxílio na transferência e importação dos dados, suporte na validação dos dados, bem como garantia da qualidade dos dados, ou seja, caso seja verificada alguma inconsistência, é necessário que a contratada dê todo o suporte para que este problema seja resolvido, até mesmo via atualização dos dados, caso necessário. Com isso, a contratada deverá disponibilizar profissionais especializados, de segunda a sexta-feira, durante o horário comercial (das 08:00 às 12:00 e das 14:00 às 18:00), que tenham conhecimento técnico para esclarecimentos de dúvidas que venham a surgir durante o uso dos dados pela contratante. É esperado um tempo máximo para o esclarecimento de 48 horas após o envio da dúvida.

10. PRAZO DE EXECUÇÃO E VIGÊNCIA CONTRATUAL

O prazo de vigência do Contrato **será de 12 (doze) meses**, contados a partir da data de sua assinatura.

O prazo de execução dos serviços **será de 9 (nove) meses**, contados a partir da data da emissão do pedido de Compra/Serviço.

O prazo para a elaboração do **Relatório de Metodologia** será de 1(um) mês, contado da emissão do pedido de Compra/Serviço.

O Relatório de Metodologia deverá ser submetido à contratante para fins de aprovação, que deverá analisá-lo e aprová-lo em até 5 (cinco) dias após o seu recebimento. Caso o Relatório de Metodologia não seja aprovado, a contratada deverá refazê-lo e submetê-lo novamente à aprovação da contratante.

Como o Relatório de Metodologia tem relação direta com a elaboração do Relatório Descritivo do Método de Geração da Base de Dados, a CONTRATADA deverá

11. PAGAMENTO

O pagamento relativo à entrega da Base de Dados, acompanhada do Relatório de Metodologia, dos Metadados, da Tabela de Dados Auxiliares e do Relatório Descritivo do Método de Geração da Base de Dados será efetuado pela CONTRATANTE no prazo máximo de 15 (quinze) dias úteis, contados do adimplemento da obrigação no valor correspondente a 40% (quarenta por cento) do valor do contrato.

O pagamento relativo à prestação do serviço de Suporte Técnico Especializado será efetuado mensalmente pela CONTRATANTE no prazo máximo de 15 (quinze) dias úteis, de acordo com o fornecimento efetivamente executado no valor mensal de 10% (dez por cento) do valor do contrato.

12. OBRIGAÇÕES DA CONTRATANTE

Prestar as informações e os esclarecimentos que venham a ser solicitados pela CONTRATADA.

Efetuar os pagamentos devidos pelos serviços prestados, desde que cumpridas todas as formalidades e exigências do contrato.

Rejeitar os serviços do objeto contratual executados em desacordo com suas orientações e/ou com as especificações constantes deste Termo de Referência e seus anexos.

13. OBRIGAÇÕES DA CONTRATADA

Executar o objeto contratual, em conformidade com as especificações técnicas e com o cronograma de execução previstos nos itens 7 e 10, respectivamente, deste Termo de Referência, e com as instruções emitidas pela CONTRATANTE.

Fornecer, todas as informações sobre a execução do objeto contratual, que a CONTRATANTE julgar necessárias conhecer ou analisar.

Providenciar, em até dez dias após a emissão do pedido de Compra/Serviço, reunião de alinhamento com a CONTRATANTE para fins de detalhamento da proposta metodológica a ser utilizada no Relatório de Metodologia. Esta reunião de alinhamento deverá ocorrer em Florianópolis (SC) em local, data e hora a serem definidos pela contratante.

Disponibilizar FTP privado para a entrega do Relatório de Metodologia, da Base de Dados, dos Metadados, da Tabela de Dados Auxiliares e do Relatório Descritivo do Método de Geração da Base de Dados, cujos acessos deverão ser controlados por intermédio de *login* e senha. Todos os acessos ao FTP devem ser disponibilizados juntamente com a entrega do Relatório de Metodologia, devendo permanecer disponíveis durante toda a vigência do contrato. O meio de disponibilização poderá ser alterado caso as partes acordem nova alternativa na reunião inicial de alinhamento.

Prover Suporte Técnico Especializado por meio de *e-mail*, telefone 0800 ou videoconferência, durante o período de 6 (seis) meses após a disponibilização da Base de Dados, com profissionais especializados, de segunda a sexta-feira, durante o horário comercial (das 08:00 às 12:00 e das 14:00 às 18:00), que tenham conhecimento técnico para esclarecimentos de dúvidas que venham a surgir durante o uso dos dados pela contratante. É esperado um tempo máximo, para o esclarecimento de 48 horas após o envio da dúvida.

14. DAS SANÇÕES ADMINISTRATIVAS

Pelo atraso injustificado, inexecução total ou parcial do objeto deste Termo de Referência, poderão ser aplicadas as seguintes sanções, garantida a prévia defesa:

- a) Juros de 0,5% (zero vírgula cinco por cento) por dia de atraso e por descumprimento das obrigações estabelecidas neste Termo de Referência, até o máximo de 1% (um por cento) sobre o valor dos serviços não executados, recolhida no prazo máximo de 15 (quinze) dias corridos, uma vez comunicada oficialmente;
- b) Multa de até 10% (dez por cento) sobre o valor dos serviços não executados, no caso de inexecução total ou parcial do objeto contratado, recolhida no prazo de 15 (quinze) dias corridos, contado da comunicação oficial, sem embargo de indenização suplementar dos prejuízos porventura causados pela não execução parcial ou total do objeto deste Termo de Referência, nos termos do parágrafo único do art. 416 do Código Civil Brasileiro (Lei n. 10.406/2002).

Referências

AGÊNCIA NACIONAL DE TELECOMUNICAÇÕES (ANATEL). **Brasil registra redução de 2,88% no número de acessos em operação na telefonia móvel em 12 meses.** Última atualização: 30 abr. 2018. Disponível em: <<http://www.anatel.gov.br/dados/destaque-1/283-brasil-tem-236-2-milhoes-de-linhas-moveis-em-janeiro-de-2018>>. Acesso em: 3 ago. 2018.

_____. **Cobertura e Zona de Sombra.** Última atualização: 3 ago. 2016. Disponível em: <<http://www.anatel.gov.br/consumidor/telefonia-celular/direitos/cobertura-e-zona-de-sombra>>. Acesso em: 11 set. 2018.

BRASIL. Ministério do Planejamento, Orçamento e Gestão. Secretaria de Tecnologia da Informação. ePING – Padrões de Interoperabilidade de Governo Eletrônico. 2017. Disponível em: <https://www.governodigital.gov.br/documentos-e-arquivos/e-ping/e-PING_v2017_20161221.pdf>. Acesso em: 26 out. 2018.

GARCÍA, P. et al. Big Data Analytics for a Passenger-Centric Air Traffic Management System: A Case Study of Door-to-Door Intermodal Passenger Journey Inferred from Mobile Phone Data. In: SESAR INNOVATION DAYS, 6th, 2016, Netherlands. **Anais...** Netherlands: Technical University of Delft, 2016. p. 1-8.

HEMMINGS, Tobias; GOVES, Carl. **Utilising Mobile Network Data for Transport Modelling:** Recommendations Paper. Reino Unido, 2016.

INSTITUTO DE PESQUISA ECONÔMICA APLICADA (IPEA). **Bases.** 2014. Disponível em: <<http://www.ipea.gov.br/ipeageo/bases.html>>. Acesso em: 14 set. 2018.

WANG, Ming-Heng; SCHROCK, Steven; BROEK, Nate; MULINAZZI, Thomas. **Estimating Dynamic Origin-Destination Data and Travel Demand Using Cell Phone Network Data.** New York: Springer Science+Business Media, 30 apr.2013. Disponível em: <https://www.researchgate.net/publication/257797926_Estimating_Dynamic-Origin-Destination_Data_and_Travel_Demand_Using_Cell_Phone_Network_Data>. Acesso: 14 set .2018.

