

Qualidade percebida de produtos e serviços turísticos em eventos: Técnicas e ferramentas para análise de conteúdo do Twitter

AUGUSTO J. W. A. DAS NEVES * [augustojw@gmail.com]

PATRÍCIA ZENI MARCHIORI ** [pzeni@ufpr.br]

Resumo | Este artigo tem como objetivo discutir o uso de ferramentas para análise de conteúdo de postagens do Twitter relativo à percepção, por usuários, da qualidade de produtos e serviços turísticos em eventos. Realizou-se um levantamento de literatura pertinente, com a finalidade de caracterizar o produto turístico, mídias sociais e a análise de conteúdo. Elaborou-se uma estrutura de categorias e itens para a coleta de postagens do Twitter que versassem sobre atitudes demonstradas por usuários relacionadas aos produtos e serviços turísticos disponibilizados para consumo durante os Jogos Olímpicos de Londres 2012. Propôs-se uma estrutura para analisar o conteúdo das informações disseminadas por usuários no Twitter, na qual se explorou a técnica *Evaluative Assertion Analysis* e ferramentas automáticas disponíveis *online*. Os 878 tweets derivados da amostra foram inseridos nas ferramentas selecionadas, tendo os resultados obtidos sido comparados com aqueles obtidos pela *Evaluative Assertion Analysis*. Concluiu-se que ambas as alternativas são viáveis para analisar o conteúdo disponível no Twitter relativo à qualidade de produtos e serviços turísticos em eventos. Ambas as alternativas apresentam convergência nos custos para a sua utilização e divergências nos custos da sua operacionalização, ou seja, no tempo despendido para análise, nos processos analíticos adotados e nos resultados obtidos..

Palavras-chave | Análise de conteúdo, Eventos, Produtos e serviços turísticos, Twitter.

Abstract | This article discusses the use of tools for content analysis in Twitter posts related to the quality of touristic products and services in events. A survey of the relevant literature was carried out in order to characterize the touristic product, social media and content analysis. Data posted on Twitter that expressed attitudes demonstrated by users related to tourism products and services available for consumption during the London Olympics 2012 were collected. It was proposed a framework for analysing the content of the information disseminated by users on Twitter, in which the Evaluative Assertion Analysis technique was explored, as well as tools available online. The 878 tweets sampled were inserted and analysed in the selected tools, having the obtained results been compared with those achieved by the Evaluative Assertion Analysis. It was concluded that both alternatives are feasible to determine the attitudes of Twitter users that are related to consumption of touristic products and services during events. Both alternatives converge in the costs for their use and diverge in the efforts demanded for their operation, especially concerning the time taken for analysis, the adopted analytical procedures and the comparison of the distinct obtained results.

Keywords | Content analysis, Events, Tourism products and services, Twitter.

* **Mestre em Ciência, Gestão e Tecnologia da Informação** pela Universidade Federal do Paraná.

** **Doutora em Ciências da Comunicação** (Ciência da Informação) pela Universidade de São Paulo. **Professora Titular** na área de Teoria e Gestão da Informação na Universidade Federal do Paraná, **Líder do Grupo de Pesquisa** em Metodologia para Gestão da Informação (UFPR), **Investigadora do Grupo de Pesquisa** CEDUS – Design de Sistemas Virtuais Centrado no Usuário (USP).

1. Introdução

Em 2008, durante o processo eleitoral para a definição do Presidente da República dos Estados Unidos da América, foi possível constatar a crescente importância das mídias sociais nas decisões do cotidiano. Durante este evento, o ainda Presidente Barack Obama (reeleito em 2012) se comunicava com os seus apoiadores através do Facebook e do Twitter. Mais recentemente, o início de 2011 trouxe a revolução egípcia na qual o Twitter surgiu como um dos principais instrumentos usados pelos populares para se comunicarem entre si e com o restante do mundo (Bhuiyan, 2011).

Organizações vêm utilizando as mídias sociais não só para divulgarem e promoverem os seus produtos e serviços, como também para dialogarem diretamente com os consumidores, esclarecendo dúvidas, desburocratizando a resolução de problemas e ajudando a criar vínculos de proximidade e de confiança. Na atividade turística essas transformações são constatadas, por exemplo, por meio de relatos que manifestam a opinião dos seus usuários em relação aos produtos e serviços turísticos que consomem ou pretendem consumir.

Neste cenário, os intervenientes da atividade turística, sejam eles integrantes da iniciativa pública ou privada, tendem a assumir a responsabilidade de monitorar e analisar o conteúdo explicitado pelos consumidores nas mídias sociais, o que pode exigir a contratação de mão-de-obra.

Baseando-se nestas premissas, a presente pesquisa teve como objetivo principal o de discutir o uso de técnicas e de ferramentas para análise de conteúdo de postagens do Twitter com foco na expressão, por usuários, da qualidade percebida de produtos e serviços turísticos em eventos turísticos. A estrutura metodológica foi testada em um fluxo de postagens (*tweets*) coletado durante os Jogos Olímpicos de Londres/2012 e, espera-se que os resultados possam vir a contribuir para o uso crítico de tais ferramentas e/ou técnicas de análise de conteúdo para o uso de informação derivada diretamente dos usuários/

consumidores, o que impulsionar ajustes e ações corretivas dos promotores e gestores de eventos.

2. Eventos turísticos: A qualidade percebida e o Twitter como espaço de expressão do consumidor

O mercado turístico é entendido como sendo o encontro e a interação entre a oferta e a demanda que está interessada e motivada em consumir os seus produtos e serviços (Ministério do Turismo, 2007). O produto turístico compreende “o conjunto de atrativos, equipamentos e serviços turísticos acrescidos de facilidades, localizados em um ou mais municípios, ofertado de forma organizada por um determinado preço” (Ministério do Turismo, 2007, p. 17).

O fato dos produtos turísticos poderem ser semelhantes entre si deriva em uma crescente segmentação de mercado, não só como forma de diferenciação em relação aos produtos semelhantes, como também para melhor atender às necessidades da demanda (Ministério do Turismo, 2007). Quando se consideram as variáveis da demanda, os critérios de idade, tipos de acompanhamento, nível de renda do turista, âmbito geográfico, duração da viagem, motivação da viagem, local da prática do turismo, e meio de transporte são utilizados como forma de segmentação (Beni, 2004). Destes, a motivação da viagem é considerada a causa primordial que instiga o consumidor a viajar e a consumir produtos turísticos (Ministério do Turismo, 2007). Entre as distintas motivações, os eventos merecem destaque, pois geram um fluxo adicional de consumidores com necessidades e características específicas para o local onde estes ocorrem (Beni, 2004).

Considerando-se que o produto turístico depende principalmente da forma como o ser humano o considera como algo único e insubstituível, torna-se importante entender como este tem se apropriado das tecnologias e ferramentas criadas e disponibili-

zadas para favorecer distintas formas de expressão e comunicação, em especial as mídias sociais.

Entre as mídias sociais existentes, as de *microblogging* adquirem especial interesse por permitirem a produção e disponibilização de conteúdo de forma rápida, clara e objetiva para todos os usuários. Um *microblogging* é uma ferramenta que permite aos usuários enviar breves atualizações de conteúdo por meio da *web* ou de dispositivos móveis e é considerado uma vertente das mídias sociais voltada para a produção e disseminação de conteúdos que apresentam um ciclo de vida curto e repetitivo (Heverin & Zach, 2010; Vieweg, Hughes, Starbird, & Palen, 2010). Jansen, Zhang, Sobel, e Chowdury (2009) afirmam que os conteúdos veiculados em serviços de *microblogging* podem expressar atitudes ou sentimentos/emoções, o que tende a influenciar o ponto de vista dos usuários em relação a produtos, serviços ou opiniões. Das ferramentas de *microblogging* existentes no mercado, o Twitter se sobressai por ser o maior em número de inscritos (Socialbakers, 2012) e, segundo Hay (2012), tem sido objeto de estudos recentes relacionados à atividade turística tais como na análise da sua aplicação como instrumento de marketing por hotéis ou por organizações que gerem a imagem e a promoção dos destinos; e na identificação de como o conteúdo produzido e compartilhado por usuários interfere no seu desenvolvimento (Barbagallo, Bruni, Francalanci, & Giacomazzi, 2012).

Fotis, Buhalis e Rossides (2011) afirmam que o conteúdo gerado e compartilhado pelos consumidores em mídias sociais apresenta maior confiabilidade do que aquele publicado por órgãos oficiais de turismo dos destinos visitados, pelo simples fato de representar a opinião do consumidor sem vínculos comerciais de publicidade e propaganda.

Portanto, um dos desafios para o efetivo uso de tais conteúdos repousa não apenas no monitoramento das plataformas e ferramentas que viabilizam o fluxo destes entre múltiplos usuários, mas especialmente a análise do teor das mensagens veiculadas.

3. Procedimentos metodológicos

Para apresentar uma estrutura de análise de conteúdo voltada ao aproveitamento de postagens do Twitter, elaborou-se um *checklist* com base nas obras de:

- Freund (2005), que explicita o funcionamento e as características do setor de alimentos e bebidas;
- Lage e Milone (2000), Cooper, Fletchner, Fyall, Gilbert e Wanhill (2001), e Beni (2004), que abordam os fundamentos e as bases para o desenvolvimento estruturado e sustentável da atividade turística;
- Ebner et al. (2010), que descrevem o uso do Twitter em eventos;
- Pak e Paroubek (2010) que abordam a análise de sentimentos no Twitter;
- Becker, Naaman e Gravano (2011), que descrevem a seleção de conteúdo de qualidade para eventos no Twitter;
- Shimada, Inoue, Maeda e Endo (2011), que abordam a análise de informações turísticas no Twitter; e
- Heverin e Zach (2012), que estudam o Twitter como ferramenta utilizada pela Polícia para compartilhamento de informações com os usuários.

As categorias de análise adotadas para a identificação da opinião dos consumidores em relação a serviços turísticos foram: alimentação, hospedagem, transporte e segurança. Para Petrocchi (2009), alimentação, hospedagem e transporte são determinantes quando se prevê a realização de eventos num determinado local para que, quando se registre o aumento de visitantes que se deslocam para participar nesse acontecimento, tanto eles como os usuários saiam satisfeitos com a sua experiência de consumo. Para a Confederação Nacional do Comércio de Bens, Serviços e Turismo (2011), o item segurança é relevante para o consumidor que participa de um evento e constrói uma boa imagem do destino visitado, tendendo a promovê-lo positivamente.

Quadro 1 | Número de *tweets* e categorias de análise: Jogos Olímpicos de Londres 2012.

Categoria de análise	Número de itens analisados	Número de <i>tweets</i> recuperados
Alimentação	14	4.576
Hospedagem	8	375
Transporte	12	3.486
Segurança	6	1.274

Fonte: Elaboração própria.

O fluxo de postagens/*tweets* dos Jogos Olímpicos de Londres de 2012 foi tomado como objeto de estudo considerando-se que eventos de grande porte surgem como uma oportunidade de promover o local onde são realizados como destino turístico cosmopolita e multicultural, envolvendo tanto comunidade local como os visitantes de distintas nacionalidades com foco na hospitalidade e qualidade (Rubio, 2007).

Como complementos às categorias de análise iniciais, e de forma a adaptar a estrutura de coleta para qualquer outro evento, foram agregados itens de análise – para cada uma das categorias acima especificadas – com base nas informações presentes no sítio oficial do evento. A categoria ‘alimentação’, por exemplo, foi desmembrada em 14 itens, tais como: *Dinner, Bar, Pub, Brasserie, Grill, Café, Lunch, Fish & Chips, Food, Drink* (entre outros).

Para a obtenção e armazenamento dos *tweets* utilizou-se o aplicativo da *Microsoft™ Analytics for Twitter®* durante o período oficial da realização do evento. Esta escolha foi respaldada pela gratuidade do aplicativo, pela interface agradável, por permitir a tabulação e acesso dos resultados em gráficos e números, e por recuperar dados e informações de forma detalhada (Microsoft, 2013).

A obtenção de dados nesta ferramenta ocorreu primeiramente pela utilização da nomenclatura #nome do evento (e.g. #beijing2008). Além do uso deste tópico, também se adotou o uso do booleano *and* seguido do item que se pretendia observar dentro de cada uma das categorias definidas para a análise. Por exemplo, para recuperar informação sobre o metrô de Pequim durante as Olimpíadas de 2008, a pesquisa no aplicativo realizar-se-ia da se-

guinte forma: #beijing2008 and underground. Para a pesquisa em pauta foram recuperados 9.711 *tweets* recolhidos diariamente entre os dias 27 de julho e 12 de agosto, espaço de tempo que representa o período oficial do evento, após as 20h (horário de Brasília), correspondendo ao horário oficial de encerramento das atividades diárias dos Jogos (24h horário de Londres) (Quadro 1).

Para o armazenamento dos *tweets* utilizaram-se tabelas em planilha de cálculos do *Microsoft™ Excel®* e, para a seleção dos *tweets* válidos, aplicaram-se os seguintes critérios: *tweet* no idioma inglês; presença obrigatória do tópico #london2012; possibilidade de se utilizar o booleano *and* para a recuperação das categorias/itens; e, postagem ocorrida durante o período oficial do evento. Excluíram-se *tweets* em formato replicado (o *retweet*) devido ao fato destes representarem apenas o compartilhamento de conteúdo publicado por outro usuário, não acrescentando conteúdo novo em relação ao original (Twitter, 2012). Esta primeira seleção resultou em 4.440 *tweets* válidos para análise aos quais se aplicou a fórmula para cálculo de amostras estratificadas para populações finitas (inferiores a 100.000 itens)¹ (Gil, 1999) (Quadro 2).

Em seguida realizou-se uma análise de conteúdo destes *tweets* utilizando-se a técnica *Evaluative Assertion Analysis* (EAA) de Bardin (2011), que tem como objetivo identificar o fundamento das atitudes, independentemente da forma que assumem quando verbalizadas.

Bardin (2011) esclarece que nem todo o material pronto para ser analisado é considerado devendo-se escolher apenas os itens que exprimem uma avaliação em relação a um determinado objeto e especificar uma nomenclatura para a avaliação dos componentes dos enunciados, a saber: (i) os objetos de atitude (AO/*Attitude Objects*): representam os itens avaliados; (ii) os termos avaliativos com significação comum (cm/*evaluative common-meaning ter-*

¹ Estabeleceu-se um nível de confiança de 95,5% (dois desvios padrão; $\sigma = 2$) e um erro máximo permitido de 5% ($e = 5$).

Quadro 2 | Cálculo de amostra para a análise dos *tweets*: Jogos Olímpicos de Londres 2012.

Categoria	Valores	Tamanho da amostra
Alimentação	$\sigma = 2$ p = 56,31% (=2.500 tweets válidos na categoria) q = 43,69% (=100%-p) N = 4.440 (total de tweets válidos) e = 5%	n = 362
Hospedagem	$\sigma = 2$ p = 2,43% (=108 tweets válidos na categoria) q = 97,57% (=100%-p) N = 4.440 (total de tweets válidos) e = 5%	n = 38
Transporte	$\sigma = 2$ p = 27,97% (=1.242 tweets válidos na categoria) q = 72,03% (=100%-p) N = 4.440 (total de tweets válidos) e = 5%	n = 301
Segurança	$\sigma = 2$ p = 13,29% (=590 tweets válidos na categoria) q = 86,71% (=100%-p) N = 4.440 (total de tweets válidos) e = 5%	n = 177

Fonte: Elaboração própria com base em Gil (1999).

ms): responsáveis por qualificar os objetos de atitude e que assumem a forma de adjetivos, substantivos, advérbios formados a partir de adjetivos (e.g. lealmente) e alguns verbos (e.g. mentir, respeitar); e, (iii) os conectores verbais (c): ligam os objetos de atitude e os termos de qualificação.

Seguindo-se as três etapas preconizadas pela técnica, a análise considerou para o conteúdo de cada um dos *tweets* da amostra a existência de objetos de atitude (AO), os termos avaliativos com significação comum (cm), e os conectores verbais (c). Normalizaram-se os enunciados obtendo-se afirmativas descritivas quanto a intenção dos *tweets*, que permitiram a análise apoiada na padronização do conteúdo dos *tweets* em frases sob uma forma sintática mais elementar (ator-ação-complemento). Este formato facilitou a atribuição de um direcionamento positivo, negativo ou neutro aos *tweets*, seus signifi-

cados² e o contexto onde as palavras foram inseridas.

Em um segundo momento de análise - dedicado a explorar ferramentas disponíveis na Internet que realizam análises automáticas de postagens do *Twitter* – recorreu-se ao sítio eletrônico do projeto *Sentiment140*, da Universidade de Stanford/EUA (*Sentiment140*, 2012). Das 130 ferramentas de análise de *tweets* gratuitas e pagas, listadas até 21 de março de 2013, seis preencheram integralmente os critérios definidos³ de seleção, a saber: *Repustate* (versão demo); *Tweenator*; *Sentimentor*; *Sentiment Analyzer*; *Sentiment Analysis with Python NLTK Text Classification*; *SentiStrength* (versão demo). Além destas ferramentas, também se utilizou o aplicativo Microsoft™ *Analytics for Twitter*®, anteriormente utilizado como suporte à realização da EAA.

Os sítios de tais ferramentas foram explorados em busca de informação sobre os algoritmos utilizados, os dicionários de tons⁴ e as respetivas escalas para a medição de tais tons. As ferramentas não apresentaram uniformidade no uso e na transparência de tais funcionalidades. Porém, verificou-se – naquelas que disponibilizavam alguns dados – o uso de escalas numéricas com variantes de (+1) a (-1) (ou de +1 e +5 e -5 e -1) para indicar força

² Utilizou-se o dicionário online de língua inglesa da Universidade de Princeton (<http://wordnetweb.princeton.edu/perl/webwn>) para a tradução.

³ Permite utilização sem custos; é intuitivo/autoexplicativo; faz análise de atitudes; faz análise de *tweets* em inglês; analisa o *tweet* pós-coleta de forma integral.

⁴ Um "tom" (*tone*) é a emoção que um determinado adjetivo está transmitindo (Bastone, 2011).

Quadro 3 | Exemplo de direcionamento de enunciados relacionados a conteúdos de tweets: Jogos Olímpicos de Londres 2012.

Tweet	Identificação e classificação de conteúdos	Direcionamento dos conteúdos	Normalização do enunciado	Direcionamento do enunciado
<i>Getting food and drink in order for #London2012 opening ceremony</i>	<i>Getting</i> – conector verbal (c)	Neutro	O usuário deixa alimentos e bebidas em ordem para a cerimônia de abertura de Londres 2012	Positivo
	<i>Food and drink in order for</i> – termos avaliativos de significação comum (cm)	Positivo		
	<i>#London 2012 opening ceremony</i> – objeto de atitude (AO)	Neutro		
<i>Wine, Mexican food and #London2012</i>	<i>Wine, Mexican food</i> – termos avaliativos de significação comum (cm)	Neutro	O usuário escreve vinho, comida mexicana e Londres 2012	Neutro
	<i>#London 2012</i> – objeto de atitude (AO)	Neutro		

Fonte: Elaboração própria com base em Bardin (2011).

positiva, neutra ou negativa (Repustate, 2013; Sentistrength, 2013); o uso de visualização (barômetro) (Sentimentalyzer, 2013); ou ainda cores: azul (amarelo ou cinza escuro) para o neutro, vermelho para o negativo, e verde para o positivo. (Sentimentor, 2013, Tweenator, 2013, Sentistrength, 2013). Quanto aos algoritmos, o *Sentistrength* (2013), o *Sentiment Analysis with Python NLTK Text Classification* (Text-Processing, 2013) e o Tweenator (2013) revelam em maior ou menor detalhe a estrutura de sua implementação e operacionalização.

Visando a comparação dos resultados obtidos pela EAA procedeu-se à realização da análise de conteúdo dos *tweets* em cada uma das ferramentas selecionadas e os resultados tabulados em planilha do *Microsoft™ Excel®*.

4. Descrição e análise dos resultados

4.1. Aplicação da *Evaluative Assertion Analysis* nos *tweets* da amostra

A realização da etapa de exploração do material iniciou-se com a identificação, classificação

e direcionamento dos objetos de atitude (AO), termos avaliativos com significação comum (cm) e conectores verbais (c). Seguiu-se a normalização dos enunciados de forma a obter afirmativas que descrevam a intenção dos *tweets* e para que o conteúdo destes se tornasse passível de análise. Após a normalização dos enunciados foi atribuído um direcionamento (positivo, negativo ou neutro) a cada um deles (Quadro 3).

Ao se analisar sintaticamente o *tweet* *Getting food and drink in order for #London2012 opening ceremony* verifica-se que o usuário está preparando comida e bebida para poder acompanhar a cerimônia de abertura das Olimpíadas de Londres 2012. Desta forma, *#London 2012 opening ceremony* (cerimônia de abertura de Londres 2012) surge como o objeto de atitude, pois é por causa dele que o usuário toma uma ação. Por se tratar de um acontecimento que nem todos os usuários teriam uma atitude positiva, optou-se por atribuir um direcionamento neutro.

O conector verbal *getting* indica que o usuário está preparando/conseguindo algo. Como o verbo *to get* (em português, obter) depende do contexto onde se insere, para ter um direcionamento positivo ou negativo, também se optou por lhe atribuir neutralidade. O conector verbal (*food and drink in order for*) indica que a comida e a bebida estão ficando prontas/em or-

dem para serem consumidas. Como é antecedido pelo verbo *to get*, o conector verbal adquire um direcionamento positivo. Entretanto, nem sempre é possível distinguir todos estes conteúdos dentro de um *tweet*.

No exemplo *Wine, Mexican food and #London2012* não se verifica a existência de um conector verbal de forma explícita. Ainda assim, por meio do significado de cada uma das palavras é possível verificar que o usuário tem vinho e comida mexicana para ver as olimpíadas. A implicação de que vinho, comida mexicana e as olimpíadas apresentam um significado positivo ou negativo depende do conector verbal e do contexto no qual eles são servidos, degustados e observados. Por este motivo optou-se por atribuir neutralidade tanto para o objeto de atitude quanto para o conector verbal.

A normalização dos enunciados permitiu transformar os *tweets* em frases sob a sintaxe mais elementar (ator-ação-complemento) o que permitiu atribuir um direcionamento para o *tweet* como um todo e suprimir a falta de um conector verbal em alguns deles, facilitando atribuição de um significado neutro, positivo ou negativo.

Verifica-se que a atribuição do direcionamento aos enunciados elaborados com base nos *tweets* é influenciada pelo direcionamento atribuído ao conector verbal (c), ao termo avaliativo de significação comum (cm) e ao objeto de atitude (AO). Mais especificamente, são os termos avaliativos de significação comum que apresentam um papel relevante nesta etapa. Caso o *tweet* apresente um 'cm' que é complementado pela ação do verbo, a sua positividade, negatividade ou neutralidade é identificável. Caso não haja a presença

de um conector verbal que indique a ação do usuário em relação à atitude explicitada por si no *tweet*, a atribuição de um direcionamento para este é dificultada e a sua realização pode não relatar com exatidão o conteúdo exteriorizado.

Identifica-se também que o direcionamento do enunciado permite determinar a atitude demonstrada pelo usuário no *Twitter* em relação às categorias e itens propostos para análise (Quadro 4).

Os itens *food*, *hotel*, *bus* e *security* foram classificados como tendo sido positivos e, neste contexto, é possível afirmar que a comida, os hotéis, os ônibus e a segurança foram considerados como satisfatórios. Embora essa atitude seja a predominante em relação a esses itens, apenas os identificados como *hotel* e *security* se pode afirmar serem seguramente satisfatórios por apresentarem mais de metade das atitudes avaliadas como positivas (55,2% de *tweets* positivos para o item *hotel* e 63,8% de *tweets* positivos para o item *security*). Os demais itens foram avaliados satisfatoriamente, mas todos com porcentagens de atitudes positivas inferiores a 50% (*food*, 49%; *bus*, 41%). Destaca-se também o fato do item *bus* ter porcentagem superior de atitudes neutras em relação às atitudes negativas, podendo indicar que os usuários tenham exteriorizado mais opiniões informativas do que negativas em relação aos ônibus disponibilizados durante a realização do evento.

O fato da aplicação da técnica depender diretamente do pesquisador confere alguma liberdade e capacidade adaptativa na sua execução e alinhamento aos objetivos propostos. Entretanto, por apresentar diversas etapas, a aplicação desta

Quadro 4 | Atitudes de usuários representadas por meio de *tweets*: Jogos Olímpicos de Londres 2012.

Categoria	Item analisado	Atitudes representadas pelos <i>tweets</i>							
		Positiva		Negativa		Neutra		Total	
		Núm.	%	Núm.	%	Núm.	%	Núm.	%
Alimentação	food	178	49,2	131	36,2	53	14,6	362	41,2
Hospedagem	hotel	21	55,2	5	13,2	12	31,6	38	4,3
Transporte	bus	124	41,2	81	26,9	96	31,9	301	34,3
Segurança	security	113	63,8	40	22,6	24	13,6	177	20,2
	Total	436	49,6	257	29,3	185	21,1	878	100

Fonte: Elaboração própria.

técnica demandou aproximadamente dois minutos e 24 segundo para a análise de cada *tweet*, o que não é indicado para situações onde os prazos para análise são curtos.

A morosidade pode ser entendida como um ponto positivo, pois os procedimentos gerais da metodologia de análise de conteúdo e a técnica de EAA respeitam uma estrutura elaborada por etapas interdependentes e indissociáveis, o que exige atenção e critério por parte do pesquisador e lhe confere coerência e consistência. As ferramentas automáticas, todavia, apresentaram processos e aspetos de análise distintos.

4.2. Uso de ferramentas automáticas para análise dos tweets da amostra

Ao realizar análise dos *tweets* que compuseram as amostras previamente definidas na ferramenta *Repustate*, os itens *food*, *hotel*, *bus* e *security* foram classificados como tendo sido positivos, tornando-se possível afirmar que a comida, os hotéis, os ônibus e a segurança foram considerados como satisfatórios. Cada *tweet* demorou, em média, dois segundos para ser analisado.

Na ferramenta *Tweenator*, os itens *food*, *hotel* e *bus* foram classificados como tendo sido negativos e o item *security* como positivo/satisfatório. Apesar da ferramenta considerar a avaliação neutra de *tweets*, na análise das amostras estas foram classificadas apenas como positivas ou negativas. Cada *tweet* demorou, em média, dois segundos para ser analisado.

Na ferramenta *Sentimentor*, os itens *food*, *hotel*, *bus* e *security* foram classificados de forma mais negativa do que positiva, o que pode indicar insatisfação em relação ao consumo de produtos e serviços relacionados a estes itens. Esta ferramenta demorou, em média, 12 segundos para analisar cada *tweet*.

Na ferramenta *Sentiment Analyzer*, os itens *food*, *hotel*, *bus* e *security* foram classificados como positivos/satisfatórios, com um tempo médio de análise por *tweet* de dois segundos.

No *Sentiment Analysis with Python NLK Text Classification* obtiveram-se resultados semelhantes àqueles resultantes do *Sentimentor*. Apesar da maioria das avaliações das amostras terem sido consideradas como neutras, verificou-se que os itens *food*, *hotel* e *bus* foram classificados com peso negativo, enquanto o item *security* foi classificado de forma positiva/satisfatória. Cada *tweet* demorou, em média, dois segundos para ser analisado.

Com relação à ferramenta *SentiStrength* – cuja média de análise por *tweet* foi de dois segundos – constatou-se que, apesar do número considerável de avaliações neutras, os itens *food*, *hotel*, *bus* e *security* foram classificados como positivos/satisfatórios.

A análise de conteúdo no *Microsoft™ Analytics for Twitter®* classificou os itens *food*, *hotel*, *bus* e *security* como neutros. Estes resultados podem indicar que a maioria dos *tweets* incluídos nas amostras selecionadas para análise não possui qualquer um dos termos que compõem o dicionário de tons com o qual esta ferramenta opera. Com relação ao tempo de análise de cada *tweet*, os resultados variaram em função da quantidade de *tweets* recuperados ainda durante a etapa de coleta de dados oscilando entre cinco e quinze segundos por *tweet*.

Considerando-se possíveis limitações das ferramentas, no caso do *Microsoft™ Analytics for Twitter®*, esta pode derivar do uso de uma escala reduzida das amostras para detectar os tons das atitudes demonstradas pelos usuários. Para o *Tweenator*, o *Sentimentor*, *Sentiment Analysis with Python NLK Text Classification* e o *SentiStrength* as limitações podem resultar do uso de algoritmos que sejam programados para identificar apenas alguns sentimentos; do uso de conjuntos lexicais que não contenham conteúdos (ex.: imagens, termos, informações) presentes em *tweets*, prejudicando a sua análise; ou, ainda, por algumas das ferramentas serem versões demonstrativas, não disponibilizando aos seus usuários todas as potencialidades da versão paga. As ferramentas *Repustate*, *Sentiment Analyzer* e *SentiStrength* apresentaram resultados mais semelhantes entre si, com uma tendência em

classificar os *tweets* como positivos ou neutros, o que pode sugerir que os algoritmos utilizados sejam semelhantes. Uma exceção é feita ao item *food* na ferramenta *Sentiment Analyzer* na qual os *tweets* foram majoritariamente classificados em positivos ou negativos.

5. Considerações finais

É possível perceber que nenhuma das ferramentas automáticas selecionadas apresenta resultados exatamente iguais aos propostos pelo pesquisador por meio da EAA. Entretanto, constata-se que a ferramenta que apresenta resultados mais próximos daqueles propostos pelo pesquisador por meio da EAA é a *Sentiment Analyzer*. Contudo, o comportamento analítico em relação aos itens de análise apenas se revelaram semelhantes em três dos quatro itens analisados (*food*, *hotel* e *bus*). Porém, não é possível afirmar se as estruturas analíticas de ambas são semelhantes, pois não se teve acesso ao algoritmo e aos conjuntos lexicais que permitem o funcionamento desta ferramenta.

O uso do *Sentiment Analyzer* pode, potencialmente, fornecer uma ideia de atitudes reveladas pelos usuários nos *tweets* mais próxima daquela obtida pela análise manual dos conteúdos. Porém, não existe a certeza de que esses resultados representam a veracidade e a intencionalidade dos usuários, o que pode prejudicar a sua interpretação e ações que sejam tomadas com base nos resultados obtidos.

Em relação aos custos de utilização, tanto a estrutura de EAA proposta como as ferramentas analisadas não apresentam custos de utilização, embora a aplicação da EAA dependa de mão de obra especializada devido às etapas de identificação e classificação de conteúdos e de normalização de enunciados.

Ao se considerar o tempo de análise, as ferramentas utilizadas apresentam maior eficiência por

permitirem a realização da análise de conteúdo em até no máximo 15 segundos, enquanto, com a EAA, o pesquisador demora um pouco mais de dois minutos para analisar um *tweet*.

Ainda assim, os procedimentos adotados para realizar a análise de conteúdo por meio da EAA são conhecidos e adaptáveis aos objetivos determinados pelo pesquisador, principalmente por dependerem das inferências que este irá efetuar durante a realização da sua pesquisa.

Apesar de apresentarem mais divergências do que convergências, tanto as ferramentas como a estrutura de análise de conteúdo proposta permitem identificar e determinar, por meio da análise de *tweets*, a qualidade percebida pelos usuários em relação aos produtos e serviços turísticos consumidos durante a realização de um evento.

Considera-se que a identificação completa dos algoritmos, e a explicitação da origem das bases de dados usadas para identificar atitudes em conteúdos textuais possam contribuir para um melhor entendimento do processo analítico utilizado pelas ferramentas disponíveis na Internet, diminuindo as divergências encontradas entre a EAA e as versões das ferramentas selecionadas.

Referências bibliográficas

- Barbagallo, D., Bruni, L., Francalanci, C., & Giacomazzi, P. (2012). An empirical study on the relationship between Twitter sentiment and influence in the tourism domain. In M. Fuchs, F. Ricci & L. Cantoni (Eds.), *Information and communication technologies in tourism 2012*. Helsingborg, Sweden.
- Bardin, L. (2011). *Análise de conteúdo*. São Paulo: Edições 70.
- Bastone, J. (2011). *Taxonomies, tone and tweets: Where text analytics and social marketing collide*. Acedido a 28 de dezembro de 2013, em <http://support.sas.com/resources/papers/proceedings11/321-2011.pdf>
- Becker, H., Naaman, M., & Gravano, L. (2011). Selecting quality Twitter content for events. In R. L. A. Adamic, A. Baeza-Yates & S. Counts (Eds.), *Proceedings of the Fifth International Conference on Weblogs and Social Media (ICWSM)*. Barcelona: The AAAI Press.
- Beni, C. M. (2004). *Análise estrutural do turismo*. São Paulo: Senac.
- Bhuiyan, S. I. (2011) Social Media and its effectiveness in the political reform movement in Egypt. *Middle East Media Educator*, 1(1), 14-20.
- Cooper, C., Fletcher, J., Fyall, A., Gilbert, D., & Wanhill, S. (2001). *Turismo: Princípios, práticas e filosofia*. Porto Alegre: Bookman.

- Ebner, M., Mühlburger, H., Schaffert, S., Shiefner, M., Reinhardt, W., & Wheeler, S. (2010). Getting granular on twitter: Tweets from a conference and their limited usefulness for non-participants. In N. Reynolds & M. Turcsán-Szabó (Eds.), *Key competencies in the knowledge society: IFIP TC 3 International Conference* (Vol. 34, pp. 102-113). Brisbane: Springer.
- Fotis, S. J., Buhalis, D., & Rossides, N. (2011). Social media impact on holiday travel planning: The case of the Russian and the FSU markets. *International Journal of Online Marketing*, 1(4), 1-19.
- Freund, F.T. (2005). *Alimentos e bebidas: Uma visão gerencial*. Rio de Janeiro: Senac Nacional.
- Gil, A. C. (1999). *Métodos e técnicas de pesquisa social*. São Paulo: Atlas.
- Heverin, T. & Zach, L. (2010). Twitter for city police department information sharing. *Proceedings of the ASIS&T 2010 Annual Meeting*. Pittsburgh, PA, USA.
- Jansen, B. J., Zhang, M., Sobel, K., & Chowdury, A. (2009). Twitter power: Tweets as electronic word of mouth. *Journal of the American Society for Information Science and Technology*, 60(11), 2169-2188.
- Lage, B. H. G., & Milone, P.C. (2000) *Turismo: teoria e prática*. São Paulo: Atlas.
- Microsoft (2013). *Microsoft*. Acedido em 28 de dezembro de 2013, em <http://www.microsoft.com>
- Ministério do Turismo (2007). *Programa de regionalização do turismo: Roteiros do Brasil – Módulo operacional 8 'Promoção e apoio à comercialização'*. Brasília: Ministério do Turismo.
- Pak, A., & Paroubek, P. (2010, 17-23 maio). *Twitter as a corpus for sentiment analysis and opinion mining*. Artigo apresentado na Language Resources and Evaluation Conference, Valletta.
- Petrocchi, M. (2009). *Turismo: Planejamento e gestão*. São Paulo: Pearson Education.
- Repustate (2013) *Repustate*. Acedido em 21 de março de 2013, em <http://www.repustate.com>
- Rubio, K. (2007). *Megaeventos esportivos, legado e responsabilidade social*. São Paulo: Casa do Psicólogo.
- Sentiment140 (2013) *Sentiment140*. Acedido em 28 dezembro de 2013, em <http://help.sentiment140.com/home>
- Sentimentanalyzer (2013). *SentimentAnalyzer*. Acedido em 21 de março 2013, em <http://sentimentanalyzer.appspot.com>.
- Sentimentor. (2013). *Sentimentor*. Acedido em 21 de março 2013, em <http://sentimentor.co.uk>
- Sentistrength. (2013). *SentiStrength*. Acedido em 21 de março 2013, em <http://sentistrength.wlv.ac.uk/>
- Shimada, K., Inoue, S., Maeda, H., & Endo, T. (2011, 19-20 dezembro). *Analyzing tourism information on twitter for a local city*. Artigo apresentado em 1st ACIS international symposium on software and network engineering. Proceedings of the ACIS International Symposium on Software and Network Engineering, Seul.
- Socialbakers (2013). *Reports*. Acedido em 28 dezembro 2013, em <http://www.socialbakers.com>
- Text-processing (2013) *Sentiment analysis with Python NLTK text classification*. Acedido em 21 de março de 2013, em <http://text-processing.com/>
- Tweenator (2011) *Tweenator*. Acedido em 21 de março de 2013, em <http://www.tweenator.com>
- Twitter (2013). *Twitter*. Acedido em 28 de dezembro de 2013, em <http://www.twitter.com>
- Vieweg, S., Hughes, A. L., Starbird, K., & Palen, L. (2010). Micro-blogging during two natural hazards events: What twitter may contribute to situational awareness. *Proceedings of the ACM Conference on Human Factors in Computing Systems*, 28, 1079-1088.