
Identificação de personalidade de atletas olímpicos: uma análise exploratória de narrativa com mineração de textos

*Katia Rubio,
Ivan Santana Rabelo,
Roberta Akemi Sinoara,
Renan Santos Barbosa,
Solange Oliveira Rezende*

Resumo

A capacidade de imaginar diferencia os humanos de qualquer outro ser animado. Analisar o passado, planejar o futuro e atuar no presente a partir de experiências vividas é uma prerrogativa humana que permitiu à espécie chegar ao nível de desenvolvimento conhecido até o momento. As questões relacionadas ao Olimpismo têm mobilizado a sociedade e a comunidade acadêmica. Assim, considerando-se as características particulares de atletas olímpicos, neste trabalho é apresentado um estudo preliminar da relação entre traços de personalidade e padrões de textos extraídos de entrevistas transcritas. São apresentadas análises conjuntas entre os resultados de um teste psicológico de traços de personalidade e uma análise exploratória do texto de uma entrevista realizada com um atleta olímpico. Tal análise do texto é realizada com o apoio de métodos de processamento automático de textos. Acredita-se que os resultados corroboram para o estudo de associações entre o conteúdo mais frequentemente presente em narrativas e o cruzamento de dados com luz às teorias psicológicas de traços de personalidade.

Palavras-chave: personalidade, esporte, atleta olímpico, entrevista, mineração de textos.

Identification of the personality of Olympic athletes: an exploratory analysis of narrative with text mining

Katia Rubio, Ivan Santana Rabelo, Roberta Akemi Sinoara, Renan Santos Barbosa, Solange Oliveira Rezende

Abstract

The ability to imagine differentiates humans from any other animate being. Analyzing the past, planning the future and acting in the present from lived experiences is a human prerogative that allowed the species to reach the level of development known to date. Issues related to Olympism have mobilized society and the academic community. Thus, considering the specificities of Olympic athletes, this paper presents a preliminary study of the relationship between personality traits and patterns of texts extracted from transcribed interviews. We present joint analyzes between the results of a psychological test of personality traits and an exploratory analysis of the transcription of an interview with an Olympic athlete. This analysis was supported by the application of automatic text processing methods. It is believed that the results corroborate for the study of associations between the content more frequently present in narratives and the crossing of data with light to the psychological theories of personality traits.

Key-words: Personality, sport, Olympic athlete, interview, text mining.

Identificación de la personalidad de los atletas olímpicos: un análisis exploratorio de la narrativa con minería de textos

Katia Rubio, Ivan Santana Rabelo, Roberta Akemi Sinoara, Renan Santos Barbosa, Solange Oliveira Rezende

Resumen

La capacidad de imaginar diferencia a los humanos de cualquier otro ser. Analizar el pasado, planificar el futuro y actuar en el presente a partir de experiencias vividas es una prerrogativa humana que permitió a la especie alcanzar el nivel de desarrollo conocido hasta la fecha. Las cuestiones relacionadas con el olimpismo han movilizad a la sociedad y la comunidad académica. Por lo tanto, teniendo en cuenta las especificidades de los atletas olímpicos, este artículo presenta un estudio preliminar de la relación entre los rasgos de personalidad y los patrones de textos extraídos de entrevistas transcritas. Presentamos análisis conjuntos entre los resultados de una prueba psicológica de rasgos de personalidad y la análisis de lo texto de una entrevista con un atleta olímpico. Este análisis fue apoyado por métodos de procesamiento automático de textos. Se cree que los resultados corroboran el estudio de las asociaciones entre el contenido más frecuentemente presente en las narraciones y el cruce de datos con la luz de las teorías psicológicas.

Palabras-clave: personalidad, deporte, atleta olímpico, entrevista, minería de textos.

Introdução

Mundialmente alguns autores apontam que a avaliação de aspectos psicológicos no esporte está pautada em procedimentos tais como observações (presenciais e por meio de tecnologia-filmagem), entrevistas, experimentos em laboratório, experimentos pedagógicos e testes psicológicos, contribuindo para o psicodiagnóstico esportivo. Observa-se que na perspectiva da relação psicofísica residem grandes debates e discussões acerca da cultura contemporânea do esporte e da expressão de crenças e valores, muitas vezes já considerados transculturais (Rubio, 2007).

Weinberg e Gould (2001) apresentaram quatro formas de estudo da personalidade especificamente no contexto do esporte, as quais denominaram de abordagem, a) Abordagem psicodinâmica - estudo da personalidade no esporte, embasado na perspectiva clínica e na psicanálise; b) Abordagem de traço - estudo da personalidade no esporte, partindo-se do pressuposto da permanência e constância da personalidade independente das situações ou circunstâncias; c) Abordagem situacional: abordagem que argumenta que o comportamento no esporte é, em grande parte, determinada pela situação ou ambiente. Noção de "estado" de humor; e d) Abordagem interacional: estudo da interação do traço de personalidade e o estado de humor, atuando em diversas situações específicas do esporte.

O modelo dos Cinco Grandes Fatores (CGF) da personalidade representa uma tentativa de descrever um conjunto amplo e complexo de características de personalidade mediante cinco traços ou fatores básicos universais (Costa & McCrae, 1985; Digman, 1990). Empiricamente, os CGF tendem a se mostrar correlacionados, uma vez que todos os fatores expressam aspectos relacionados à aptidão à vida social (Just, 2011). Não obstante, à parte a discussão em andamento sobre a ortogonalidade versus obliquidade do modelo, as suas cinco dimensões básicas têm sido replicadas em diversas culturas e ao longo de diferentes etapas do desenvolvimento. Portanto, trata-se de um dos principais modelos que embasam a avaliação da personalidade em seres humanos e, como sugerem alguns autores, também em outras espécies animais (King, Weiss, & Sisco, 2008).

Quando classificada como método, a história oral implica em um conjunto de procedimentos éticos que faz parte do trabalho de busca, armazenamento e utilização dos depoimentos, o qual conta com a concordância do colaborador sobre o uso de suas narrativas mediante termo de consentimento, bem como o compromisso assumido pelo pesquisador de retornar os bens da pesquisa ao colaborador (Ferreira Junior, 2013; Rubio, 2013, 2001).

Portanto, narrativas e histórias de vidas são fontes valiosas que estão presentes em vários momentos da pesquisa acadêmica (Rabelo & Rubio, 2018). Conforme destacado em Rubio (2014), quando os atletas referem-se a sua trajetória esportiva eles trazem em suas narrativas a lembrança de pessoas e profissionais que influenciaram e determinaram o desejo pelo esporte, pela busca de melhores condições de vida e de treinamento ou a convivência com outros atletas que também competiam naquele momento histórico e cujas carreiras se cruzaram apontando para a necessidade premente de contextualizar essas situações para promover o entendimento de episódios marcantes de suas vidas e de seus resultados. Em se tratando de atletas de modalidades coletivas essa condição é ainda mais evidente porque vários deles narram suas memórias sobre um mesmo

conteúdo vivido a partir de diferentes pontos de vista, apresentando novos conteúdos, uma nova história, apontando para a subjetividade que envolve a construção e elaboração desse tema ainda que vivido coletivamente.

Em um ensaio intitulado "O que é comunicação", presente no livro "O mundo codificado" de Flusser (2007, p. 97), é aplicada uma distinção entre comunicação dialógica e comunicação discursiva. Segundo o autor, para produzir informação, os homens trocam diferentes informações disponíveis na esperança de sintetizar uma nova informação, se referindo à forma de comunicação dialógica. Já com o fim de preservar, manter a informação, os homens compartilham informações existentes na esperança de que elas, assim compartilhadas e compartilhadas, possam resistir melhor ao efeito entrópico da natureza, tratando-se da forma de comunicação discursiva.

Porém, conforme enfatizado por Flusser (2007, p. 98), uma das grandes dificuldades é "produzir diálogos efetivos, isto é, de trocar informações com o objetivo de adquirir novas informações". Sobretudo quando o volume de informações é muito alto, dificultando ainda mais a interpretação destas informações contidas em narrativas discursivas, seja com o fim de produzir e descobrir novas informações, seja com o fim de manter estas memórias ao longo do tempo.

Pode-se afirmar que o uso de métodos manuais para análise e obtenção de conhecimento a partir de dados torna a tarefa dispendiosa, em termos financeiros ou de tempo, totalmente subjetiva e muitas vezes inviável considerando-se um grande volume de dados. Neste contexto, técnicas especializadas de extração automática de conhecimento têm sido desenvolvidas para a extração de conhecimento de base de dados, sendo referenciadas na literatura como knowledge discovery in databases (KDD), data mining ou mineração de dados (Fayyad et al., 1996). Segundo Rezende (2003), a mineração de dados é uma área multidisciplinar, que incorpora técnicas utilizadas em áreas como Inteligência Artificial, especialmente Aprendizado de Máquina, Base de Dados e Estatística. Com o processo de mineração de dados, ou seja, o processo de extração de conhecimento de grande volume de dados, busca-se a descoberta de conhecimento com a identificação de padrões novos, válidos, úteis e, muitas vezes, compreensíveis.

Quando os dados disponíveis são textos escritos em língua natural, ou seja, dados não estruturados, o processo de extração de conhecimento é chamado de mineração de textos ou text mining. Assim, a mineração de textos pode ser vista como uma especialização da mineração de dados, correspondendo à aplicação de um conjunto de técnicas para descoberta de conhecimento inovador em textos (Aggarwal & Zhai, 2012). O processo de mineração de textos pode ser visto como uma sequência de etapas genéricas, que devem ser instanciadas de acordo com os dados disponíveis e o conhecimento que se espera obter (Aggarwal & Zhai, 2012; Rezende, 2003). As possíveis aplicações desse processo são diversas, como, por exemplo, análise exploratória de conteúdo textual, organização de coleções de documentos, análise de sentimentos e sistemas de recomendação (Sinoara et al., 2019; Marcacini et al., 2018; Sinoara, Antunes & Rezende, 2017; Sinoara, Scheicher & Rezende, 2017). A aplicação do processo de mineração de textos em análises exploratórias de documentos, como a

realizada neste trabalho, possibilita a identificação e o mapeamento de padrões no conteúdo de documentos.

Independente da aplicação ou da técnica utilizada, a natureza não estruturada dos dados traz grandes desafios para o processo de mineração de textos. Nas línguas naturais, as palavras, bem como as sentenças, apresentam diversas relações que podem impactar no significado dos textos, tais como sinonímia, hiperonímia, polissemia e ambiguidade (Pietroforte, 2010). Essas relações semânticas entre palavras e sentenças influenciam como as pessoas interpretam os textos e podem ser importantes para a mineração de textos. O entendimento de textos escritos em língua natural é um processo complexo, que se dá por meio do conhecimento das palavras e de seus significados, das relações existentes entre as palavras, bem como do conhecimento de mundo e do contexto no qual o texto foi escrito. Assim, recursos e técnicas de processamento de língua natural podem auxiliar o processamento de textos. Entre as análises sintática e semântica de textos que podem ser realizadas automaticamente pode-se citar a identificação de classes morfossintáticas (como substantivos e verbos) e a identificação de papéis semânticos (como os agentes de ações descritas por verbos) para as palavras presentes em documentos.

Neste contexto, busca-se realizar uma análise exploratória de documentos, com o objetivo de identificar padrões em narrativas e a relação destes padrões com traços de personalidade. Neste trabalho, técnicas de mineração de textos são utilizadas para apoiar a análise dos textos e, assim, verifica-se também a viabilidade do uso da mineração de textos como um mecanismo de apoio ao processo de identificação de traços de personalidade. Constituído como um dos principais fenômenos sociais contemporâneos, o esporte tem se estabelecido como um campo privilegiado de estudo e intervenção, seja nos aspectos específicos de sua prática tática e técnica, como também nos aspectos educativo e sociocultural.

Métodos

Nesta fase da pesquisa de desenvolvimento de uma metodologia de identificação de traços de personalidade em narrativas de atletas, buscou-se realizar uma análise exploratória da transcrição de uma entrevista com um atleta olímpico, realizada com o apoio de técnicas de mineração de textos, frente aos resultados de um teste psicológico de traços de personalidade. Essa análise pode ser vista como um estudo piloto das possibilidades de busca de palavras-chave na narrativa discursiva do atleta, não se propondo, neste manuscrito, ser uma análise exaustiva sobre a metodologia. É necessário, assim que definido a metodologia, investigar em um número muito maior de atletas, inclusive com o objetivo de analisar, por exemplo, comunalidades entre traços de personalidade em mesmas modalidades esportivas.

A hipótese da pesquisa é de que seja possível identificar características associadas às emoções, sobretudo, de traços de personalidade, a partir da aplicação de técnicas de mineração de textos nas entrevistas realizadas com atletas. Para trabalhar essa hipótese, buscando indicativos da validade do uso da mineração de textos neste estudo piloto, definiu-se o método apresentado na Figura 1. A ideia central é realizar uma análise comparativa entre os resultados obtidos com a realização de um

teste psicológico e termos identificados na narrativa da história de vida de um atleta, obtida por meio de entrevista.

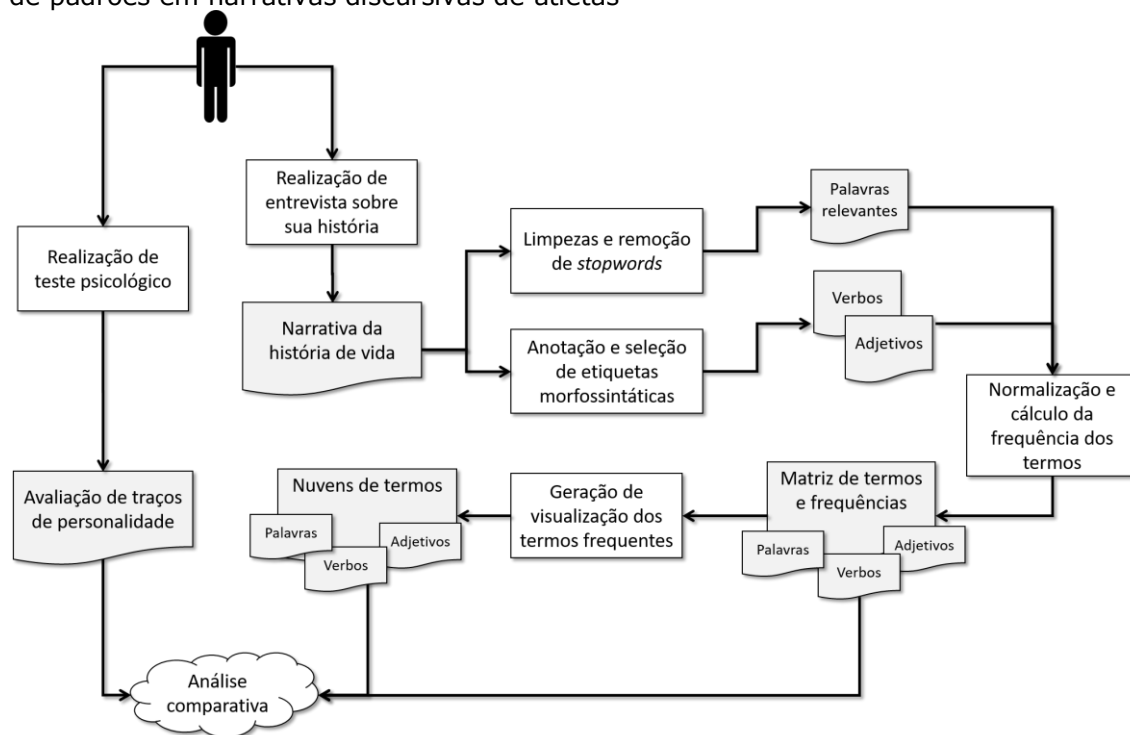
Para o tratamento das palavras, são realizadas atividades de pré-processamento comuns em mineração de textos, realizando uma limpeza dos textos, como remoção de pontuações, de números e de stopwords. A remoção de stopwords visa a eliminação de palavras que não trazem informação relevante para o processo de mineração de textos, de acordo com os objetivos estabelecidos. Essas palavras, chamadas de stopwords, normalmente são palavras que possuem as funções de artigos, preposições, pronomes e conjunções. No entanto, também podem ser identificadas stopwords específicas do domínio de aplicação do processo, ou seja, palavras que sabidamente são frequentes na coleção e que não distinguem classes ou grupos que se espera identificar com a mineração de textos.

Para a identificação automática de verbos e adjetivos realiza-se a anotação morfossintática. A anotação morfossintática é uma tarefa de processamento de língua natural na qual é realizada a atribuição de etiquetas morfossintáticas às palavras de uma sentença. Essas etiquetas identificam a função sintática de cada palavra, tais como substantivo, verbo, adjetivo e preposição. Essa é uma tarefa importante do processamento de língua natural e que serve como base a várias outras tarefas.

Outra atividade importante nesse processo é a normalização dos termos. A normalização, tanto das palavras quanto dos verbos e adjetivos, pode ser realizada com um processo de radicalização. Esse processo reduz cada palavra ao seu radical (ou palavra raiz) e é realizada com o objetivo de eliminar as diversas variações que as palavras podem sofrer, como por exemplo variações de gênero e número dos substantivos e as conjugações dos verbos. Com isso, a dimensionalidade (número de termos) do documento é reduzida, facilitando a identificação de padrões interessantes. Visando melhorar a interpretabilidade dos termos, após a radicalização, o radical pode ser substituído pela palavra completa (palavra, verbo ou adjetivo) mais frequente no corpus.

Após a normalização é realizada a contagem automática da frequência dos termos, sejam eles: palavras, verbos ou adjetivos normalizados. Com essa contagem, gera-se uma matriz de frequências para cada abordagem. Para facilitar a visualização dos dados, pode ser gerada uma nuvem de termos para cada matriz. Nessa técnica de visualização de dados, conhecida como nuvem de palavras, o tamanho de cada palavra corresponde à frequência da mesma no texto analisado. Portanto, quanto maior a frequência de uma palavra maior será o seu tamanho na nuvem de palavras. A partir das nuvens de termos e das matrizes de frequências é realizada uma análise comparativa com os resultados do teste psicológico.

Figura 1. Método para verificação do uso de mineração de textos para identificação de padrões em narrativas discursivas de atletas



Instrumentos

Medida de personalidade - A avaliação da personalidade foi realizada pelo teste Bateria Fatorial de Personalidade (BFP) que avalia traços de personalidade por meio do modelo dos Cinco Grande Fatores (CGC). A escala possui 126 itens, respondidos em escala *Likert* de sete pontos, na qual o sujeito indica a concordância com as afirmativas, que descrevem a expressão da personalidade por meio da forma como as pessoas pensam, agem e sentem (Nunes, Hutz, & Nunes, 2010). A BFP avalia os fatores Extroversão, Socialização, Realização, Neuroticismo e Abertura.

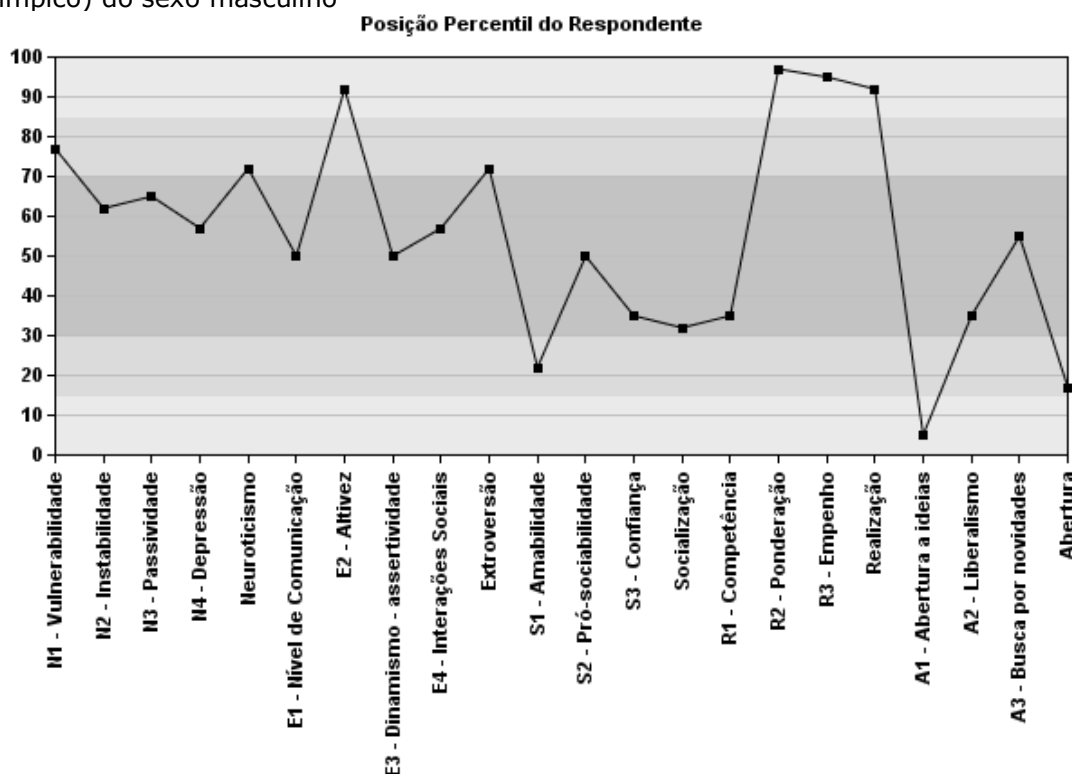
Mineração de texto - Conforme apresentado na Figura 1, como uma análise preliminar da adequação do uso de técnicas mineração de textos para a análise de narrativas de atletas em busca da identificação de relações entre o conteúdo de tais narrativas e os resultados de avaliações de personalidade, realizou-se a análise dos textos em três abordagens: palavras, verbos e adjetivos. A anotação morfossintática foi realizada com o modelo pré-treinado para o idioma português da ferramenta *nlpnet* (Fonseca & Rosa, 2013) e a radicalização dos termos foi realizada por meio do *Snowball Stemmer*. Todo o pré-processamento dos textos foi realizado com o apoio do pacote *Text Mining Infrastructure in R*, pacote *tm* da ferramenta *R* (Feinerer et al, 2008). Já as nuvens de palavras foram geradas com o apoio da ferramenta *Wordle* (Feinberg, 2010).

Resultados e discussão

Considerar nas histórias de vidas as histórias das emoções possibilita a compreensão não só da rede comunitária da qual o narrador faz parte, mas também, entender as “comunidades emocionais” – fundamentalmente o

mesmo que comunidades sociais – famílias, bairros, sindicatos, instituições acadêmicas, entre outros (Rosenwein, 2011; Rubio, 2013). Assim, as análises preliminares apresentadas neste manuscrito, contribuem para os estudos iniciais de identificação de traços de personalidade, levantados com um atleta olímpico, da modalidade Tênis de mesa, por meio da análise do cruzamento dos dados apresentados pelo participante no relatório informatizado de correção do teste psicológico e, em seguida, a apresentação da extração de adjetivos da própria entrevista realizada com o atleta. Na Figura 2 são apresentados os dados coletados de um atleta olímpico do Tênis de Mesa no teste psicológico BFP.

Figura 2. Perfil de personalidade de um atleta da modalidade Tênis de Mesa (olímpico) do sexo masculino



A seguir, na Figura 3, é possível observar uma representação gráfica, nuvem de palavras, gerada a partir da extração e tratamento de adjetivos extraídos da entrevista transcrita do referido atleta. Reforça-se, contudo, que o objetivo deste manuscrito é de estudar possibilidades de aplicação de metodologias para extração de dados via mineração de textos de entrevistas das narrativas discursivas de atletas, sem propor, neste momento, qualquer psicodiagnóstico ou avaliação psicológica ou de personalidade deste respondente em específico.

Análises preliminares dos termos obtidos mostraram que os adjetivos apresentam um potencial maior para serem indicadores de personalidade, portanto, neste artigo, é apresentada apenas a análise dos adjetivos. Esse fato é consistente com pesquisas anteriores que apontam a existência de marcadores de personalidade. Nesse contexto, o termo 'marcadores' designa adjetivos correntemente empregados para descrever

características salientes, presentes nos indivíduos de uma determinada cultura e que, em geral, remetem às dimensões básicas dos Cinco Grandes Fatores de Personalidade. Se a avaliação da personalidade via marcadores linguísticos é uma parte intrínseca à abordagem léxica dos aspectos fundamentais da personalidade humana e os adjetivos são tratados como fortemente atrelados aos aspectos culturais de uma população, ainda, agrega-se a isso, a vantagem de que adjetivos como marcadores são especialmente convenientes em se tratando de obter instrumentos breves de avaliação (McCrae & John, 1992; Machado et al., 2014).

Os pesquisadores Hauck, Machado, Teixeira e Bandeira (2012), por exemplo, desenvolveram uma medida breve para a avaliação da personalidade no modelo dos CGF, baseado em 64 marcadores utilizando adjetivos associados aos traços de personalidade (Hutz et al., 1998), uma medida análoga aos marcadores de Goldberg (1992), o instrumento possui 25 adjetivos-marcadores, sendo cinco para cada um dos fatores de personalidade. Em um estudo independente, o instrumento se mostrou também adequado para o uso com adolescentes, embora os autores tenham sugerido utilizar apenas quatro itens para cada dimensão (Machado et al., 2014).

A nuvem de palavras (Figura 3) permite uma exploração visual de padrões presentes no discurso do atleta, sendo que quanto mais frequente é um termo no discurso, maior é o seu tamanho na visualização. Conforme apresentado no método da Figura 3, foram analisados termos referentes a palavras, verbos e adjetivos normalizados.

Figura 3. Nuvem de termos gerada com os adjetivos normalizados identificados na entrevista transcrita



Portanto, continuando com os objetivos deste manuscrito, procurou-se analisar a frequência dos termos presentes na entrevista do referido atleta, que podem ser verificados nas Tabelas 1 e 2. Nessas tabelas os termos referem-se aos adjetivos presentes na entrevista após normalização.

Tabela 1. Termos mais frequentes na entrevista transcrita do atleta

Freq.	Termo	Freq.	Termo	Freq.	Termo	Freq.	Termo
34	grande	9	vencido	6	adulta	5	seguinte
33	melhor	8	aberto	6	curiosas	5	triste
29	brasileira	8	forte	6	nova	5	voltado
23	primeira	8	ganhado	6	ótima	4	chamada
20	juvenil	8	ininteligível	5	americano	4	classificado
16	olímpica	8	maior	5	certo	4	feliz
15	individual	8	paulista	5	duro	4	fraca
14	Bom	8	segunda	5	européu	4	intercontinental
14	Latina	8	técnica	5	infantil	4	interessada
14	mundial	8	última	5	necessária	4	internacional
13	importada	7	latinoamericana	5	pequena	4	motivada
9	diferente	7	passada	5	profissional	4	nervosa
9	Difícil	7	rápida	5	ruim	4	psicológica

Tabela 2. Termos menos frequentes na entrevista transcrita do atleta

Freq.	Termos
3	acostumado, argentino, boa, dado, decidido, francesa, impossível, legal, lotado, naturalizada, panamericana, ranqueados, sozinho, sul
2	animal, apoiado, atual, bons, carentes, chinês, claro, contratado, convidado, curto, devagarzinho, direito, doente, enciumado, filiado, física, focada, fominha, francês, gratificante, histórico, ideal, internacionais, inúmeras, japonesas, livre, lógico, machucado, máximo, necessária, nordeste, normal, novinho, organizado, perdido, presente, remunerado, seguidos, seletiva, simples, sudeste, terceiro, tranquilo
1	adequada, agente, alojado, alto, anterior, assado, ativo, balanceada, barato, capacitados, carentes, chateado, chegado, classificatório, comec, complicado, confiante, conhecido, consecutivos, conseguido, continuado, corrido, criado, décima, dedicado, deitado, desenvolvido, desgastante, devagarzinho, diferenciadas, disputados, dividido, doendo, empolgado, enfileirados, enorme, errada, escolhido, esgotado, especial, experiente, falecido, finais, firme, futuras, geral, gostado, grave, horrível, húngaro, igual, inédito, investido, invicto, jantado, japonês, jogado, juntos, juvenil, largado, ligado, limpa, local, maduro, máximo, medicado, nacionais, oeste, paitrocínio, perfeccionista, perfeito, pessoal, pior, positiva, possível, potencial, privadas, próximas, públicas, rankeado, real, recordista, recuperado, regionais, resultado, satisfeito, soltinha, sulamericanos, surpreendido, tensos, terminas, terrível, torneiozinhos, ultima, única, utópica, valorizado, velha, vivo

Levando em consideração se tratar de um estudo inicial, exploratório, de verificação de metodologias para análises qualitativas e quantitativas de termos extraídos de entrevistas de atletas, e não propondo ser um psicodiagnóstico de personalidade, mas, e sobretudo, uma análise da relação entre os dados capturados pelo teste de personalidade, em formato de autorrelato (escala fatorial) e os dados minerados, por meio de adjetivos, à partir das transcrições das narrativas, a seguir será feita uma relação entre as principais palavras presentes e sua frequência no discurso em relação aos traços de personalidade observados por meio do gráfico e do relatório do atleta no teste.

Portanto, observa-se uma alta frequência de palavras, tais como, grande (34), melhor(33), bom(14), vencido(9), maior(8), forte(8),

ganhado(8), técnica(8), rápida(7), certo(5), profissional(5), classificado(4), motivada(4), decidido(3), ranqueados(3), que somariam um total de 153 palavras, que poderiam estar mais associadas ao Fator Realização, também chamado de Conscienciosidade. Ao mesmo tempo, se for verificado no gráfico, por meio da Figura 2, os dados relacionados ao Fator Realização apresentados pelo atleta no teste de autorrelato de personalidade, ficam classificados como “muito altos” também nesse fator. De maneira que acredita-se haver uma relação, para um sujeito que apresenta escores muito elevados neste fator de Conscienciosidade, um número mais elevado de palavras associadas a adjetivos relacionados ao conteúdo interpretativo deste fator.

De acordo com a interpretação, segundo Nunes et al. (2010), o fator se refere a quão ativamente as pessoas buscam atingir seus objetivos, bem como a predisposição para fazer sacrifícios pessoais para tanto. Sugere uma percepção realista sobre si mesmo no tocante à sua capacidade de realização de acordo com a situação. Elevada dedicação às atividades profissionais e acadêmicas. Relaciona-se a pessoas que geralmente gostam de obter reconhecimento por seu esforço e que tendem ao perfeccionismo. Também descrevem uma tendência a planejar detalhadamente os passos para a realização de alguma atividade e a necessidade de revisões cuidadosas dos trabalhos antes de expô-los aos outros. Ainda em acordo com isso, verificou-se transcrito a partir da entrevista, ainda que em menor frequência, as palavras: focada(2), organizado(2), experiente(1), firme(1), dedicado(1), perfeccionista(1), empolgado(1), com baixa frequência, mas também presentes na transcrição da entrevista com o atleta. Tais palavras se mostram associadas também ao fator de personalidade Realização.

Ainda sobre o fator Realização, segundo autores que têm trabalhado com marcadores reduzidos de traços de personalidade, entre eles, Hutz et al. (1998), Hauck, Machado, Teixeira e Bandeira (2012), Machado et al. (2014), os cinco marcadores de traço associados a este fator estão os adjetivos: dedicado, esforçado, responsável, organizado, cuidadoso. Assim, acredita-se, portanto, haver uma relação entre este fator de personalidade demonstrado pelo sujeito em um teste padronizado de autorrelato, com os adjetivos mais presentes nas suas narrativas discursivas de memórias de vida, por meio de mineração de texto da entrevista com o mesmo.

No mesmo sentido, porém, em relação ao Fator Extroversão, verifica-se por meio da Figura 2, que o sujeito apresentou um valor classificado como “muito alto” também no subfator (faceta) da extroversão, denominado de Altivez. Segundo explicações sobre o referido fator, na faceta Altivez escores muito altos podem relacionar-se a necessidade de receber atenção dos outros, tendência a supervalorizar suas experiências, predisposição para falar excessivamente sobre si e sobre suas qualidades e posses. Em consonância, verificam-se palavras presentes no discurso, tais como, grande (34), melhor(33), bom(14), maior(8), que poderiam indicar uma propensão a supervalorização de discursos associados a grandeza, elevada autoestima. Outras palavras também presentes na entrevista foram os objetivos legal(3), confiante(1), enorme(1), especial(1), perfeito(1), que também estão relacionados a escores de extroversão, sobretudo a respeito do fator altivez.

Um outro fator que se destacou com classificação fora da média, com palavras associadas ao seu significado empírico, refere-se a faceta Amabilidade, integrante da Socialização, haja vista que os resultados do teste de personalidade demonstraram rebaixada classificação neste componente, podendo ser retratada na entrevista, por meio da mineração de palavras, tais como, individual(15), sozinho(3) e enciumado(2).

Ainda quanto a interpretação de outros fatores que apresentaram algum traço fora da média da padronização do teste BFP neste atleta, destaca-se por fim o fator Abertura, já que o atleta apresentou classificação rebaixada na faceta Abertura a ideias, o que indicaria uma pessoa pouco curiosa quanto a novos temas, mais conservadoras e fiéis a seus gostos artísticos, ao mesmo tempo, que foram minerados adjetivos da entrevista do atleta, com palavras tais como, diferente(9), aberto(8), nova(6), curiosa(6). Isso poderia indicar uma contradição a relação da presença de adjetivos contrários ao traço captado pelo teste.

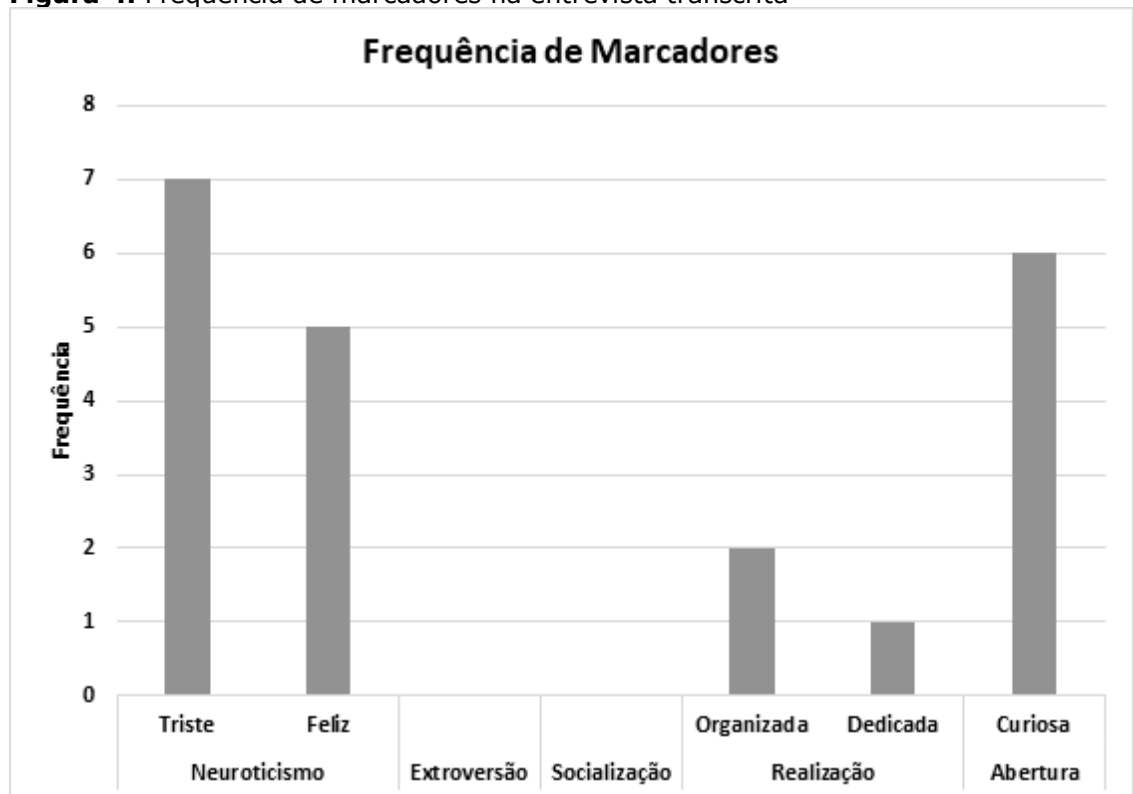
No entanto, conforme relatório do teste o atleta apresentou resultados dentro da média na faceta Liberalismo, ainda dentro do mesmo fator Abertura, o que indicaria também um padrão comportamental adequado com valores morais e sociais, mas também ciente da noção de que estes valores podem ser relativizados, que podem mudar ao longo do tempo e ser diferentes em várias culturas e regiões. No mesmo sentido, na faceta Busca por novidades, revelou um padrão comportamental também dentro da média no que se refere ao quanto gosta e busca vivenciar novos eventos e ações, costuma lidar bem com a quebra da rotina, apesar de também não se importarem quando essa mudança não ocorre (Nunes et al., 2010). De maneira que estas palavras mineradas, associadas a abertura, podem ser em razão de um padrão mediano nos demais traços associados ao fator Abertura, que equilibrariam esta única faceta muito rebaixada, de abertura a ideias.

Por fim, ainda quanto a tentativa de relacionar qualitativamente os resultados do atleta no teste de personalidade e as palavras mineradas da entrevista transcrita, os resultados de Neuroticismo mostram-se, no atleta, segundo pode ser observado na Figura 2, dentro dos escores medianos de padronização do teste na amostra normativa. Ainda que, considerando ainda dentro da média, os resultados mostraram-se no quadrante “médio alto”, o que poderia indicar já um padrão um pouco mais elevado nestas características. O que fica ainda mais evidenciado, quanto a faceta “Vulnerabilidade ao sofrimento”, que mostrou-se no atleta com classificação “alto”.

Segundo o manual do teste BFP (Nunes et al., 2010), altos escores em vulnerabilidade pode indicar baixa autoestima e algum medo de que pessoas importantes venham a deixá-los em decorrência de seus erros, ou mesmo, mostrar-se capazes de atitudes que vão contra a sua vontade, relatar insegurança, dependência de pessoas próximas e certa dificuldade ao tomar algumas decisões, mesmo em situações cotidianas. No mesmo sentido, observa-se na mineração de adjetivos, a presença das palavras: triste(5), ruim(5), nervosa(4), fraca(4), impossível(3), sozinho(3), carente(3), doente(2), machucado(2), chateado(1), complicado(1), desgastado(1), pior(1) entre outras com baixa frequência.

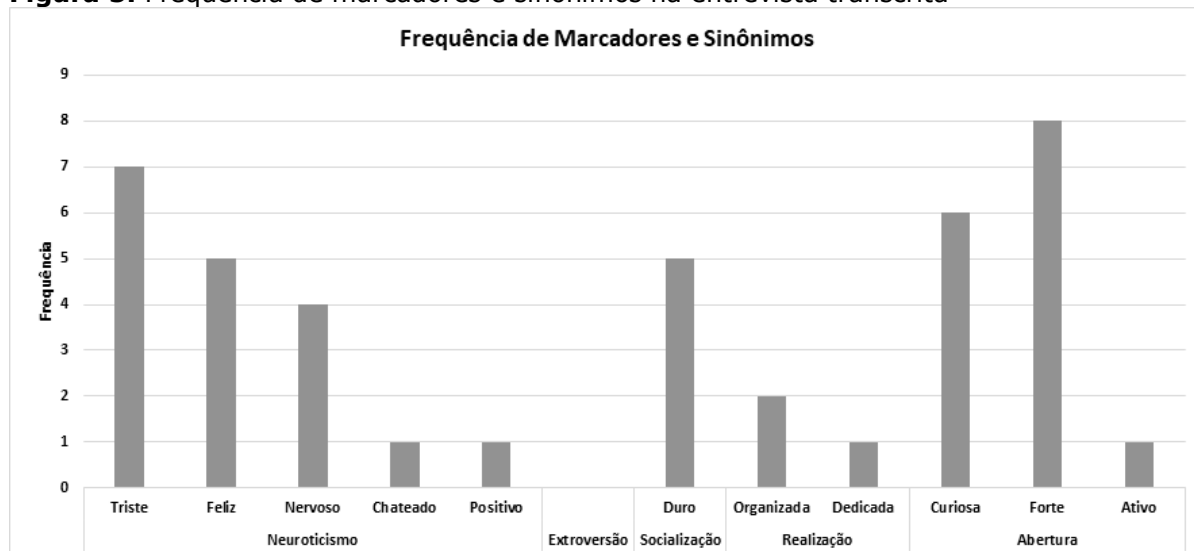
Considerando as relações encontradas entre os dados obtidos com o teste de personalidade aplicado e os adjetivos presentes na narrativa do atleta, também foi realizada uma análise sobre a presença dos 64 marcadores indicados por Hutz et al. (1998). Na narrativa do atleta foram identificadas 696 ocorrências de adjetivos, sendo apenas 21 as ocorrências de marcadores de personalidade, considerando-se também as variações de gênero e de número dos adjetivos. Apenas 6 dos 64 marcadores foram observados na narrativa do atleta, conforme é apresentado na Figura 4. Pode-se verificar que os marcadores identificados correspondem a 3 fatores: Neuroticismo, Realização e Abertura. Caso sejam considerados apenas os 25 marcadores apresentados por Machado et al. (2014), apenas o fator Realização é identificado por meio dos marcadores “Dedicada” e “Organizada”. Vale notar que o fator Realização é o fator com maior posição percentil no teste psicológico realizado com o atleta (Figura 2).

Figura 4. Frequência de marcadores na entrevista transcrita



Uma segunda análise dos marcadores foi realizada considerando-se também sinônimos dos 64 marcadores. Com isso, foram identificadas 41 ocorrências de marcadores e sinônimos na narrativa do atleta, conforme é apresentado na Figura 5. Apesar do acréscimo na quantidade de ocorrências, ainda não foi observada nenhuma ocorrência para o fator Extroversão.

Figura 5. Frequência de marcadores e sinônimos na entrevista transcrita



Em resumo, parece haver uma relação entre os dados apresentados por meio do teste de personalidade e os adjetivos mais frequentes presentes na entrevista do atleta. Contudo, a análise realizada indica a necessidade de ir além dos 64 marcadores de personalidade, buscando o tratamento de aspectos semânticos dos textos. Reforça-se que estes são estudos exploratórios iniciais, com o objetivo de entender as relações entre as variáveis associadas a investigação da personalidade e a maneira como isso pode também ser expresso nas narrativas das histórias de vida dos atletas. E, sobretudo, analisar possibilidade de estudar métodos de capturar traços de personalidade de maneira automatizada, porém, com acurácia quanto ao seu objeto de mensuração, os traços de personalidade segundo o modelo dos Cinco Grandes Fatores de Personalidade (Rabelo & Rubio, 2018).

Estes são estudos preliminares no sentido de começar a organizar e, se necessário, desenvolver um corpus para contribuir na mineração de textos em entrevistas com os atletas, para identificação dos cinco grandes fatores de personalidade. De maneira que, a todo momento, sugere-se que sejam continuadas as pesquisas, assim como realizados novos estudos, com o objetivo de desenvolvimento de métodos de extração e organização de dados quantitativo, a partir de medidas qualitativas, tais como as entrevistas.

Nesse sentido, acredita-se que os atletas precisam de ser investigados por meio de ferramentas construídas dentro do seu próprio contexto, pois, o uso de testes psicológicos de personalidade construídos fora do contexto esportivo, para populações gerais, quando aplicados em atletas apresentam resultados menos consistentes. Isso pode estar relacionado ao fato de que o atleta está exposto a situações diferentes da população normativa do teste de personalidade. Um exemplo refere-se a forma como é tratado ansiedade, a percepção da dor, o esforço etc., no contexto esportivo, pois muitas vezes os atletas estão expostos a situações que, na forma como é perguntado na escala, podem gerar respostas dúbias, entendimentos duplicados, sobre o mesmo fenômeno podendo o atleta responder com base em um aspecto ou em outro, diminuindo a confiabilidade dos resultados oriundos da escala, assim como, nestes

mesmos estudos, podem ser observadas variações na estrutura fatorial da escala, em grupos específicos, como no caso de atletas (Rabelo, 2013).

Portanto, em linha com Williams (1991), o autor traz questionamentos a respeito da personalidade e perfil psicológico de atletas de alto rendimento, identificando três fontes de ajuda para que se possa identificar as características psicológicas que se encontram na base das atividades esportivas, entre elas, as informações provenientes dos próprios atletas, por meio das percepções subjetivas experimentadas durante os momentos de alto desempenho de suas carreiras, a outra característica relacionada às informações geradas por estudos que comparam as características psicológicas de atletas exitosos com outros menos prósperos e, por fim, os dados recebidos pelas pessoas que prepararam os atletas mais talentosos de alto desempenho, como treinadores. Em especial para a análise e compreensão da percepção por parte do próprio esportista, tem-se os testes psicológicos como uma possibilidade relevante e de cunho científico para complementar tais investigações.

Considerações finais

Levantar informações em relação às pessoas com foco em investigações apuradas e embasadas empiricamente poderão contribuir para uma compreensão pessoal e profissional do atleta, adequando-se metodologias de treinamento e suporte às especificidades de cada modalidade e acima de tudo de cada esportista, seja do ponto de vista fisiológico, cognitivo, assim como também das emoções. Afinal de contas, considera-se que estímulos e condições ambientais provocam respostas físicas, psicológicas e sociais em distintas circunstâncias e situações nas quais os atletas e esportistas estão envolvidos. As exigências são respondidas com base nos esforços empreendidos pelas pessoas, conforme sua capacidade, habilidade, competência e funcionalidade (Rubio; Rabelo; & Cruz, 2018). Portanto, não é um ponto de vista dos autores deste artigo, considerar o resultado aqui apresentado como estanque, como algo estático, mas apenas, com base nas construções discursivas dos atletas entrevistados, contribuir para um levantamento de aspectos relacionados aos traços de personalidade e emoções, entendendo que, e exclusivamente, o atleta é o único a basilar e poder compreender o que realmente responde, para si, o papel das suas emoções em sua vida, refletida também em sua prática esportiva, em sua profissão.

No Brasil, ainda são raras as publicações de pesquisas que investigam a personalidade do atleta em seu aspecto global, ainda mais considerando extração de conhecimento com uso de tecnologias, mas encontramos uma variedade de trabalhos que buscam estudar individualmente alguns componentes da personalidade do indivíduo, como ansiedade traço-estado, motivação intrínseca, agressividade, determinação, persistência, liderança, extroversão-introversão. No estudo apresentado neste artigo, foi realizada uma análise exploratória do texto de uma narrativa de história de vida de um atleta olímpico visando a identificação de traços de personalidade. Esse processo foi apoiado por técnicas de mineração de textos, e por meio do método proposto e aplicado verificou-se também o potencial do uso de técnicas de mineração de textos na busca de indicativos de personalidade a partir de narrativas de história de vida.

Com este estudo piloto, verificou-se a relação entre adjetivos utilizados na narrativa com traços de personalidade e este trabalho servirá de base para o desenvolvimento de métodos, baseados em mineração de textos, para capturar traços de personalidade segundo o modelo dos Cinco Grandes Fatores de Personalidade. Vale ressaltar que, alguém que se propõe a estudar a personalidade de um indivíduo que pratica esporte ou atividade física tem de estar ciente de que o tema personalidade é amplo, que abrange muitos elementos cujas investigações ainda não foram esgotadas e qualquer tipo de avaliação realizada não corresponde a totalidade, mas sim, a um levantamento de algumas características e tendências que podem ser relevantes ou não no ambiente esportivo (Rubio, Rabelo, & Cruz, 2018).

Ao final, vale reforçar que, para avaliar e sobretudo interpretar resultados associados aos aspectos psicológicos no contexto do esporte, é imprescindível ter uma ampla noção das questões relacionadas ao universo do esportista, desde noções de fisiologia, biomecânica, características específicas das modalidades, regras, questões socioculturais, a psicodinâmica de grupos esportivos entre outros aspectos.

Por fim, destaca-se que a discussão qualitativa versus quantitativa já alimentou longos debates, sendo necessário o esclarecimento de que diferentemente de buscar uma saída conciliatória, devemos ter a compreensão de que não há quantificação sem qualificação, assim como, não há análise estatística sem interpretação, e que diante da complexidade dos fenômenos estudados é cada vez mais desejável e necessário o pluralismo metodológico para que nosso objetivo de investigação seja atingido.

Referências

- Aggarwal, C. C.; & Zhai, C. (2012). *Mining Text Data*. Springer.
- Costa, P. T. & McCrae, R. R. (1985). *The NEO Personality Inventory manual*. Odessa, FL: Psychological Assessment Resources.
- Digman, J. M. (1990). Personality structure: Emergence of the Five-Factor Model. *Annual Review of Psychology*, 41, 417-440.
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI magazine*, 17(3), 37.
- Feinerer, I.; Hornik, K.; & Meyer, D. (2008). Text mining infrastructure in R. *Journal of Statistical Software*, 25(5):1-54.
- Ferreira Junior, N. S. (2013). *Transição de carreira em tempos de amadorismos na história de vida dos pós-atletas da seleção brasileira bicampeã mundial de basquetebol. Qualificação Mestrado*. Escola de Educação Física e Esporte da Universidade de São Paulo, 2013.
- Flusser, V. (2007). *O mundo codificado*. São Paulo: Cosac Naif.
- Fonseca, E. R.; & Rosa, J. L. G. (2013). Mac-morpho revisited: Towards robust part-of-speech tagging. In: *STIL 2013: Proceedings of the 9th Brazilian Symposium in Information and Human Language Technology*, p. 98-107.
- Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, 4, 26-42.
- Hauck, N., Machado, W. L., Teixeira, M. A., & Bandeira, D. R. (2012). Evidências de validade de marcadores reduzidos para a avaliação da

personalidade no modelo dos Cinco Grandes Fatores. *Psicologia: Teoria e Pesquisa*, 28(4), 69-76.

Hutz, C. S., Nunes, C. H., Silveira, A. D., Serra, J., Anton, M., & Wiecek, L. S. (1998). O desenvolvimento de marcadores para a avaliação da personalidade no modelo dos cinco grandes fatores. *Psicologia: Reflexão e Crítica*, 11(2), 395-411.

Just, C. (2011). A review of literature on the general factor of personality. *Personality and Individual Differences*, 50(6), 765-771. doi: 10.1016/j.paid.2011.01.008

King, J. E., Weiss, A., & Sisco, M. M. (2008). Aping humans: age and sex effects in chimpanzee (*Pan troglodytes*) and human (*Homo sapiens*) personality. *Journal of Comparative Psychology*, 122(4), 418-427. doi: 10.1037/a0013125

Machado, W. D. L., Hauck Filho, N., Teixeira, M. A. P., & Bandeira, D. R. (2014). Análise de teoria de resposta ao item de marcadores reduzidos da personalidade. *Psico. Porto Alegre*, 45(4), 551-558. DOI 10.15448/1980-8623.2014.4.13138

Marcacini, R. M., Rossi, R. G., Matsuno, I. P., & Rezende, S. O. (2018). Cross-domain aspect extraction for sentiment analysis: A transductive learning approach. *Decision Support Systems*, 114, 70-80. <https://doi.org/10.1016/j.dss.2018.08.009>

McCrae, R. R. & John, O. P. (1992). An introduction to the Five-Factor Model and its applications. *Journal of Personality*, 60(2), 175-215.

Nunes, C. H. S. S.; Hutz, C. S.; & Nunes, M. F. O. (2010). *Bateria Fatorial de Personalidade (BFP): manual técnico*. São Paulo: Casa do Psicólogo.

Pietroforte, A. V. S. (2010). Semântica lexical. In: *Introdução à Lingüística II: Princípios de Análise*. Editora Contexto, p. 111-135. Rabelo, I. S.; Rubio, K.; Gonçalves, G. de C. M.; & Silva, P.V.C. (2015). Monitoring of Personality Traits among Candidates of an Athletics Program. *International Journal of Applied Psychology*, 5(5), 119-125.

Rabelo, I.S. (2013). *Investigação de traços de personalidade em atletas brasileiros: análise da adequação de uma ferramenta de avaliação psicológica*. Tese de Doutorado. Universidade de São Paulo.

Rabelo, I.S., Rubio, K (2018). Literatura científica sobre a mineração de textos aplicada à identificação da personalidade de atletas. *Olimpianos – Journal of Olympic Studies*, 2(1), 274-303. DOI: 10.30937/2526-6314.v2n1.id37

Rezende, S. O. (2003). *Sistemas inteligentes: fundamentos e aplicações*. Editora Manole.

Rosenwein, B. H. (2011). *História das emoções*. São Paulo: Letra e Voz.

Rubio, K. (2014). Memórias e narrativas biográficas de atletas olímpicos brasileiros. In K. Rubio (org.). *Preservação da memória: A responsabilidade social dos Jogos Olímpicos*. São Paulo: Laços.

Rubio, K. (2007). A avaliação em Psicologia do Esporte e a busca de indicadores de rendimento. In Angelo LF, Rubio K. (org.). *Instrumentos de*

avaliação em psicologia do esporte. São Paulo: Casa do Psicólogo.

Rubio, K. (2001). *O Atleta e o Mito do Herói*. São Paulo: Casa do Psicólogo.

Rubio, K. (2013). Multiple Identities and Trans-culturalism: Joaquim Cruz, a Brazilian Olympic Hero. *American International Journal of Contemporary Research (Print)*, 3, 42-51.

Rubio, K.; Rabelo, I. S.; & Cruz, R. M. (2018). Avaliação de aspectos psicológicos em Educação Física e Esporte. In: Maria Tereza Silveira Böhme (org). *Avaliação e Desempenho em Educação Física e Esporte*. São Paulo: Editora Manole.

Sinoara, R. A., Camacho-Collados, J., Rossi, R. G., Navigli, R., & Rezende, S. O. (2019). Knowledge-enhanced document embeddings for text classification. *Knowledge-Based Systems*, vol 163, 955-971. <https://doi.org/10.1016/j.knosys.2018.10.026>

Sinoara, R. A., Antunes, J., & Rezende, S. O. (2017). Text mining and semantics: a systematic mapping study. *Journal of the Brazilian Computer Society*, 23:9. <https://doi.org/10.1186/s13173-017-0058-7>

Sinoara, R. A., Scheicher, R. B., & Rezende, S. O. (2017). Evaluation of latent dirichlet allocation for document organization in different levels of semantic complexity. In: *Proceedings of 2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, (p.1-8). IEEE. <https://doi.org/10.1109/SSCI.2017.8280939>

Weinberg, R., & Gould, D. (2001). *Fundamentos de Psicologia do Esporte*. Porto Alegre: ArtMed.

Williams, J. M. (1991). *Psicologia aplicada al deporte*. Madrid: Biblioteca Nueva.

Sobre o autor

Katia Rubio

Universidade de São Paulo, Escola de Educação Física e Esporte da USP, (EEFE/USP), SP, Brasil

Ivan Santana Rabelo

Universidade de São Paulo, USP, SP, Brasil.

Roberta Akemi Sinoara

Universidade de São Paulo, USP, SP, Brasil.

Renan Santos Barbosa

Universidade de São Paulo, USP, SP, Brasil.

Solange Oliveira Rezende

Universidade de São Paulo, USP, SP, Brasil.

Contato

ENDEREÇO PARA CORRESPONDÊNCIA

Katia Rubio

Universidade de São Paulo, Escola de Educação Física e Esporte da USP, Departamento de Pedagogia do Movimento do Corpo Humano. - Av. Prof. Mello Moraes, 65 – Butantã – CEP: 05508900 - São Paulo, SP - Brasil

E-MAIL

katrubio@usp.br

TELEFONE

+ 55 (11) 39013135

Fax: (11) 39312808

Sobre o trabalho

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) Código de Financiamento 001. Com apoio da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP - 2009/54.846-2), submetido CEP, sob protocolo nº. 30756114.6.0000.5391, e também apoiado por meio do Programa de Incentivo à Atração de Pós-doutorandos (PIAPD) na Escola de Educação Física e Esporte (EEFE/USP) sob Proc. USP nº. 16.1.1201.1.5.