

Apreensibilidade e Qualidade da Informação: Bases de uma Avaliação Textual Automática na Área da Saúde

Asdrubal Falavigna, Carine G. Webber, Fernando A. Abel, Marco A. Koff, Maurício K. Santos, Natália Lisboa

Universidade de Caxias do Sul
Rua Francisco Getúlio Vargas, 1130 - CEP 95070-560
Caxias do Sul (RS) - Brasil

cgwebber@ucs.br

Abstract. *The internet is usually employed to research health-related information. However, available texts lack proper evaluation. In this context, a systematic review of literature on the online evaluation of texts on spinal pathologies has been performed. Selected articles evaluate the quality and the readability of the texts by means of metrics widely accepted in the scientific community, which serve as a basic source to develop a hybrid software for automatic text analysis, based on Artificial Intelligence techniques.*

Resumo. *A Internet é comumente utilizada para pesquisa de informações na área da saúde. Contudo, os textos disponíveis carecem de devidas avaliações. Neste escopo, realizou-se uma revisão sistemática da literatura referente à formas de avaliação de textos sobre patologias da coluna vertebral. Selecionou-se artigos que avaliam a qualidade e a apreensibilidade (ou readability) dos textos por meio de métricas difundidas na comunidade científica para construção de um software para automação da análise textual, em uma abordagem híbrida que empregue técnicas de Inteligência Artificial.*

1. Introdução

Variadas fontes de informação são observadas com a expansão da Internet, nem sempre abordando um tema de forma clara e correta. Em certas áreas, sobretudo naquelas relacionadas à Saúde, é muito importante que existam procedimentos para avaliar informações textuais com relação a critérios como a apreensibilidade (ou *readability*) e a qualidade da informação.

A apreensibilidade da informação textual foi definida por Klare (1963) como sendo a facilidade do entendimento ou compreensão de um texto a partir do estilo da escrita utilizada. Desde então, diversos métodos foram criados a fim de avaliar a apreensibilidade de um texto. Tais métodos vêm sendo, em grande parte, desenvolvidos e aplicados para textos em língua inglesa. A avaliação da apreensibilidade corresponde a uma análise quantitativa da informação (contagem de frases, palavras e sílabas de um texto conforme regras gramaticais), que pode produzir resultados numéricos. É escasso o estudo da aplicabilidade de métodos quantitativos para a língua portuguesa.

A qualidade pode também ser verificada por meio de técnicas de Inteligência Artificial para análise textual (aprendizagem automática e processamento de linguagem natural), extraíndo-se padrões, agrupando e classificando textos [Luger, 2009]. As análises textuais qualitativas em domínios restritos produzem resultados coerentes e úteis [Weiss, 2010; Feldman e Sanger, 2006].

O acesso facilitado de pessoas e pacientes a fontes de informação sobre saúde ampliou a necessidade de que tais informações disponíveis sejam revisadas e analisadas, principalmente em contextos onde o paciente deve participar da decisão do seu tratamento. Nesta problemática, já foi observado em uma amostra populacional que as informações disponíveis não foram compreensíveis a usuários não especialistas, ocasionando indefinições e insegurança no andamento dos respectivos tratamentos. O objetivo deste artigo é promover uma revisão sistemática sobre os métodos e modelos de avaliação de conteúdos da internet sobre doenças da coluna vertebral, servindo como base para o desenvolvimento de um software de análise textual.

2. Materiais e Métodos

2.1 Fase 1: Revisão Sistemática como base científica para desenvolvimento de inovação

Foi realizada uma revisão sistemática da literatura com o objetivo de avaliar a qualidade e apreensibilidade das informações disponíveis na internet sobre doenças da coluna vertebral. A metodologia utilizada foi de pesquisa no banco de dados do PubMed utilizando a seguinte lista de palavras-chave: “education”, “readability”, “evaluation”, “back pain”, “spinal diseases” e “internet”.

Os critérios de seleção dos artigos foram: (1) estudos que continham avaliação da informação em sites; (2) avaliação de sites com doenças da coluna vertebral; (3) publicações em inglês; (4) publicação no período compreendido entre Janeiro de 2000 e Setembro de 2015. Os critérios de exclusão foram: (1) avaliação de um único site; (2) sites com informação contida em vídeos e (3) sites não disponíveis ao paciente.

2.2 Fase 2: Desenvolvimento do Software para Avaliação Textual Automática na Área da Saúde

Para solucionar o problema da avaliação textual na área de saúde na Língua Portuguesa propõe-se o desenvolvimento de um software que empregue técnicas quantitativas e qualitativas, integrando técnicas de Inteligência Artificial.

3. Resultados

3.1 Fase 1: Revisão Sistemática como base científica para desenvolvimento de inovação

O término do levantamento bibliográfico foi na data de 05/10/2015. Foram selecionados 167 artigos, dos quais 104 não estavam relacionados a patologias, sintomas ou cirurgias de coluna. Os resumos foram analisados em 63 artigos, sendo a avaliação do texto completo realizada em 24 artigos. Os 23 artigos foram incluídos no estudo.

As principais ferramentas utilizadas para avaliação da qualidade da informação e apreensibilidade dos sites foram o Flesch-Kincaid Grade Level, o DISCERN tool e o Flesch Reading Ease (Tabela 1).

Tabela 1. Ferramentas Utilizadas na Análise da Informação

Modalidade de Avaliação	Ferramenta de Análise	Número de Artigos Presentes
Avaliação da Qualidade	DISCERN tool/instrument + 3 revisores	5
	JAMA Benchmark Criteria	3
Avaliação da Apreensibilidade	Flesch Reading Ease	4
	Flesch-Kincaid Grade Level	7
	SMOG Grading	2
	Coleman-Liau Index	1
	FORCAST Formula	1
	New Dale-Chall Readability Formula	1
	Gunning Fog Index	3
	New Fog Count	1
	Fry Graphical Analysis	1
	Raygor Readability Estimate	1

Os doze diferentes programas de análise de textos utilizados na avaliação da informação dos sites são em Língua Inglesa. Posteriormente foi verificada a possibilidade de utilizar os programas de análise de texto da Língua Inglesa para aplicação ou adaptação em Língua Portuguesa e de processamento por um computador (Tabela 2).

Tabela 2. Estudo das ferramentas encontradas

Ferramenta de Análise	Aplicação para o português?	Computável?	Considerações
DISCERN Instrument	Sim	Não	Mecanismo realizado manualmente por avaliadores, utilizando-se 16 perguntas sobre o tema.
JAMA Benchmark Criteria	Sim	Não	Sem considerações.
Flesch Reading Ease	Sim	Sim	Métrica é calculada usando o número médio de sílabas por palavra e comprimento médio da frase. Já existe uma adaptação para a língua espanhola.
Flesch Kincaid Grade Level	Sim	Sim	Semelhante ao Flesch Reading Ease. Utilizado para medir o nível de escolaridade de pessoas nos EUA.
SMOG Grading	Não	Sim	Requer um mínimo de 30 frases para aplicação, menos que isso resulta em resultados estatisticamente inválidos. Existem avaliações de textos farmacêuticos de Portugal, no entanto os textos tiveram de ser traduzidos em inglês.
Coleman-Liau Index	Sim	Sim	Métrica calcula o nível de ensino baseado apenas nas médias dos comprimentos das sentenças e a média do número de caracteres por palavra.
FORCAST formula	Sim	Sim	Métrica ideal para manuais técnicos e formulários (textos sem narrativas).A fórmula não utiliza o tamanho das frases, portanto pode ser utilizada em textos sem frases.
New Dale-Chall Readability Formula	Não	Sim	Diferente de outras métricas, essa fórmula utiliza uma lista de 3000 “hard words” para cálculo do grau de escolaridade.
Gunning Fog Index	Sim	Sim	Semelhante ao Flesch Reading Ease.
New Fog Count	Sim	Sim	Métrica ideal para documentos técnicos e manuais. Desenvolvida pela Marinha dos EUA.
Fry Graphical Analysis	Talvez	Sim	Método inicialmente desenvolvido para o Inglês.
Raygor Readability Estimate	Não	Sim	Métrica de mensuração do grau de escolaridade de textos em Inglês. Necessita grande adaptação para o Português.

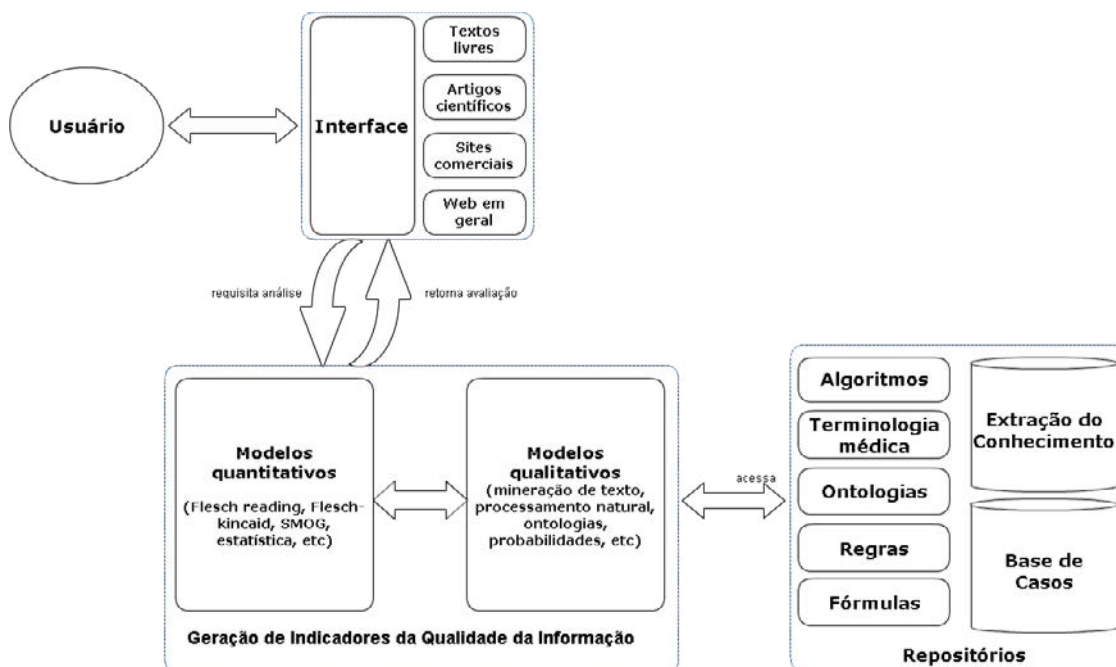
3.2 Fase 2: Desenvolvimento do Software para Avaliação Textual Automática na Área da Saúde

A concepção e modelagem do software encontra-se em desenvolvimento. O software apresentará uma interface visual e textual, com possibilidade de navegação na web para seleção e avaliação de textos (figura 1).

A análise quantitativa será realizada utilizando as ferramentas elencadas na tabela 1. O processamento qualitativo será gerado através de algoritmos de aprendizagem automática e processamento da linguagem natural (máquinas de suporte vetorial, aprendizagem simbólica e análise de latência semântica). O software deve ainda contemplar uma camada de persistência para armazenar dados coletados, calculados, resultados de análises, bem como termos da área da saúde, vocabulários, stop words e bases de casos. O resultado das análises será apresentado na forma de

interfaces gráficas com recursos de visualização que permitam inferir sobre a qualidade da informação e apreensibilidade identificadas automaticamente.

Figura 1 - Arquitetura de software de análise de apreensibilidade da informação



4. Contextualização e Perspectivas Futuras

A Internet, apesar de ser uma fonte útil e relevante na busca por conhecimento, deve oferecer mecanismos objetivos de verificação, usando métricas de apreensibilidade e qualidade. A fim de construir uma fundamentação sólida para a avaliação textual optou-se por uma abordagem híbrida (quantitativa e qualitativa). Em termos quantitativos pretende-se avaliar características de apreensibilidade. Utilizando técnicas oriundas da Inteligência Artificial busca-se implementar mecanismos que permitam avaliar a qualidade da informação. Tais técnicas se complementam, trazendo ao usuário dados que lhe permitam julgar um conteúdo textual a partir de métricas bem definidas. As próximas etapas deste trabalho envolvem a implementação do software, testes e análise dos resultados por meio da comparação da avaliação realizada por especialistas humanos em uma amostra de documentos extraídos da Internet. Espera-se desta forma contribuir com um mecanismo de avaliação correto e preciso dos sites de conteúdos de doenças da coluna vertebral, permitindo que informações claras e verdadeiras sejam oferecidas aos usuários.

5. Referências

- Feldman, R.; Sanger, J. The Text Mining Handbook Advanced Approaches in Analyzing Unstructured Data. New York: Cambridge University Press, 2007.
- Klare, G. R.: The measurement of readability. Iowa: Iowa State University Press,(1963)
- Luger, George.F. Artificial Intelligence. Person Education, 2009,774 p.
- Weiss, Sholom M., Indurkha, Nitin, Zhang, Tong. Fundamentals of Predictive Text Mining, Springer-Verlag London, 2010.