

心身ストレスの可視化のためのストレス・エビデンス検索エンジンの提案

A Proposal of Stress Evidence Search Engine for Visualization of Human Stress

中山 友紀(岩手大), 山口 昌樹(岩手大学大学院)

Yuki NAKAYAMA, Iwate University

Masaki YAMAGUTHI, Graduate school of Engineering, Iwate University

Abstract: The purpose of this study was to develop a "diagnostic algorithm for human stress" to connect between web information and medical information. Stress Evidence Search Engine foresaw a change and judged it and could suggest a recovery support technology and the evaluation standard to connect it with appropriate coping, recovery, and it could expect application to health care business thereby.

Key Words: Stress evidence search engine, Evidence, Recovery support technology

1. 緒言

ストレスのように病気ではなく、心身の状態に影響を及ぼすクライシス(危機)を予測するには、モニタリングが重要な役割を果たす。現在、多様なセンサを活用した生体情報を可視化する技術があるが、単方向に偏り、一望できる可視化技術が必要である¹⁾。現在社会は、デジタルデータが爆発的に増大するビックデータ(情報爆発)時代を迎えたと言われているが、不安や疑問を解決するための、データベースは構築されていない。そこで、疾患・心身状態との間を取り持つ情報の整理・分析を「意図的に」行うために、非医学的(Web, つぶやき)情報と医学的(学術論文)情報の統合を実現し、非医学的情報の中で発生したイベントと医学的情報(エビデンス度)の関連性を解釈する、「ストレス・エビデンス・検索エンジン」を開発し、診断アルゴリズムの構築を目的とする。

2. 方法

2.1 情報

本情報を搭載する環境として、Unix系OS Linux(2.6.32 x86_64; CentOS 6.4)データベースkey-value型を用いた。医学的情報は、論文検索サイトScopusで、ストレス×コルチゾール×唾液×(鬱病+大鬱病+双極性+不安症+心的外傷+過敏性腸症候群+統合失調症)で検索した2,000件のアブストラクトと、非医学的情報として、Webサイトから記事が独立している、はてなブックマーク、はてなキーワード(525.5KB)を情報基盤とする「ストレス・エビデンス・検索エンジン」

を試作した。

2.2 エビデンス度

TF-IDF²⁾は、文書中に出現する頻度と重要度を示している。形態素解析(MeCab)を用いて、Web情報と学術論文のキーワード抽出を行い、抽出したキーワードをTF-IDF計算(東大中川研究室開発)により、頻出度及び重要度計算をした。(1)TF(単語の出現頻度)とIDF(逆数頻度)の2つの指標にもとづいて計算されている。(2)ここで、 $n_{i,j}$ はある単語*i*の文書*j*における出現回数、 $\sum_k n_{k,j}$ は文書*j*内における全ての単語の出現回数の和。(3)|*D*|は総文書数、 $|\{d:d \in t_i\}|$ は任意の単語*i*を含む文書数を表している。

$$tfidf = tf \cdot idf \quad (1)$$

$$tf_{i,j} = \frac{n_{i,j}}{\sum_k n_{k,j}} \quad (2)$$

$$idf_i = \log \frac{|D|}{|\{d:d \in t_i\}|} \quad (3)$$

Web情報と学術論文の同一キーワードのスコア(TF-IDF値)を乗算して、加算した値(ベクトル内積)を計算し、算出することで示しており、Tiny SVMはキーワードごとの距離の測定を行い、エビデンス度を出している。

3. 結果と考察

3.1 重要頻出キーワードの抽出

論文検索サイトScopusで検索した2,000件のアブストラクトの医学的情報と、はてなブックマーク、はてなキーワード

などの Web コンテンツ等の非医学的情報を TF-IDF により数値化し、また、双方の関連性を、SVM を用い可視化した結果、Fig.1 に示す、エビデンス度を確認できた。これにより、ストレスに関連するキーワードとして、「人・自分・社会」が抽出され、それらをコアキーワードとしたエビデンス度を示す事が可能となったことから、ストレスの原因検索に寄与できると考えられる。

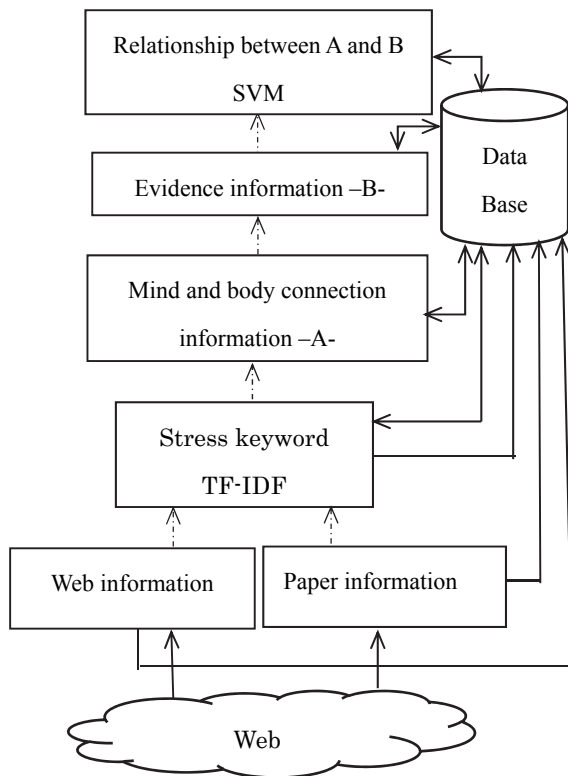


Fig.1 Evidence algorithm

3.2 診断アルゴリズム

個人ユーザーインターフェースに、ストレス情報と論文情報を画面で表す機能を、Fig.2 の「KOKORONOMA」で表示させることで、個人ユーザーが一人で判断・診断できるアルゴリズムを構築することができた。このことにより、「未病」の時からいつでもどこでも、ユーザー一人で、インターネット環境下であれば使用可能であり、ストレス関連疾患の発症前診断として有用であると考えられ、ストレス回復支援に繋がる可能性を示唆する事ができた。

4. 結語

慢性疾患は、ひとたび発生すれば、完全な治療が困難となる病気が多い。早期診断は、その予防や重症化への防止が重要である。そのため、発症→早期診断→治療→回復に至る 4 つの過程を断ち切らず、ポイント・オブ・ケアとして、アナリティック・データベースによる新しい手法として、本シス

テムは発揮できると考える。

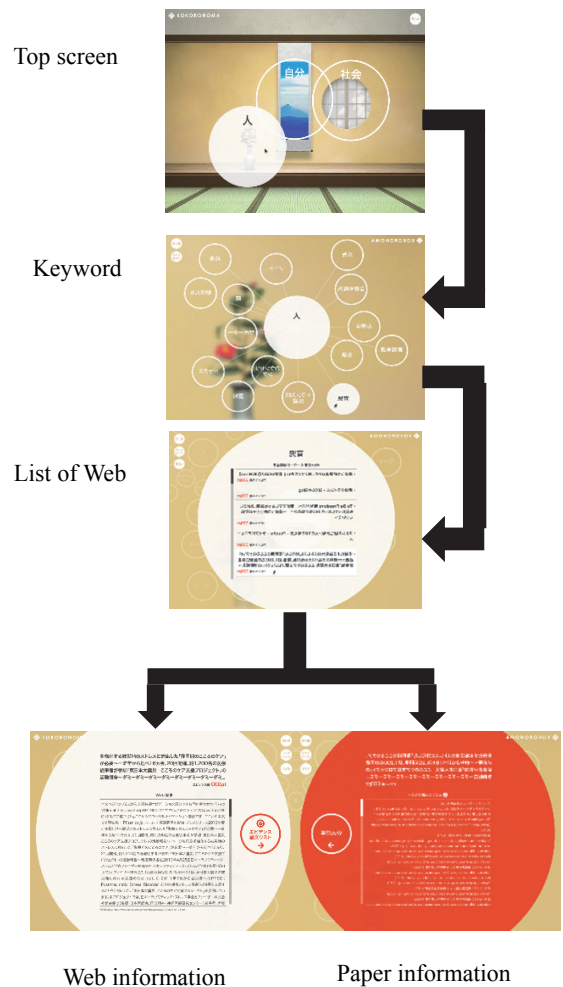


Fig. 2 diagnosis algorithm

参考文献

- 1) 六反 一仁, 加藤 宏一, 奈良原 正俊, 富田 裕行, 齋藤 俊郎, 杉山 寿: ストレス評価用 DNA チップを用いたメンタルジェネティクスの展開, Bio Industry Organization, **19**, 2002, 19-24
- 2) 佐藤 翔輔, 林 春雄: TFIDF/TF 指標を用いた危機管理分野における言語資料体からのキーワード自動検出手法の開発, 地域安全学会, **8**, 2006, 367-376