

理学療法士の新人教育のための電子カルテ評価支援システム

An electronic medical record evaluation support system for the education of new physical therapists

鍛冶宏旭^{1*} 砂山渡¹ 堀寛史²
Hiroaki Kaji¹ Wataru Sunayama¹ Hirofumi Hori²

¹ 滋賀県立大学工学部

¹ School of Engineering, The University of Shiga Prefecture

² びわこリハビリテーション専門職大学

² Biwako Professional University of Rehabilitation

Abstract: In recent years, the number of young physiotherapists has increased rapidly. As a result, the burden of new employee education has increased, the quality of education for each student has declined, and the overall quality of physiotherapists has declined. Therefore, in this study, we support educators' scoring by performing automatic evaluation and text highlighting in electronic medical records, which is one of the flows of physical therapy. This will lead to an improvement in the quality of scoring by educators and a reduction in the burden.

1 諸言

近年、理学療法士養成施設や理学療法士が急激に増加している傾向がある [1]。また、理学療法士について、日本理学療法士協会が公開する理学療法士の年齢分布と平均年齢 [2] によると 12 万人中、20 代がおよそ 5 万人であり、他の年代と比べると最も多く、若い理学療法士が増えていることがわかる。この影響もあり、理学療法士の新人教育への負荷の増加や経験の浅い理学療法士が新人教育を担当することが問題点としてあげられている。これにより、理学療法士の教育不足による全体的な質の低下を招き、患者の退院時の生活の質 (QOL) が望むように向上しない結果となっている [3]。

こういった現状から理学療法士の新人教育に対する負荷の軽減が重要であると考え、負荷の軽減として教育方法の効率化や新人のフィードバックを簡単にする事で時間の短縮につながる。時間の余裕が生まれれば新人一人一人に的確な教育ができ新人の成長につながると思う。

そこで、本研究では理学療法の流れの一つである電子カルテの評価負荷の軽減と評価の質の向上を目的として行う。評価負荷の軽減として電子カルテの自動評価により評価の時間短縮を行う。評価の質の向上として電子カルテを項目に分けそれぞれの重要な単語をハイライトすることで細かく評価を行い質の向上を行う。

本研究は電子カルテを自動評価する処理と重要な単語をハイライトする処理を組み込んだシステムを作成した。本システムを用いて評価者を支援し、評価負荷の軽減と評価の質を向上させることを目的とする。

本論文は、2 章では関連研究、3 章では理学療法計画における電子カルテの構成、4 章では電子カルテ評価支援システムの構成、5 章では電子カルテ評価支援システムの精度と評価の検証結果を述べ、6 章で本論文を締めくくる。

2 関連研究

レポートの評価基準の不安定さを抑えるツールを作成、評価を行う研究がある [4]。これはレポートの評価時に前の内容の質によって影響し、評価が不安定になる点に着目したものである。この研究では類似したレポートをまとめ評価することで評価の不安定さを抑え、評価者の負荷を軽減している。本研究と同じで評価者に支援を行い負荷の軽減を行っている。しかし、本研究では自動評価や単語のハイライトを行い評価のしやすさを重視している点で違いがある。

国語の記述式の答えを 4 項目に分けて自動評価を行う研究がある [5]。評価項目ごとのスコアを出力する問題を、深層学習を使った自動評価のモデル [6] を拡張することで項目ごとの点数を予測している。また、項目を分けることにより学習者は項目の点から改善する箇所を知ることができ、フィードバックを可能とした。本

*連絡先： 滋賀県立大学 工学部 電子システム工学科 鍛冶宏旭
〒 522-8533 滋賀県彦根市八坂町 2500
E-mail: ov23hkaji@ec.usp.ac.jp

表 1: 症例データを構成する情報

症例データの情報	その情報に含まれる内容
一般情報	氏名, 年齢, 職業, 診断名など, 基本的な情報が含まれている
医学的情報	血液検査や血圧検査などの医学的検査の結果が含まれている
他部門からの情報	医師 (Dr), 言語聴覚士 (OT), 看護師 (Nr), 作業療法士 (St) によるコメント情報が含まれている
理学的検査	運動機能障害の度合いや, 特定の動作に対する補助の必要性の有無などの情報が含まれている
問題点の列挙	この症例における問題点の候補が含まれている

研究も電子カルテを 4 項目に分け自動評価を行い, 教育者の評価の支援や学生のフィードバックに用いる. この研究では自動評価のデータとしておよそ 5000 枚の解答を使用している. 本研究では様々な問題を自動評価できるようにするため少ない解答データで自動評価を行う.

英語エッセイの自動レベル判定システムと手動評価結果を比較を行う研究がある [7]. この研究ではまず, 3 つの仮説を立てる. その仮説をもとに手動評価と自動評価で別々に評価を行いそれぞれの課題を考察していた. 本研究と同じで自動評価と手動評価を用いて研究を行っているが, 自動評価の結果をもとに手動で評価を行い手動評価の支援をする違いがある.

3 理学療法計画における電子カルテの構成

本章では症例データの構成, 電子カルテの構成と, 電子カルテの点数評価における症例データと電子カルテの関係について述べる. 本論文では, 症例データとは様々な医学分野での患者の診断結果や状態が記載されているデータのこと, 文献とは各症例データごとに問題のある特定の身体部位や動作に対する治療方法や対処方法が含まれている情報群とする. 症例データ, 文献と電子カルテとの関係として, 症例データと文献を基に電子カルテを作成する. 電子カルテによって今後の治療計画を定め, 治療を行った結果を症例データに更新する. これを繰り返すことにより患者の症状を改善させる

3.1 症例データの構成

症例データは表 1 の 5 つの情報群で構成されている. また, 今回用いた症例データは, ペーパーペイシェントと呼ばれる完全な架空のデータである.

表 2: 電子カルテの項目単語の意味

単語	意味
心身機能	身体系の生理的機能 (心理的機能を含む) [9]
活動	課題や行為の個人による遂行のこと [9]
統合	症例データに含まれる情報から該当する問題点を特定すること
解釈	特定された問題点から, 今後の治療計画の策定につながる考察を行うこと

3.2 電子カルテの構成

電子カルテの文は心身機能の統合, 心身機能の解釈, 活動の統合, 活動の解釈の 4 つの項目とどの項目に属さない文に分けられる. 点数を評価する際この 4 つの項目で評価を行い, 総合点を算出する. 表 2 に心身機能, 活動, 統合, 解釈の意味を示す. 各項目に分類される文の特徴を述べる.

3.2.1 心身機能の統合

症例データの【問題点の列挙】の項目に含まれる問題点の候補のうち, 心身機能に異常があると推測される問題点の選択とその根拠について【一般情報】、【医学的情報】、【他部門からの情報】、【理学的検査 (理学療法評価)】を基に述べている文がこの項目に分類される

3.2.2 心身機能の解釈

心身機能に関する問題点を解消するための治療計画を, 文献を参考に提案されている文がこの項目に分類される. このとき, 心身機能の統合に分類される文に含まれる単語とのずれが少ないことが条件となる.

3.2.3 活動の統合

症例データの【問題点の列挙】の項目に含まれる問題点の候補のうち, 活動に異常があると推測される問題点の選択とその根拠について【一般情報】、【医学的情報】、【他部門からの情報】、【理学的検査 (理学療法評価)】を基に述べている文がこの項目に分類される.

3.2.4 活動の解釈

活動に関する問題点を解消するための治療計画を, 文献を参考に提案されている文がこの項目に分類される. このとき, 活動の統合に分類される文に含まれる単語とのずれが少ないことが条件となる.

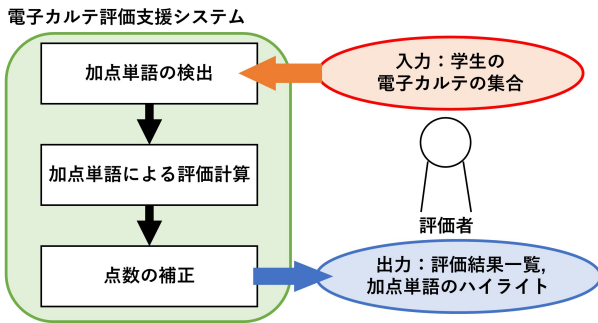


図 1: 電子カルテ評価支援システムのフローチャート

4 電子カルテ評価支援システムの構成

4.1 電子カルテ評価支援システムの構成

電子カルテ評価支援システムは前処理として症例データと文献を用いて作成した専門家の電子カルテと学生の電子カルテ、電子カルテの点数データを用意する。この電子カルテを教師データとし、教師データと点数データを用いて自動評価に使用する計算の重みや単語を設定する。自動評価に用いる単語を加点単語とする。

本システムではこの前処理が必要なため、あらかじめ定めた症例データと文献をもとに作成した教師データを用いて自動評価に用いる計算の重みや単語を設定する必要がある。そのため、新しく症例データを用いて評価支援を行う場合、4.2 加点単語の決定や 4.3 電子カルテの点数評価アルゴリズムから相関係数の高い重みを設定する必要がある。

本研究では、架空の脳卒中の患者の症例データと文献を用いて作成した電子カルテを教師データとして用いる。今回、教師データは専門家が作成した電子カルテと学生が作成した電子カルテ 86 枚の計 87 枚である。この教師データを用いて自動評価に用いる計算の重みや単語を設定する。

本研究における電子カルテ評価支援システムのフローチャートを図 1 に示す。電子カルテ評価支援システムの処理の流れは、まず複数の学生の電子カルテを段落で分割したテキストを入力する。次に学生のカルテを各項目の加点単語で、数と位置を検出する。加点単語の検出した数を用いて項目の点数を算出する。項目の合計点に補正を加える。その計算を段落の数、つまり学生の数行う。最後に学生の点数と加点単語をハイライトしたテキストを出力する。

表 3: 大加点単語・中加点単語の単語

項目	心身機能の統合	心身機能の解釈	活動の統合	活動の解釈
大加点単語	障害, 体, 錐, 路, 点	回復, レベル, 立ち上がる	困難, 移動	レベル, 麻痺, 立
中加点単語	反射, 筋, 運動, 緊張, 深部	考える, 筋, 運動, 歩行, 下肢, 麻痺, 緊張	介助, トランス, フェア, 車椅子, 可能, レベル, 手すり, 動作	自立, トランス, フェア, 歩行, 考える, 車椅子, 動作

4.2 加点単語の検出

加点単語の設定には専門家が作成した電子カルテと学生が作成した電子カルテ 86 枚の計 87 枚の教師データと点数のデータ、電子カルテの文を 4 項目に分類したデータを用いる。ここで 4 項目は電子カルテの構成の「心身機能の統合」「心身機能の解釈」「活動の統合」「活動の解釈」の 4 つである。また、4 項目に含まれない文は加点単語の設定に使用しない。教師データの点数のうち、8 点以上を高得点、7 点以下を低得点とし、高得点に分類された教師データの各項目で高頻出単語を 10 個取得する。低得点に分類された教師データも同じように各項目で高頻出単語を 10 個取得する。

これにより取得した単語のうち、高得点のみに含まれる単語を大加点単語、高得点と低得点どちらにも含まれる単語を中加点単語と設定する。

今回、研究に用いた「脳卒中」の症例に対する電子カルテについては、全部で 28 個（重複を含めると 40 個）の加点単語が設定された。設定した加点単語を表 3 に記す。この加点単語を用いて、入力された学生の電子カルテ集合の各カルテで、各項目の加点単語の数と位置を検出する。

4.3 電子カルテの点数評価アルゴリズム

4.3.1 加点単語による評価

各項目の最高点は、「心身機能の統合」「活動の統合」を 2 点とし、「心身機能の解釈」「活動の解釈」は記述で重要なため 3 点とする。各項目の合計点の最高点を 10 点、最低点を 0 点とする。

項目ごとの点数の計算は式 1 に記す。専門家が作成した電子カルテ文章の中で、項目 C の加点単語の使用回数を W_{TC} 、大加点単語の使用回数を W_{TCL} とする。評価したい電子カルテ文章の中で、項目 C の加点単語の使用回数を W_C 、大加点単語の使用回数を W_{CL} とする。最高点数を M_C とし、重みを 0.5 として各項目の点数を求める。項目の点数を S_C とし、 S_C が項目 C の最高点を上回る場合は $S_C = M_C$ 、 S_C が 0 を下回る場合は $S_C = 0$ とし、点数の範囲外の値となることを

避ける. 式 (1) で求めた各項目の点数 S_C の合計 S_{SUM} を式 (2) で求める.

$$S_C = \frac{W_C + W_{CL} \times 0.5}{W_{TC} + W_{TCL} \times 0.5} \times M_C \quad (1)$$

$$S_{SUM} = \sum S_C \quad (2)$$

4.3.2 点数の補正

点数の補正として文字数と非線形の補正を加える.

文字数の補正を加える理由は単語のみの場合, 説明が不十分でも評価されるため, それを防ぐためである.

非線形の補正としてロジット関数を用いる. ロジット関数を用いる理由は計算した点数が実際より高く, 低くなりやすい傾向があり, それを防ぐためである.

4.3.1 で求めた S_{SUM} に文字数の重みを追加した点数 S_{length} を式 (3) で求める. 文字数の重みを 710, 評価したい電子カルテの文字数を L として計算する. また, S_{length} が最高点の 10 点を上回る場合は $S_{length} = 9.99$, S_{length} が 0 を下回る場合は $S_{length} = 0.01$ とし, 点数の範囲外の値となることを避ける.

$$S_{length} = S_{SUM} \times \frac{L}{710} \quad (3)$$

S_{length} に非線形の補正を追加した点数 S_{logit} を式 (4) で求める. また, S_{logit} が最高点の 10 点を上回る場合は $S_{logit} = 10$, S_{logit} が 0 を下回る場合は $S_{logit} = 0$ とし, 点数の範囲外の値となることを避ける. この S_{logit} が最終的な電子カルテの総合点となります.

$$S_{logit} = \log \left(\frac{S_{length} \times 0.1}{1 - S_{length} \times 0.1} \right) + 5 \quad (4)$$

以上の式の重みは専門家に評価した点数と計算し求めた点数との相関係数が高い重みを設定している. 実際に専門家の点数と計算した点数の相関係数は 0.739 と高い値をとる.

4.4 電子カルテ評価支援システムのインタフェース

本節では, 教育者が学生の電子カルテを評価するためのインタフェースについて述べる. 図 2 の電子カルテの評価支援システムのインタフェースは, 4 つの部分から構成されている.

図 2 より①, ②は学生の電子カルテを加点単語で評価した点数を表示する. ②は名前, 総合点, 各項目の点数, 文字数を表示している. 複数の学生の点数の中から, 一人の学生を選択し, 下のフォーカスボタンを押すことで一人分の点数を表示できる. ①は学生用の



図 2: 電子カルテ評価支援システムのインタフェース

点数を表示する. 本システムでは検討段階だが学生が採点を終えた電子カルテを修正する際の支援を考えている. その際, どのように見えるか評価者が確認するため使用する.

③は電子カルテの加点単語をハイライトし表示している. 大加点単語を橙色, 中加点単語を緑色でハイライトしている. 下の項目のボタンにより各項目に設定されている加点単語をハイライトすることができる. ④は加点単語の単語と使用回数を表示する. ②のフォーカスボタンを使用することで一人の学生の加点単語使用回数を確認できる.

また, 本システムはテキストデータの分析を支援する枠組みを有するため統合環境 TETDM (Total Environment for Text Data Mining)[8] 上で構築した.

5 電子カルテ評価支援システムの効果の検証

5.1 検証方法

評価者に本システムを用いて, 学生が架空の症例データから作成した電子カルテ 86 枚を 10 点満点で採点してもらい評価の質の向上を検証する. この電子カルテはすでに評価者に採点されているが, その結果を参照せずに採点をしてもらった. 今回, 評価者がシステムの点数に引っ張られる可能性を考慮して, システムの評価に加点単語を考慮せず, カルテの長さのみによって評価した結果を用いて検証を行った. この評価結果は加点単語を考慮した評価結果より高い点数が出力されるように調整した. 続いて, 加点単語を考慮したシステムが出力する点数より評価者が 1 点以上高くつけたカルテについて, 再度加点単語を考慮したシステムで採点してもらい, その点数の変化を確認した. ここで加点単語を考慮しないシステムの評価結果を高評価

システムとし、加点単語を考慮したシステムの評価結果を評価システムとする。

検証後、データをもとに評価者にインタビューを行い、電子カルテ評価支援システムによって評価の負荷が軽減したかを評価した。

5.2 検証結果と考察

5.2.1 電子カルテ評価支援システム使用前後の評価結果との相違

図3は評価者が採点した点数と評価者が最初に採点した点数（以下、元の点数）との差を示す。図3より元の点数より評価者の採点に変化していることがわかる。変化が起こる理由として加点単語や文のハイライトの支援により電子カルテの評価の着目ポイントが変わったからだと考える。全体的に点数が上がっていることから高評価システムに点数が誘導されている可能性も考えられる。

図4は評価者が採点した点数と評価システムとの差を示す。図4より評価システムの点数の差が0に近いものが多いことがわかる。評価システムは加点単語を考慮して評価された点数であり、その点数に近いことから加点単語、文のハイライトによる評価の支援は行えていると考える。

図5は評価者が採点した点数と高評価システムとの差を示す。図5より全ての点数の差が負の値をとることがわかる。高評価システムの点数だけに着目するのではなく、加点単語、文のハイライトによって電子カルテの評価を行っていると考える。図4,5より-3、-4と大きな差をとることがわかる。これはシステムが10点と評価した電子カルテで多く、現在の評価方式で評価できていない要素があると考えられる。今後、システムの評価方式に新しい要素を加える必要があると考える。

図6は評価者の2回目の採点結果を含めた点数と評価システムの点数の相関をグラフにしたものである。両者の相関は0.89と高い値となる。相関係数より評価システムと評価者の採点した点数に相関があることから、システムの評価を参考に評価が行えると考える。

5.2.2 電子カルテ評価支援システムの評価

検証後に電子カルテ評価支援システムについて評価者にインタビューを行った。「単語のハイライトや点数を参考にすることでスムーズに採点が行えた」ことから、本システムの点数やハイライトにより評価負荷が軽減されたとわかる。また、「システムで採点を行うことで学生に不利益になる採点をなくせた」ことから、本システムを用いることで採点の不平等をなくし評価の質も向上していることがわかる。このように評価者か

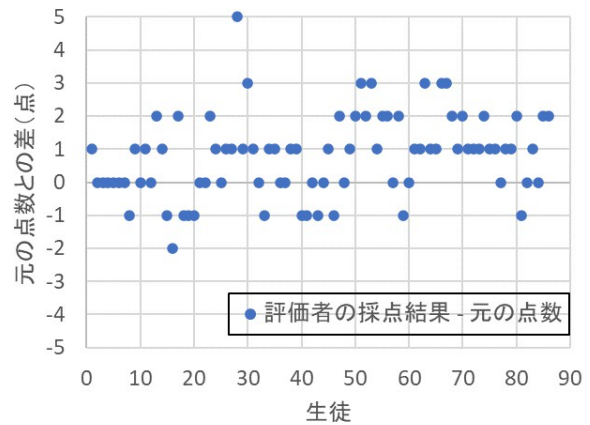


図3: 評価者が採点した点数と元の点数の差

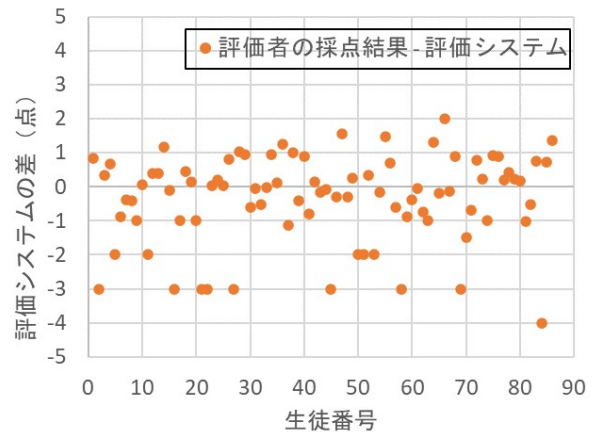


図4: 評価者が採点した点数と評価システムの点数の差

ら見ても評価負荷の軽減、評価の質が向上していることがわかる。

しかし、評価者の採点結果とシステムの評価で差が-3、-4と大きくなった原因として「電子カルテに患者をどのようによくしていくかのプログラム（計画）が書かれてないためシステムの評価と差が出た」と回答をされた。本システムの採点方法では項目ごとの加点単語を用いているか見ているため、患者のプログラムについては見れない。採点結果の差を小さくするためには患者のプログラムを考慮した採点方式を考える必要がある。

6 結言

教育者の負担の軽減、評価の質の向上を目的として、電子カルテの評価を支援する電子カルテ評価支援システムを実装した。

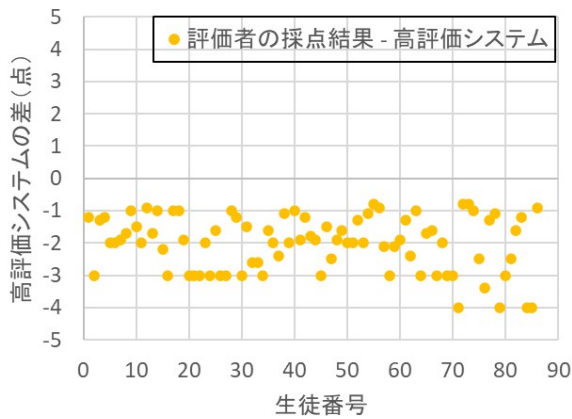


図 5: 評価者が採点した点数と高評価システムの点数の差

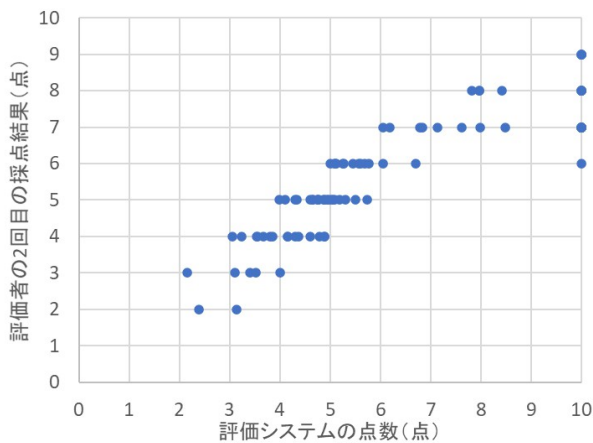


図 6: 評価者の 2 回目の採点結果と評価システムの点数の相関

実装した電子カルテ評価支援システムが教育者の評価負担の軽減、質の向上につながるか検証を行った。検証結果より電子カルテ評価支援システムの評価結果に教育者が多少影響されるものの加点単語、文のハイライトにより安定して評価が行えることがわかった。図6の評価者と評価システムの点数の相関から相関が強く評価者はシステムの点数を参考に評価が行えたことがわかった。また、5.2.2の評価者の回答内容からシステムの評価した点数と加点単語、文のハイライトにより電子カルテを平等に評価が行えることがわかった。以上のことから、電子カルテ評価支援システムは教育者の負担の軽減、評価の質の向上が達成できたと考える。

しかし、課題として患者に対するプログラムの内容が本システムでは評価ができず、評価者の評価結果と差が生じてしまう点である。これは評価方式に患者の

プログラムの内容が十分に書かれているか評価できるように対応を行っていく。

参考文献

- [1] 甲田宗嗣：理学療法士の卒前教育と臨床実習，卒業教育，理学療法の臨床と研究，Vol. 26，No. 1，pp. 9-15(2017)
- [2] 日本理学療法士協会：年齢分布と平均年齢。資料・統計，(URL) <http://www.japanpt.or.jp/about/data/>(2021年1月31日引用)
- [3] 芳野純，白田滋：医療施設における理学療法士の継続教育の現状，理学療法科学，Vol. 25，No. 1，pp. 55-60(2010)
- [4] 椿本弥生，赤堀侃司：主観的レポート評価の系列効果を軽減するツールの開発と評価，日本教育工学会論文誌，Vol. 30，No. 4，pp. 275-282(2006)
- [5] 水本智也，磯部順子，関根総，乾健太郎：採点項目に基づく国語記述式答案の自動採点，言語処理学会第24回年次大会発表論文集，pp. 4-17(2018)
- [6] Kaveh Taghipour and Hwee Tou Ng (2016): A Neural Approach to Automated Essay Scoring, Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, pp.1882-1891
- [7] 中谷安男：英文エッセイの自動レベル判定システムと手動採点結果の比較検証：CEFR-J ライティング・テストタスク構築のための予備調査，法制大学経済学部学会，Vol87，No1・2，pp. 21-50(2019)
- [8] Total Environment for Text Data Mining (テキストデータマイニングのための統合環境)，(URL) <http://tetdm.jp>，砂山渡，高間康史，西原陽子，徳永秀和，串間宗夫，阿部秀尚，梶並知記：テキストデータマイニングのための統合環境 TETDM の開発，人工知能学会論文誌，Vol. 28，No. 1，pp. 1-12 (2013)
- [9] 世界保健機関：ICF 国際生活機能分類，中央法規 (2002)