

■ 原著-技術 ■

標準用語集の普及を目的としたツールの開発

長井 美和 姜 長安 森岡 慶 小野 大樹
横井 英人

医療情報学 第32巻 第6号 (2012)

Japan Journal of Medical Informatics Vol. 32 No 6 2012

日本医療情報学会/篠原出版新社

標準用語集の普及を目的としたツールの開発

長井 美和*¹ 姜 長安*^{1,3} 森岡 慶*² 小野 大樹*¹
横井 英人*¹

医療情報システムにおける情報の標準化を図る方法として、標準用語集の普及が必要である。われわれは、作成された用語集の現実的な運用体制を用語集の作成者とユーザが、互いに用語集を編集しながら使用していく「ライフサイクルモデル」としてモデル化した。さらに、用語集運用の過程をサポートし、標準用語集の国際規格である ISO/TS17117 の理念に基づいたシステムを開発した。

今回、3つの用語集に本システムを適用し、機能評価を行った。また、用語集運用のライフサイクルモデルのどの過程の効率化に寄与できるか、機能の有効性を検証し、さらなる改良を実施した。

作成者とユーザの相互作業を考慮した用語集運用の「ライフサイクルモデル」における効率的な運用を本システムによりサポートし、精度の高い改訂に寄与できる可能性を示せた。今回の経験を元に、本システムを電子カルテのマスタ整備に使用できるのではないかと考え、今後検討していく予定である。

■キーワード：標準化，用語集

Development of Tool for Popularization of Standard Terminology: Nagai M*¹, Changan J*^{1,3}, Morioka K*², Ono H*¹, Yokoi H*¹

The spread of standard terminology and the model for effective operation of standard terminology is indispensable as the method of attempting the standardization of the medical information system. In this research, a handling system is developed to facilitate the work and to make the terminologies quality improving. Functional evaluation of its system was conducted by three terminologies.

Moreover, the validity of the function was verified to see it could respond to the increase in efficiency of which process of terminology operational model, and further improvement was carried out.

*¹香川大学医学部附属病院 医療情報部
〒761-0793 木田郡三木町大字池戸 1750-1

*²株式会社ミトラ

*³独立行政法人理化学研究所 東海ゴム人間共存ロボット連携センター

E-mail : nagai@med.kagawa-u.ac.jp

受付日：2012年2月6日

採択日：2012年8月21日

【第15回日本医療情報学会春季学術大会推薦論文】

*¹Department of Medical Informatics, Kagawa University Hospital

1750-1 Ikenobe, Miki-cho, Kita-gun, Kagawa, 761-0793, Japan

*²Medical IT Laboratory, MITLA Inc.

*³RIKEN-TRI Collaboration Center for Human-Interactive Robot Research

In the future, the developed system will provide to reduce the work for terminology operational. It showed that this system supported the model for effective operation of the terminology considering the interaction between administrator and user.

Key words: Standardization, Terminology

1. はじめに

医療情報システムにおける情報の精度向上には、用語の標準化・コード化が必須である¹⁾。ひいては、標準用語集（用語集）の普及が必要不可欠となり、そのためには、用語集の現実的な運用体制を確立することが重要である。

しかし、用語集運用は非常に複雑な作業を伴い、用語集の作成者・管理者やユーザに多くの負担がかかっている。例えば、作成者・管理者から入手した用語集を独自のシステムのマスタとして使用する際には、ユーザは表計算ソフトやデータベース管理ソフトを元に、マスタの運用管理のためのツールやデータベースを作成し、マスタの維持をしている。また、用語集に対して用語の追加等の要望がある場合、ユーザは管理者が作成した意見や要望を提出するためのフォームを用いて変更要求を行う。一方、管理者は、ユーザから提出された変更要求から改訂に必要な内容を抽出・評価し、改訂版へ反映するかどうかを判断する。この際に行われる変更要求の集計や抽出作業も、非常に多くの手間がかかる。以上のように、ユーザと管理者の双方が、一連の用語集運用に伴う編集作業に大変な労力や時間を要している。

現在、用語集編集作業をサポートするためのツールはいくつか存在している^{2~4)}。しかし、上記のような用語集の運用全体を見据えて作成されたツールの存在は少ない上、ある用語集に特化して作られたものが多い^{3,4)}。われわれも以前より、用語集の編集作業に表計算ソフトで作成したツールを使用してきた。しかし、用語のコピーや貼り付け等の編集機能や入出力機能などの基本的な実装に留まり、運用全体まで考慮した機能の実装には及んでいなかった⁵⁾。したがって、用語集の編集作業には膨大な時間や仕事を必要とし、

高度な IT リテラシーが要求され、必ずしも使いやすいものではなかった。

2. 開発目的

本研究では、用語集運用のモデルを提案し、その過程で生じる用語集の編集作業の効率化を図る、用語集ハンドリングシステムを開発することを目的とする。さらに、以前より編集作業を行っていた種々の用語集に本システムを適用し、実装した機能の有効性を検証する。

3. システム概要

1) 用語集運用のライフサイクルモデル

われわれが提案する「用語集運用のライフサイクルモデル」（ライフサイクルモデル）を図 1 に示す。ライフサイクルモデルには、以下の 8 過程が存在する。用語集は、①作成者により作成され、②作成者・管理者からユーザに配布される。次に、配布された用語集は、③電子カルテなど様々なアプリケーションにおけるマスタとして二次活用する目的で、ユーザにより必要に応じて編集される。その後、編集した内容について、④ユーザから用語集の作成者・管理者へ改訂要求を行う。用語集の作成者・管理者は、⑤改訂要求の検討を行い、⑥必要な改訂を行った上で、⑦改訂版として用語集の再配布を行う。ユーザは、⑧再配布された用語集と自分たちが用いている現状の用語集を比較し、改訂された部分との差分を確認する。そこで、再度必要と思われる内容があれば、④に戻り、さらに改訂要求を行う。ライフサイクルモデルの④から⑧の繰り返しにより、精度の向上した用語集へと改訂が進んでいく。

2) 用語集ハンドリングシステム

これまで蓄積した用語集編集作業のノウハウから、以下 4 項目の開発方針を立案する。

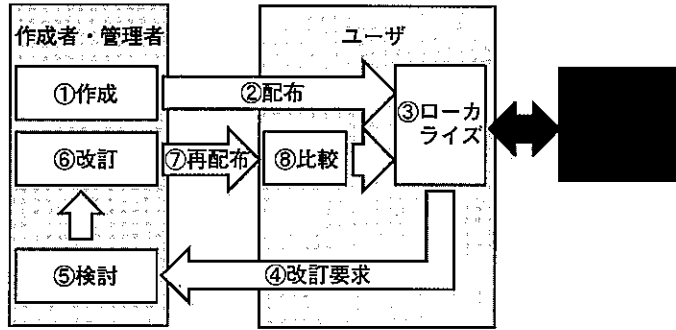


図1 用語集運用のライフサイクルモデル

- 用語集の作成者・管理者とユーザの双方が使用できること
- 高度なITリテラシーを必要としない簡便な操作性であること
- 用語集の構造がヒトの目で理解しやすい閲覧性があること
- 任意の用語集に対して適用できること

また、医療用語集に関する国際規格である、ISO/TS17117 Health informatics-Controlled health terminology-Structure and high-level indicators (ISO/TS17117)⁶⁾の理念に基づき、前述したライフサイクルモデルの8過程をサポートする機能を導入する。

(1) 基本画面

本システムの基本画面を図2に示す。画面の右側には、用語集がツリービューで表示される。画面の左側では、ツリービューで選択した用語の属性等の閲覧・編集が可能である。

(2) 各種機能

- a. 言語の切り替え機能 (図1の①③⑥に使用)
本システムの機能名などの表示を、日本語または英語の表記に変更できる。
- b. 用語集ごとの情報属性のカスタマイズ機能 (図1の①③⑥に使用)
属性の名称・個数等を、独自にカスタマイズできる。
- c. 用語の編集機能 (図1の①③⑥に使用)
用語の追加・削除・コピー・貼り付け、編集作業を元に戻す、取り消した作業をやり直すなど。

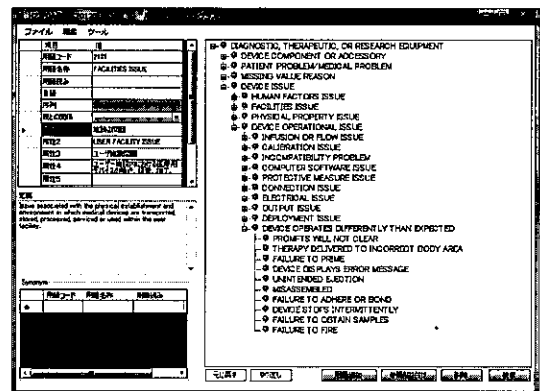


図2 用語集ハンドリングシステムの基本画面

- d. 同義語管理機能 (図1の①③⑥に使用)
ある用語の同義語が、固有のコードや属性を持ち、個別に表示される。
- e. 「親との関係」のビジュアル化 (図1の①③⑥に使用)
用語の属性として、親との関係を規定できる (is-a, part-of など)。アイコンの設定を行い、ツリービュー上で関係性を明示する。
- f. 履歴管理機能 (図1の①③⑥に使用)
用語集編集の作業記録・実施者・日時を、自動的にログファイルに保存する。
- g. 検索機能 (図1の①③⑥に使用)
用語集中の用語や数値を、部分一致で検索できる。検索条件には、大文字/小文字・半角/全角・スペースの無視を設定できる。検索範囲には、コードや定義や同義語を含めることができる。

h. 編集制限機能 (図1の②⑦に使用)

編集できる用語, もしくは属性の範囲を指定する。

i. インポート・エクスポート機能 (図1の②④⑦に使用)

CSV形式で記述された既存の用語集をインポートできる。また, インポートして編集された用語集を, CSV形式か Classification Markup Language (ClAML)形式でエクスポートできる。

j. 用語単位のインポート・エクスポート機能 (図1の②④⑦に使用)

選択した用語とその下位の階層を1つの単位として, CSV形式でインポート・エクスポートができる。

k. 参照貼り付け機能 (図1の①③⑥に使用)

ある「元となる用語」を別の場所に参照貼り付けると, 貼り付けた場所で「参照した用語」として使用可能となる。「元となる用語」の内容を変更すると, 「参照した用語」の内容は自動で変更される。OSなどのファイル管理システムにおける, ショートカットやエイリアス作成に近い仕組みである。これにより, 用語集の中で, 1つの用語が複数の親を持つことが可能となる。

1. 比較機能 (図1の⑤⑧に使用)

異なるバージョンの用語集を比較し, その差分を抽出する。差分は, 背景色を変えて表示される (図3), CSV形式で出力できる。大文字/小文字表記や半角/全角など比較条件を選択できる。

(3) 動作環境

<OS> Windows XP (Microsoft. NET Framework 3.5 必須), VISTA, 7

<CPU> Pentium 4/Celeron 1.5 GHz 以上,
Pentium M 1 GHz 以上

<有効メモリ> 512 MB 以上

4. 用語集ハンドリングシステムの機能評価

以下に挙げる3つの用語集に本システムを適用し, 実装した機能の有効性を検証する。

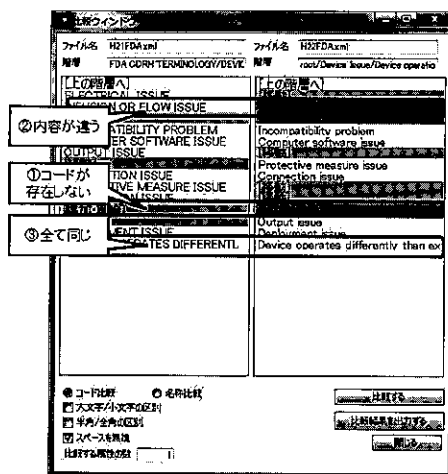


図3 比較機能の表示画面

1) 対象

(1) FDAの医療機器の安全情報に関する用語集

米国では, 行政への医療機器の安全情報報告に, 米国食品医薬品局 (FDA) で作成された用語集⁷⁾が使用されている。この用語集は, 医療機器の有害事象の用語集である Event Problem Codes⁸⁾と, その有害事象を評価する用語集である Evaluation Codesの, 大きく2つのカテゴリーに分かれた階層構造で表現された用語集である。本研究では, 2009年6月 (2009年度版)と2010年8月 (2010年度版)に入手した, 時期の異なる2つのEvent Problem Codesを対象とする。両ファイルは, エクセル形式で配布されている。

(2) Minimal Standard Terminology (MST)

MST⁹⁾は, 消化器内視鏡データの電子記録に必要な所見用語集として, 1995年に世界消化器内視鏡学会により作成され, 現在Ver.3が公表されている。日本語をはじめ, 約10カ国語に翻訳されている国際用語集である^{9,10)}。Ver.2までは, 臓器 (部位) ごとに必要な所見を, 階層構造で表現した用語集であったが, Ver.3では, 複数の表の組み合わせで所見を表現する用語集となり, 大幅な用語集構造の変更がみられる。本研究

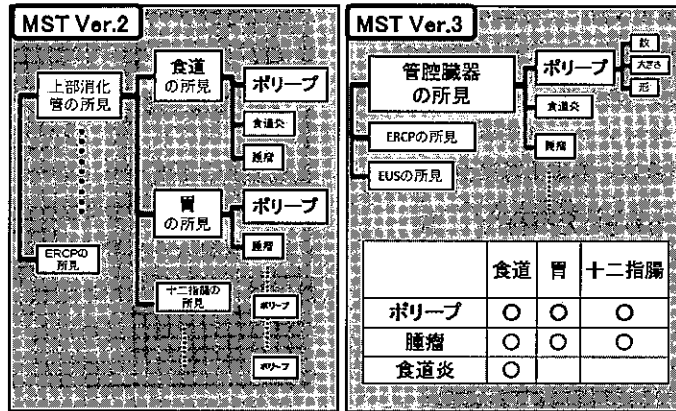


図4 MST Ver.2とMST Ver.3の構造の概略図

では、Ver.3を対象とする(図4)。

(3) 医療機器の不具合等に関する用語集

現在、われわれは、日本医療機器産業連合会(医機連)と合同で、医療機器の不具合報告に用いる用語集を作成している。この用語集は、医療機器ごとに起こりうる不具合の用語集と、健康被害の用語集の2つから構成されている。不具合用語集は、医療機器の名称をカテゴリー別に分類し、不具合事象を総称と名称の二階層とその原因となる部品・構成品で構成している。健康被害用語集は、不具合用語集と同様に医療機器の名称をカテゴリー別に分類し、健康被害を総称と名称の二階層で表している。

2) 評価方法

既存の用語集を本システムに取り込む際の前処理として、エクセル形式のファイルを、本システムに適応する項目に変更し、CSV形式に変換する。

(1) 比較機能

FDAの用語集の2009年度版と2010年度版に、比較機能を用いて差分を出力する。改訂による変更を出力された差分から確認し、その件数を調査する。

(2) 参照貼り付け機能

MST Ver.3の階層構造より、所見記載に使用する用語を、参照貼り付け機能を用いて、食道・胃・十二指腸など部位の下位にそれぞれ貼り付け

る。例えば、「ポリープ」は、食道・胃・十二指腸で使用されるため、それぞれの下位に貼り付けるが、「食道炎」は、食道の下位のみ貼り付ける。

(3) 用語単位のインポート・エクスポート機能

本機能を用いて、医機連の不具合用語集と健康被害用語集を結合する。

まず、健康被害用語集から、健康被害総称と名称の階層構造部分のみを用語単位のエクスポート機能を使用しエクスポートする。それを、不具合用語集の不具合事象と同じ階層に、用語単位のインポート機能によりインポートする。さらに、不具合事象ごとに、該当する部品・構成品を参照貼り付け機能を用い、それぞれの不具合事象の下位に貼り付ける。

3) 評価結果

(1) 比較機能

2009年度版と2010年度版を比較した結果を表1に示す。改訂により、すべての内容が小文字表記から大文字表記に変更されていた。そのため、小文字表記と大文字表記を区別する条件を選択できるように修正をした。その後、比較機能を使用して得られた結果は、用語の追加(6件)、コードが別階層に移動した(1件)、用語名称の変更(2件)、定義のスペース数の変更(3件)、2009年度版で使用されていたコードに別の用語名称が

表1 FDAの用語集の改訂による差分結果

変更内容	件数
小文字表記から大文字表記への変更	すべて
用語の新規追加	6
コードが別階層に移動	1
用語名称の変更	2
コードと用語名称の組み合わせの入れ替わり	36
定義文中の空白の数が変更	3

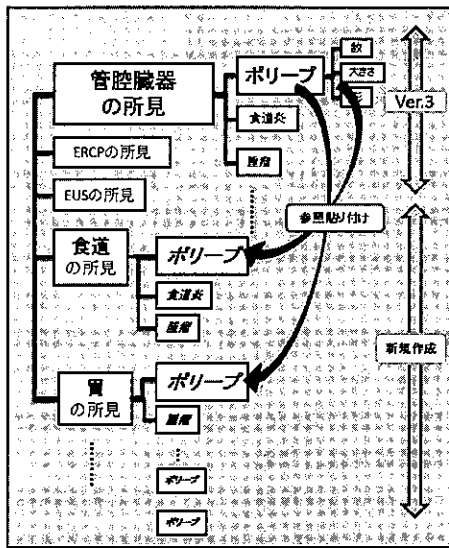


図5 MST Ver.3を階層構造に変更した概略図

組み合わされていた(36件)であった。

(2) 参照貼り付け機能

MST Ver.3を本システムにインポートし、Ver.3の用語から、部位を表す用語とその部位の所見記載に使用する用語を選び、上位下位関係となる階層構造をVer.3中に作成した(図5)。その結果、冗長性が排除されたVer.3の階層構造を保持したまま、Ver.2のような部位別の階層構造も表現できた。これにより、用語集の全体像を把握することができた。また、必要に応じて、参照元の用語の属性を変更する場合、参照貼り付けした他の位置にある同じ用語の属性も、自動的に変更される。

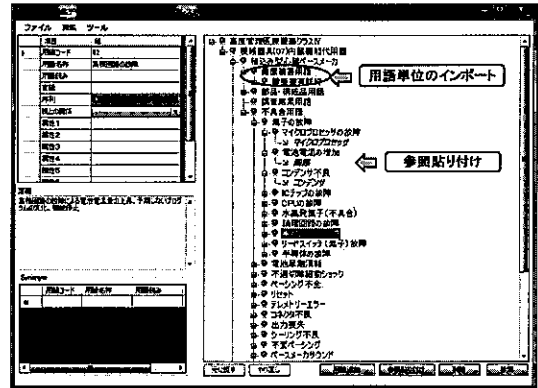


図6 不具合用語集と健康被害用語集の本システム上での結合図

(3) 用語単位のインポート・エクスポート機能

健康被害用語集から、健康被害総称と名称の階層構造部分のみをエクスポートし、それらを不具合用語集の階層構造中にインポートし結合した。さらに、不具合名称ごとに対応する部品・構成部品を、参照貼り付け機能を使用し貼り付けた。その結果、われわれが開発した医療機器の不具合報告システムに使用するマスタとして、新たな用語集を容易に作成できた(図6)。

5. 考察

1) 用語集運用を考慮した編集機能の開発

本研究で実施したように、編集作業の効率化に適切なツールを用いることは重要である。用語集は、使用される中で、時代背景や状況に応じて絶えず変化する。特に試作段階では、表1に示したように、用語の追加・削除だけではなく、階層構造が変化することも想定される。加えて、用語集ポリシーが不完全な場合には、コードの重複や再利用など様々な状況がありうる。ISO/TS17117⁶⁾で示されるように、コードの再利用は用語集を運用する上で好ましくない。われわれもコードはユニークであるべきだと考え、原則的には認めていない。よって、本システムには、コードの再利用を発見する機能は実装していない。しかし、実際には作成ミスなどにより、コードの重

複や再利用が行われている。今回、「比較機能」を用いることで、同一コードで属性が異なる場合の検出は可能であった(表1)。

また、複数の入力場面で、同じ入力ポリシーを持つ用語を使用する場合は、複数箇所の用語間にリンクを持たせる「参照貼り付け機能」が有用であった。一方、同じ用語名称であるが属性が異なり、別の用語と判断されるような場合は、参照貼り付け機能ではなく、編集機能にある「コピー」と「貼り付け」を用いることで、表記は同じであるが異なるコードや属性を持つ用語を作成することができる。

一般的に用語集は、複数の作成者が共同して作成すること多い。複数の作成者の編集に対応するため、「用語単位のインポート・エクスポート機能」を実装した。用語単位のインポート・エクスポート機能を用いれば、個々に作成された用語集を、最終的に管理者が容易に結合することができる。また、一つの利用集を複数の作成者が共同して作成する場合、作成者ごとに編集内容や範囲を制限することができる「編集制限機能」を用いることで、用語集の全体像を俯瞰しながら、自らに依頼された編集内容のみの作業を行うことができる。

現在、北米放射線医学会では、Web用の protégé を元に開発した用語集編集ツールを用いて、RadLex という放射線科領域の用語集を作成・運用している³⁾。RadLex のような大規模な用語集で、世界中に複数の作成者がいる場合、インターネットを介した用語集の一元管理は非常に有用である。しかし、本システムは、国内規模や施設内など比較的少数のグループでの使用を想定していたため、現時点では、Webを使用したシステムを想定していなかった。それよりむしろ、作成者やユーザなど、用語集運用に関わる人の編集作業に必要な機能を追加し、ユーザビリティに重きを置いて開発した。また、何らかの用語集に特化したものではなく、非常に汎用性の高い用語集編集システムを目指している。しかし、RadLex のように対象とする用語集の規模が大きくな

り、それに関わる作成者やユーザの数も多くなれば、Web 上での運用も可能なシステムの開発も必要になってくるとされる。

2) 用語集ハンドリングシステムを使用した用語集運用管理への活用

ユーザが用語集を入手した際、収載されている用語と自らが必要とする用語が完全に一致していることはまれである。また、使用途中に、必要な用語が出現することもある。そのような場合、ユーザが元となる標準用語集へ反映する必要があると判断すれば、所定の文書で変更要求を行う場合が多い。その際、本システムを適用し、同じ形式の用語集データを作成者・管理者とユーザの双方が用いていれば、ユーザがローカライズした後の用語集データを管理者にそのまま受け渡すことができる。その結果、作成者・管理者とユーザ間でシステマティックに情報共有が行える。一方、管理者は、提出された同じ形式の出力データを集計することもできる。この集計により、複数の施設から同じ変更要求が何件提出されたかなど量的データの入手が容易に行える。さらに、ユーザがローカルな用語として取り扱っていた用語の情報も入手でき、そこに伏在している重要な用語を発見する手段ともなりうる。

また、本システムを適用し、同じ形式の用語集データを、作成者・管理者とユーザの双方が用いれば、複数の異なる用語集のフォーマット変更の必要がない。すなわち、電子カルテのマスタのように、複数の用語集をマスタとして取り入れるときにも、同一のファイル形式をとることが可能となり、その後のマスタの整備が効率的に行える。このように、様々なケースに対応できるよう編集機能に幅をもたせたシステム開発をすることは重要である。

6. 結論

今回、用語集運用のモデルを提案し、その過程で生じる用語集の編集作業の効率化を図る、用語集ハンドリングシステムを開発した。さらに、以前より編集作業を行っていた種々の用語集に本シ

システムを適用し、実装した機能の有効性を検証した。結果、用語集のポリシーや作成ミスなどから起因する用語やコードの重複など既存の機能だけでは対応が困難なケースがあることが明らかとなった。しかし、当初考えていた仕様に細かな修正を加えることで対応可能であった。

複数の作成者が用語集を作る場合や、ユーザが入手した用語集を二次利用する際のローカライズ、改訂に伴うユーザと作成者や管理者の編集作業など、各プロセスで生じる作業を効率的にサポートすることが用語集運用の鍵となる。本研究で開発した用語集ハンドリングシステムの機能が、用語集の作成者・管理者とユーザの相互作業を基にした用語集運用のライフサイクルモデルに生じる様々な編集作業の効率化を促せたと考える。また、ライフサイクルモデルに基づき、本システムを適用することで、用語集の問題点を発見し、用語集のさらなる精度向上に寄与できる可能性を示せた。

医療機器の不具合用語集を行政報告に用いるシステムとリンクさせたように、用語集の表現方法を吟味し、電子カルテのマスダ整備などにも使用できるように今後更なる改良を行っていく予定である。

謝辞

本研究は、厚生労働科学研究費補助金（医薬品・医療機器等レギュラトリーサイエンス総合研

究事業 平成 20～21 年度「医療機器の不具合用語の標準化とコード化に関する研究」の助成を受けたものである。

参考文献

- 1) 大江和彦. 医療情報の電子化と用語・コードの標準化. 医学のあゆみ 2007; 221, 12: 1013-1017.
- 2) 古崎晃司, 溝口理一郎. オントロジー構築ツールの現状. 人工知能学会誌 2005; 20, 6: 707-714.
- 3) Rubin DL, Noy NF, Musen MA. Protégé: a tool for managing and using terminology in radiology applications. *J Digit Imaging*. 2007; 20 (Suppl 1): 34-46.
- 4) Tao Y, Mendonça EA, Lussier YA. A "Systematics" Tool for Medical Terminologies. *AMIA Annu Symp Proc*. 2003: 1028.
- 5) 横井英人. 用語集の標準化と個別化の狭間で. 医療情報学 2009; 29 (Suppl): 125-128.
- 6) ISO/TS 17117:2002 Health informatics—Controlled health terminology—Structure and high-level indicators 保健医療情報—制御保健用語—構造及びハイレベルインジケータ. 2002.
- 7) FDA の医療機器安全情報に関する用語集. <http://www.fda.gov/MedicalDevices/default.htm>
- 8) FDA の医療機器安全情報に関する用語集 (Event Problem Codes). <http://www.cancer.gov/cancertopics/cancerlibrary/terminologyresources/fda>
- 9) Minimal Standard Terminology Ver.3 <http://www.worldendo.org/mst.html>
- 10) 日本消化器内視鏡学会用語委員会. 消化器内視鏡用語集. 医学書院, 2011.