

短 報

医師および看護師国家試験問題の電子辞書化の試み

梅村俊彰

Development of electronic dictionary of questions from the National Examination for Doctors and Nurses

Toshiaki UMEMURA

Department of Adult Nursing, Graduate School of Medicine and Pharmaceutical Sciences, University of Toyama

和文要旨

背景：医師および看護師の国家試験問題をEPWING形式の辞書データとし、電子辞書として検索できるようにする。

方法：対象とした国家試験問題は、医師の第95～114回9,890問、看護師の第79～109回7,290問である。問題のテキストデータをHTML形式のデータに変換し、EPWING形式に変換した。

結果：医師および看護師の国家試験問題をEPWING形式の辞書データとした。閲覧ソフトを用いることで、前方一致検索、メニュー検索、複合検索、全文検索が可能である。

考察：辞書データはオフラインで検索でき、串刺し検索で他の辞書と同時に参照できる。EPWING形式は古い規格であるが、国家試験問題を電子辞書とする試みは今後も生かすことができると考えられる。

英文抄録

Introduction: To provide dictionary data in EPWING format for the national board examinations of doctors and nurses, and to enable searching as an electronic dictionary.

Method: The national board examination questions that were considered were the 95th to 114th of 9,890 questions for doctors and the 79th to 109th of 7,290 questions for nurses. The text data in question are first converted to an HTML format and then to an EPWING format.

Results: The national board examinations of doctors and nurses were used as dictionary data in EPWING format. The search software enables forward-matching, menu, compound, and full-text searches.

Discussion: Dictionary data can be searched offline and can be searched together with other dictionaries. The EPWING format is an old standard, but the attempt to make an electronic dictionary of national examination questions can be utilized in the future.

Key words: national exam, EPWING, electronic dictionary

はじめに

医師、看護師を目指す学生にとって、国家試験に向けた学習は欠かせない。医療系の教員にとっても、これら国家試験を知ることは重要である。国家試験の過去の問題は、厚生労働省のホームページでPDFファイルとして公開されている¹⁾。実施された形で見られる反面、機械可読形式でなく、検索には難がある。国家試験問題について、学習用のWebサービスもあるが、多くは会員制であり有料である。

パソコン用の電子辞書の規格として、EPWING形式がある。1988年に提唱された辞書データの共通規格であり、かつては多種多様な辞書が市販されていた。閲覧ソ

フトにユーザ自身で辞書データを入れて使うため、動作は軽快で、インターネット接続を要しない。閲覧ソフトは検索・閲覧の機能に特化しており、検索方法も多彩である。古い規格であるが、ライフサイエンス辞書²⁾やWikipedia³⁾など、継続して更新されている辞書データもある。また、医療における辞書を自作した例として、医薬品などに応用したものがある⁴⁻⁶⁾。

今回、医師国家試験および看護師国家試験の問題をEPWING形式の辞書データとし、パソコンやスマートフォンで電子辞書として検索できるようにすることを目的とした。

対象および方法

1. 対象

対象としたのは、医師および看護師の国家試験問題で、医師は第95回（2001年）～第114回（2020年）、9,890問、看護師は第79回（1990年）～第109回（2020年）、7,290問である。問題データは、属性（回（年）、ブロック、設問形式、科目、問題番号）、問題文、選択肢、解答の各項目が並んだテキストデータとした。

2. 方法

国家試験問題のテキストデータを、辞書の元となるHTML形式のデータに変換し、次にオンラインウェアを用いてEPWING形式のデータに変換する。辞書では一般に、見出し語に続いて、品詞、語義、用例等の本文が続く。国家試験問題を辞書とするにあたって、見出し語には、回、ブロック、問題番号から成る問題コードを用いた。例えば、医師国家試験における問題コード「108A001」は、第108回のAブロックの問1を表す。看護師国家試験における問題コード「113AM001」は、第113回の午前の問1を表す。本文は、実際の試験のように問題文と選択肢を表示し、合わせて解答を示す。また、辞書の品詞情報に相当するものとして、設問形式や科目を提示する。

辞書の元となるHTML形式のデータは、見出しを<h3></h3>タグ、本文を<p></p>タグとした簡易なHTML文書である。また、回（年）を<h1></h1>タグ、ブロックを<h2></h2>タグとすることで、辞書が階層メニューを持つようにした。画像はJPEGファイルとし、文中にタグで挿入した。この方針に基づき、国家試験問題のテキストデータからHTML形式のデータに、Perlスクリプトを用いて一括変換した。

次に、HTML形式のデータを、オンラインウェアのEBStudio2⁷⁾を用いて、EPWING形式の辞書データに変換した。EBStudio2は、様々なテキストデータからEPWING形式の電子辞書を作成するツールである。別途、複合検索用の定義ファイル、凡例ファイルを準備する。その他、変換時のオプションで設定したのは主に、「前方一致表記」、「複合検索」、「<h1>..

######

結 果

1. EPWING形式の電子辞書

医師国家試験および看護師国家試験の問題を、EPWING形式の辞書データとした（<https://square.umin.ac.jp/tt>）。閲覧ソフトにこの辞書データを入れることで、種々の方法により国家試験問題を検索することができる。

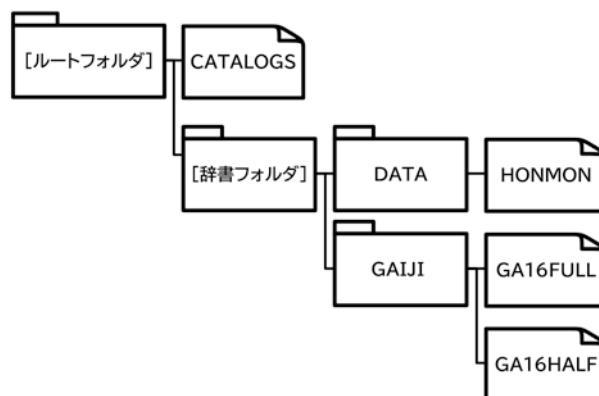


図1 EPWING形式のファイル構造

EPWING形式の辞書を使うためには、EPWING形式に対応した閲覧ソフトが必要である。代表的なものにEBWin、EBPocketがある⁷⁾。閲覧ソフトによって、辞書データの配置場所や設定方法は異なる。EPWING形式の辞書データは、フォルダ中に決まった構成で辞書データや目次ファイルが存在しなければならない（図1）。そのため、データを含むフォルダを、そのままコピーする必要がある。その後、閲覧ソフト内の設定で、フォルダあるいは目次ファイルの「CATALOGS」ファイルを指定する。これらの設定後、閲覧ソフトで検索した画面の例を示す（図2）。

2. 検索方法

国家試験問題の辞書データでは、前方一致検索、メニュー検索、複合検索、全文検索が可能である。検索方法の呼称は、閲覧ソフトによって異なることがある。それぞれについて以下に述べる。

前方一致検索では、見出し語である問題コードから、問題を引くことができる。問題コードは、回、ブロック、問題番号から成り、いつどの問題か指定して探す場合、一番早い。問題コードの問題番号は、頭に0を補って3桁に揃えているが、0を省いてもよい。例えば、「108A001」を検索するには「108A001」でも「108A1」でもよい。後者では、同時に「108A10」なども検索される。また、閲覧ソフトによっては該当項目だけでなく、書籍のように周辺の問題も続けて閲覧することが可能である。

メニュー検索は、分類された項目の階層をたどって、目的の項目に至る検索方法である（図3）。最初の画面で回（年）のリストが表示される。回を選べるとブロックのリストに進む。さらに、ブロックを選べると問題のリストとなり、問題を選んで見ることができる。リストでは、回やブロック名に続けて、問題数などの内訳を付した。この検索方法では、回やブロックの階層から、問題を俯瞰することができる。

複合検索は、回やブロック、科目、種類といった問題

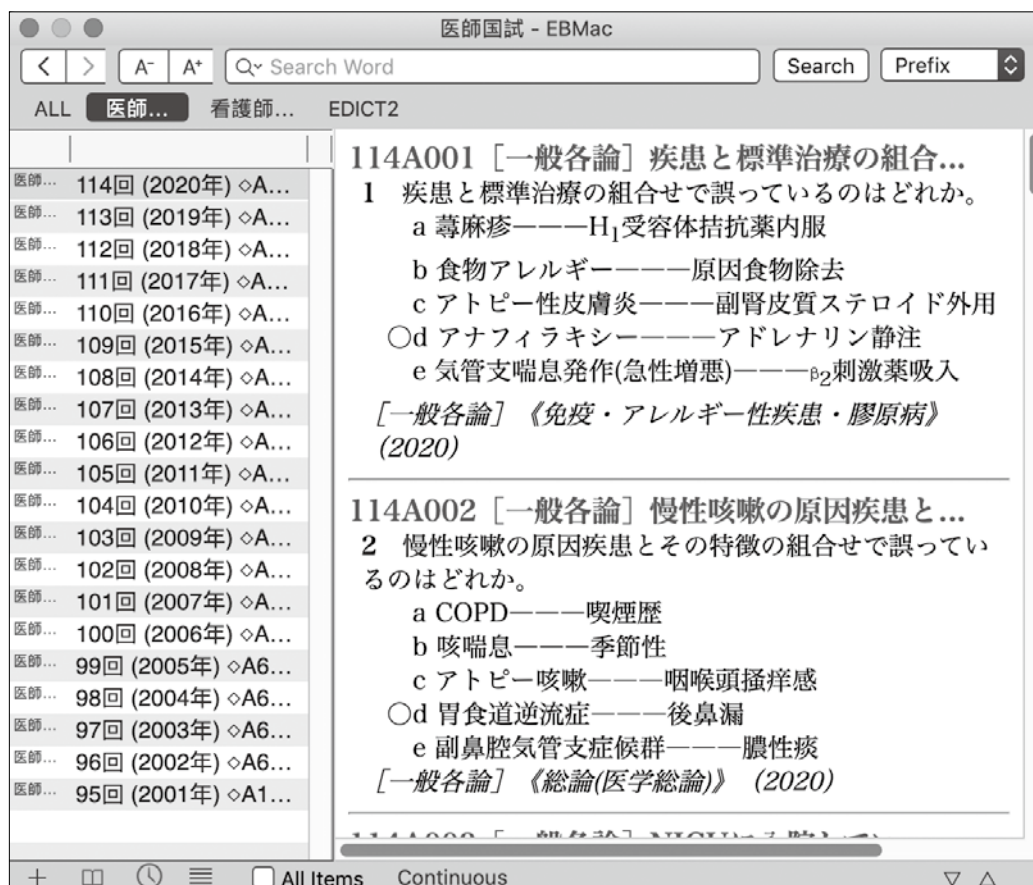


図2 閲覧画面

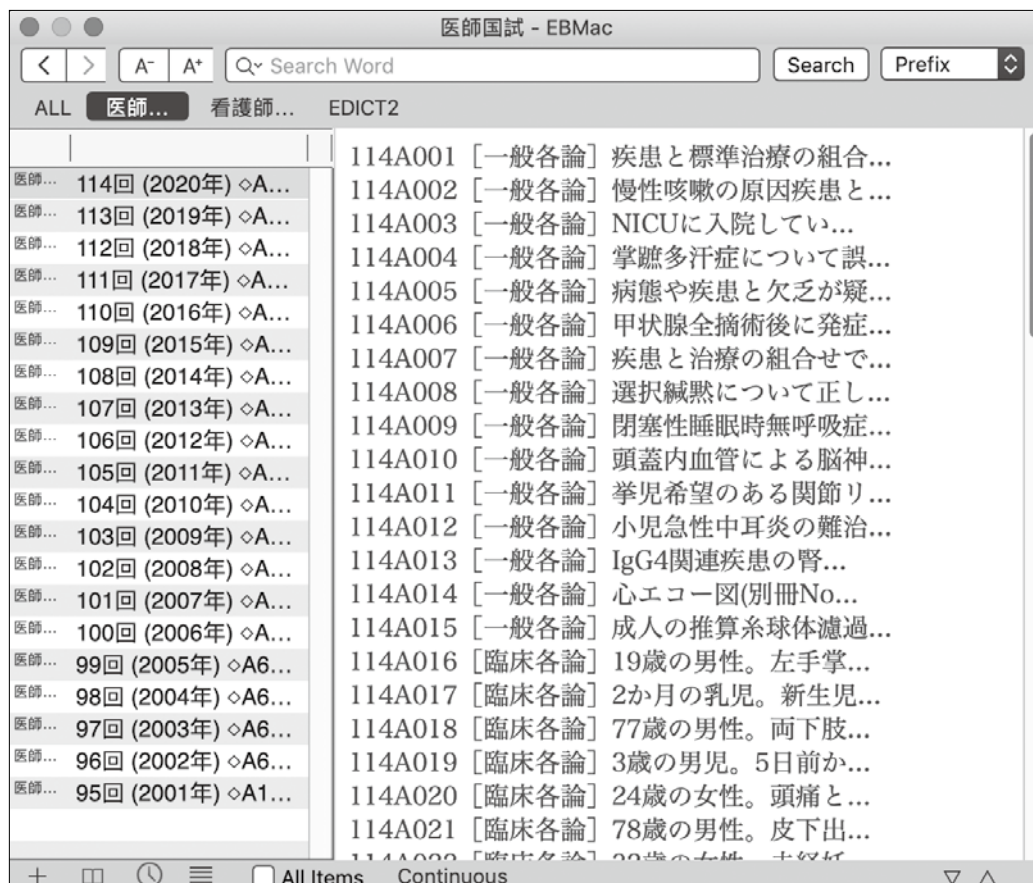


図3 メニュー検索画面

の属性を指定して検索できる。検索ソフトで複合検索を選ぶと、回や科目など、属性を選択するメニューが表示される（図4）。指定したいものをメニューから選ぶことで、条件にあった問題が抽出される。特定の属性の問題を、一覧する際に有用である。

全文検索では、国家試験の問題文すべてを対象として、指定した文字列を含む問題を検索できる。病名や用語などから、問題を探すことができる。全問題のデータを走査するため時間がかかるが、今のパソコンでは概ね、他の検索方法と遜色ないスピードで検索できる。なお、全文検索では、2つの注意点がある。1つは、用語の表記揺れである。検索では入力通りの文字列で検索され、同一用語でも表記の異なるものは検索されない。例えば、「インスリン」と「インシュリン」は別に検索する必要がある。もう1つ、外字を含む検索はできないが、これはEPWING形式に起因する。

考 察

国家試験問題を電子辞書として検索できるようにした。教員にとって、国家試験問題を縦横に検索でき、過去問を閲覧できる。これにより、教材として広く活用する基盤となると考えられる。

EPWING辞書データとしたことによる利点は、はじ

めに述べた。パソコンやスマートフォンなど多くの環境で、実績のある閲覧ソフトを使うことができる。閲覧ソフトの多彩な検索機能を使って、オフラインで検索ができる。パソコンが大容量、高速化したことで、かつて大きかった辞書データが問題なく扱えるサイズとなり、多量の辞書の串刺し検索や、全文検索が実用的となった。

現在、パソコンやスマートフォン上の辞書では、アプリとしてインストールするだけで使えるものも多い。また、Webサービスの辞書であれば、インターネットを通してアクセスすれば使うことができる。これに比べEPWING形式の電子辞書は、閲覧ソフトの他に、別に辞書データを入れて設定する必要がある。これは一面で、音楽データと再生機、書籍データと閲覧装置の関係に似ている。辞書データ自体は、将来、パソコン環境が変わっても、その環境での閲覧ソフトを用意することで、辞書として使うことができる。

今回、見出しには問題コードを用いた。問題コードは参考書などでも類似の表記が用いられ、問題情報をやり取りするのにポピュラーな方法と考えられた。国家試験問題を活用する辞書として、今後も、見出しや本文を工夫する余地がある。

この辞書データでは、不完全であるがキーワード検索の機能を持つ。これは、辞書自体にあらかじめ設定され



図4 複合検索画面

たキーワードを元に検索するため、全文検索よりスピードが速く、複数のキーワードを指定して絞り込むこともできる。しかし、キーワードは機械的に生成しており、必ずしも個々の問題の正しいキーワードとはなっていない。キーワードの精度を高め、正しく検索できるようにするのが今後の課題である。その場合、キーワードとして、問題に含まれる用語に限らず、標準化された医学用語を含めることも考えられる⁸⁾。

パソコンで扱える文字の集まりを文字集合という。現在、多くのパソコンやスマートフォンでは文字コードとしてユニコードが用いられており、JIS第3水準、第4水準を含む多くの文字を扱うことができる。しかし、EPWING形式では、扱える文字はJIS第1水準、第2水準までに限られる。それ以外の文字は外字として扱われることになるが、外字の定義は辞書によって異なる。国家試験問題の場合、ギリシャ文字、修飾付きアルファベット、漢字の一部などが外字となる。医師、看護師の国家試験に登場する外字を表に示す(表1)。外字の表示について、閲覧ソフトの側で適宜、相当するユニコードの文字に置き換え、内蔵のフォントで美しく表示する機能を持つものもある。また、外字の検索についても、ユニコードの文字を入力することで、閲覧ソフトで外字に置き換えて検索するものもある。いずれも閲覧ソフトによる拡張であるため、今回、漢字についてはなるべくJIS第1水準、第2水準内のものに置き換え、必要となる外字を少なくした。

EPWING形式は古い規格であり、表示や検索の面でやや難がある。一方、電子辞書において、EPWING形

式に代わる共通規格はまだない。EPWING形式は、これまで出された辞書や変換ツールの蓄積も多く、今後、あらたな辞書データの共通規格が生まれても、EPWING形式から変換して、辞書データ自体は活用できると考えられる。

今後の課題

さらに広い年代の国家試験問題を収録することが考えられる。新しい問題を追加すること、古い問題を遡って収録することで、より辞書としての価値が高まる。また、保健師や助産師など、他の資格に拡大すれば、複数の資格の国家試験問題を横断して検索できる。

結 語

- ・医師国家試験、看護師国家試験の問題のEPWING形式の辞書データを作成した。これにより国家試験問題を電子辞書として活用することができる。
- ・可能な検索方法は、前方一致検索、メニュー検索、複合検索、全文検索である。
- ・今後も新しい問題を追加し、教育に生かせるツールとしたい。

謝 辞

EBStudio 2 の開発者、および、EPWINGビューアを熱心にサポートしてくださっている開発者に感謝申し上げます。

表1 国家試験に現れる外字

ユニコード	名称/よみ	用例
°	U+00B0 DEGREE SIGN	90°
±	U+00B1 PLUS-MINUS SIGN	潜血(±)
×	U+00D7 MULTIPLICATION SIGN	0.4×-0.5D
÷	U+00F7 DIVISION SIGN	50÷200,000
ç	U+00E7 LATIN SMALL LETTER C WITH CEDILLA	Behçet病
é	U+00E8 LATIN SMALL LETTER E WITH GRAVE	Café au lait斑
è	U+00E9 LATIN SMALL LETTER E WITH ACUTE	Ménière病
ö	U+00F6 LATIN SMALL LETTER O WITH DIAERESIS	Sjögren症候群
ü	U+00FC LATIN SMALL LETTER U WITH DIAERESIS	Kübler-Ross
α	U+03B1 GREEK SMALL LETTER ALPHA	インターフェロンα
β	U+03B2 GREEK SMALL LETTER BETA	β-D-グルカン
γ	U+03B3 GREEK SMALL LETTER GAMMA	γ-GTP
δ	U+03B4 GREEK SMALL LETTER DELTA	δ-アミノレブリン酸
μ	U+03BC GREEK SMALL LETTER MU	950/μl
曖	U+566F あい	曖気
瘰	U+75E4 ざ	瘰瘡
癩	U+762D ひょう	癩疽
癰	U+7664 せつ	鼻癰

文献

- 1) 厚生労働省：第113回医師国家試験の問題および正答について：https://www.mhlw.go.jp/seisakunitsuite/bunya/kenkou_iryuu/iryuu/topics/tp190415-01.html（参照日：2020年3月）。
- 2) ライフサイエンス辞書プロジェクト：ライフサイエンス辞書：<https://lsd-project.jp/ja/project/pc.html>（参照日：2020年3月）。
- 3) Bookends：ウィキペディア日本語版：<https://ja.osdn.net/projects/boookends/>（参照日：2020年3月）。
- 4) 金子周司：薬学英語教育に有用なスマホ・タブレット辞書. 医学のあゆみ 210 (13) : 1062-1063, 2004.
- 5) 佐藤篤郎, 天野哲史, 鈴木厚志ほか：携帯端末 iPhone (R) を利用した高速検索が可能な電子院内 医薬品集の構築事例. Jpn. J. Drug Inform 13 (1) : 8-12, 2011.
- 6) 谷村晋：国際保健用語集を電子辞書に変換するプログラムの開発. Journal of International Health 29 (2) : 75-79, 2014.
- 7) hishida：EB series support page：<http://ebstudio.info>（参照日：2020年3月）。
- 8) 日本医学会：日本医学会医学用語辞典Web版について：<http://jams.med.or.jp/dic/Notes2.html>（参照日：2020年3月）。