デジタル辞書の現在とこれから

2022年7月8日

株式会社ディジタルアシスト
永田 健児

自己紹介

1990　日外アソシエーツ(株)入社

- 書誌DB、新聞記事DB、電子ブック、EPWING…

2000　イースト(株)入社

- 官報XML、DicX、三省堂WebDictionary…

2001　ディジタルアシスト起業

- 辞書事典類のデジタル化専業
- 辞書専用XMLフォーマット「LeXML」の策定・公開
- IEC TC100/TA10で国際標準策定

◆日本電子出版協会(JEPA)理事／レファレンス委員会委員長

◆次世代辞書研究会発起人・幹事

デジタル辞書の基本と進化

電子辞書の進化を知ろう

電卓から電子辞書へ

- 1979年、シャープがポケット電訳機「IQ-3000」を発売
- 1980年代前半、液晶やボタンなど共有部品の多い電卓メーカーが次々に商品化
- 1990年代前半、フルコンテンツ収録モデルが登場
- 21世紀に入って収録コンテンツ数競争が勃発
- 音声再生、カラー液晶、メモリカード対応など高機能化
- 年度モデルビジネスが成功し、現在は学生向けが主流

辞書検索の基本を理解しよう

EPWINGと電子ブック

- 1987年、WINGフォーマットによる『広辞苑 第三版 CD-ROM』が発売される。
- 1990年電子ブックコミッティ
  1991年EPWINGコンソーシアム
  →CD-ROM媒体で多数の辞書事典が商品化される。
- 辞書検索のユーザインタフェースの基礎が確立。
  現在のダウンロードアプリに引き継がれる。

身近なデバイスを活用しよう

ケータイ辞書とスマホアプリ

- 1999年、iモード版「三省堂辞書検索」サービス開始。
- 2005年、ビッグローブが全キャリアで「ジーニアス」ほか辞書検索サービスを開始。
- 2006年～、各キャリア/各携帯端末メーカーが辞書バンドルを推進。
- 2008年、ソフトバンクがiPhone 3Gの取り扱いを開始。
- ヘビーユースから“お守り”へ。

価値ある情報には適正な対価を

オンライン辞書サービス

- 2001年、「三省堂WebDictionary」「ジャパンナレッジ」「ウィキペディア日本語版」がサービス開始。
- 前後して各社ポータルサイトが競うように辞書検索サービスを開始。
- ネットユーザの間で“辞書は無料”のイメージが定着してしまう。
- タダより高いものはない。

デジタル辞書の現在と課題

平易でわかりやすい説明を

辞書の記述は約束だらけ

- 辞書独自の“約束事”が多い。凡例をきちんと読んで理解しないと、辞書の記述は正しく伝わらない。
- 見出し語の省略とか

  travel ～ed; ～･ing; ((英))～led; ～･ling
  resort to ～ 暴力に訴える
  May the *F*～ be with you. 理力がきみと共にあらんことを
  join* [*combine*] ～*s 〈…と〉力を合わせる，協力する〈*with*〉.

- 言い換え･省略可のカッコとか

  anhidrosis 〔医〕無(発)汗(症).
  antifungal 抗真菌薬[物質,因子].
  acinous 〔植〕小核(果)[(腺)小胞]からなる[を含む]

- 基本的に書籍版の都合→デジタルなら解決できるはず。
- 現況では音声合成･自動読み上げも無理。

もっと用例を

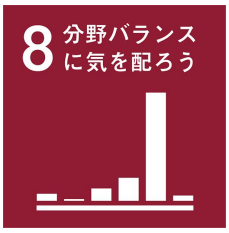
用例に語らせよう

- 書籍編集でスペース的に一番割を食っているのが用例（という印象）
- デジタルならば
  - もっと多くの用例が収録できるはず
  - 用例中、用例前後の省略も最低限ですむはず
  - 用例が多すぎても、
    - 適切な分類で見やすい配列にすることもできるはず
    - ユーザに合わせた優先順位を与えることも(将来的には)できるはず
- だから、用例データは大切に蓄えておきましょう。

数で勝負するのはやめよう

収録数とかサバ読まない

- 「類書中最大、〇万語を収録!」みたいな宣伝文句をよく見かけますが…
- デジタル化すると見出し語も用例も1桁単位で明確に。
  ：少なくともベンダーには筒抜けで、版元から「なんとか数字を合わせて」と言われることもしばしば。
- 各辞典でカウント方法に独自ルール（+四捨五入の幅）
  ：少なくとも学習辞典においては、カウント方法を明らかに、もしくは統一した方がよいのでは？
- やはり学習辞典においては、項目数より優先したいものがあるとも思います。



分野バランスに気を配ろう

百科語・専門語

- 企画編集や(常時)改訂の中で、
  何かしらの分類体系基準をもち、
  ユーザニーズに気配りしながら、
  各分野の比率や実数を把握・意識しながら
  編集(立項)していくことはとても重要。

- ラベル記述にも気を配りましょう
  例1）〔クリ〕
  例2）〔コン〕〔コンピ〕〔コンピュ〕〔コンピュータ〕〔コンピューター〕〔電算〕

辞書ユーザを育てよう

凡例とチュートリアル

- 辞書の記述は約束だらけ
  ⇒辞書引き熟練者以外にとってはとっつきづらい
  ※確かに慣れてしまえばとても便利な完成形なのだけれど
- 多くの人が、あまり凡例を読まない
  そして、デジタル版での凡例の使い勝手がもうひとつ
- デジタル版での凡例誘導の工夫、凡例参照の必要を減らす工夫などを通して辞書引きのハードルを下げるべき。
- チュートリアルがあってもいいのでは？

正確で、より新しい情報を

定期更新と常時改訂

- 国語辞典、英語(外国語)辞典、百科(専門分野)事典の別にかかわらず、言葉は次々に生まれ、常に変化している。
  (ついでに言うと、日々、誤植なども見つかる)
- デジタル版では書籍版の改訂を待たずに、新語追加や時事に則した内容修正(そして誤植修正)が可能。
- 常時改訂の体制を整えて、定期的な更新をデジタル版のアドバンテージとして積極的に進めるコンテンツも増加。

デジタル編集に適応しよう

編集者も進化しないと

- よりデータベース的な視点で。
- でも「とにかくエクセルで」はやめて。お願い。
- 「見た目の編集」から「スタイルシート的思考」へ。
- 責任とコスパを考慮して、仕事の切り分けを（餅は餅屋）。
- デジタルデータの査収に必要な知識はしっかりと。
- テキストエディタ使いましょう。
- 正規表現、覚えると便利ですよ。

デジタル辞書のこれから

12 国際標準の重要性を理解しよう

国際標準の重要性を理解しよう

IEC62605とLeXML

- 辞書コンテンツの、①構造上のクォリティを一定のレベルに保つため、②デジタル化およびシステム投入のコストを抑えるため、標準的なデータ規約は有効。
- 国際競争の観点からも（WTO:TBT協定/GP協定）多くのコンテンツおよびデジタル辞書デバイス/サービスを実績として持つ日本が国際標準化のイニシャティブを取ることは重要。
- 2011年、日本が主導した辞書XMLの国際標準化が「IEC 62605」として発行された。2021年にed.3。
- 国内で一定のシェアをもつ「LeXML」がそのベースとなっている。

13 外字のない世界へ

外字のない世界へ

「字種」と「字体」と「字形」

- 辞書・事典はホントに“外字”が多い。
- 組版段階、表示デバイス/システムのユニコード対応で、以前に比べるとかなり改善されてきた。
- 一方で同一コンテンツ内での字形の揺れや全文検索対応の問題なども。

  ※「叱」(U+53f1)と「叱」(U+20b9f)とか。
- 異体字セレクタの活用が理想的。しかし、コスト面などでまだハードルが高い。

より良い理解とファクトチェック

辞書比較とソース確認

- より詳しく、より正確に、言葉の意味・内容を知るためには「辞書の引き比べ」と「資料の参照」が有効。
- 複数辞書の串刺し検索、検索結果を効率的に比較できるUIなど、辞書引き環境のさらなる発展に期待。
- 専門性の高い辞書事典、原典や統計資料などにシームレスにアクセスできる知的空間の拡大・整備も。

マルチメディアを再考しよう

ワンソース・マルチユースも

- Windows95が起爆剤となった“マルチメディア”ブーム。
- デバイスの基本性能やネットスピードが飛躍的に向上した現在、より適切で効果的なメディア参照が必要。

- 図版は書籍版流用が最適解なのか。写真かイラストか。権利関係処理の問題。

- 見出し語発音におけるネイティブ音声の評価。朗読音声の活用。
- 音声合成と自動読み上げへの対応。

多様性とパーソナライズ

コンテンツ制作段階でできること

- 利用者のニーズや理解レベルに応じたコンテンツを揃える責務。これは書籍版でもいえることですが。
- ワンソースで異なるニーズ/レベルに対応できるコンテンツ（の模索）。
- ワンソースで利用者（の指定）に応じて表示を変化させることができるUIの可能性。
- 検索方法も、検索結果の選択方法も、項目内容の提示方法も、検索履歴の管理方法も、すべて多様化の方向。
- まずは“利用のされ方”に対応できるコンテンツを準備していく必要がある。

次世代の辞書を創り出そう

まとめに代えて

- **まとまりのない話が続き、申し訳ございませんでした。**
- 電子出版の先駆けとなったデジタル辞書は、一旦の完成形を得たが故に、ここしばらくはエポックメイキングな革新を体験していません。やや下降気味ながら、それでも一定のビジネスモデルを維持しています。
- とはいえ、ここまで見てきた通り、数多くの「やればできること」「やらないといけないこと」が存在します。
- ボーンデジタルの辞書コンテンツ、書籍の段階から構造変革した新しい辞書の誕生が望まれます。
- 次世代の辞書を一緒に創り出しましょう。