

辞書作成のリエンジニアリング

岡地優司・山崎 昶

国立電気通信大学大学院 情報システム学研究科
国立電気通信大学 電気通信学部 電子物性工学科

本論文では、限られた短いスケジュールで辞書を完成する1つの方法が記述されている。大部分の辞書は、多数の専門家に執筆を依頼し、膨大な時間をかけ、それに加えて大人数で多数回の編集会議を経た結果として作成されている。例えば「化学大辞典」(共立出版)は完成まで10年を要し、約百人の編集委員のもと、延べ400人強の専門家が執筆にあたっていた。これでは辞典が完成する頃には「古語辞典」になってしまうことが度々起こる。そこで我々は約2200語からなる情報分野の辞典を、限られた数の専門家達と新進気鋭の大学院生達の力を借りて、期間3ヶ月で完成させることに成功した。

RE-ENGINEERING OF PRODUCTION
A DICTIONARY OF INFORMATION

Yuji Okachi, Akira Yamasaki

The University of Electro-Communications
Graduate School for Information System
Department of Applied Physics and Chemistry

This paper describes a way to make a dictionary of information science in very tight schedule. Many dictionaries are usually the production of numerous famous professionals under the busy and tedious editorial committees. For example, 「ENCYCLOPEDIA CHIMICA (Publisher :Kyoritsu-Shuppan)」 needed almost 10 years and 400 professionals. Therefore, it needs many years until publication, and the contents and explanations are occasionally inclined to become old-fashioned and meaningless. We have succeeded in preparation of the complete manuscripts to overcome these problems for a novel "Dictionary for Information Science" in only almost three months. Although this dictionary is pocket-book size, it contains above 2200 terms, whose explanations have been written by limited number of special writers including fresh graduate students of proper research area aiming for the beginner's guide and readers' good understanding.

1. 序論

によって現ル問はか日みのした
部署に於ける力なで。て業のれ
分野はあく大一であるにた多論
業、葉をテ、間界で業しは議
企く言目たば世業明企う業に
、多く注しれ、の自いこ企め
にが言がうみしそはなてるた
うととスこてきはのらめいる
よこ」一。し歩境るな含てめ
るうドべるに人環あはをし進
あ使一夕あ達一たでてど手く
がをカーで人てし題くな着良
慣言「デ常いしう問な0に率
習方、な通などそなて0成能
と、ば々がら項。き育0作を
化りえ様の知事る大を9のれ
文た例、すく解い、れ0書そ
にっ。が指全理てばそS辞そ
れ創いたをてのいれ、1のて
ぞを多い」いア招み近、用て
れ葉がてスつ二をて最で内た
そ言とし一にマ果しるこ社あ
て、にこ味べ世界の結にいそに
て手な意夕世界の結にいそに
い勝激を一のす新れる実す
おが急」デそてくるわあ、成す
にれがド型、しくす問でく作い
合れ変力一題えれう性問す書な
集そのチカ問、らよ産な決辞去、
のほソ「のりれし生要解、過で
トてはバはムあ活躍の重をこ
おいて「で一でけ活人ば題ると
おは代夕題受ら本れ問かこ
典く作辞らて短自る
かく「品す

2. 辞書の存在意義

ところで、辞書を作成することの重要性は以下の点にある。

- ①テクニカルタームの問題は、その世界について全く知らない人々にとって大きな障壁となることが多く、一部のマニアのみで理解事項として一人歩きし、世間にはなかなか受け入れられないという傾向を生み出す。
- ②それぞれの業界の新人にとって大きな問題である。新人を迅速に育成して稼ぎ手になってほしいる企業にしてみれば、どうしても見逃せない問題でもある。（この言葉に関する問題はこれから話題になるIS09000にも関係している。）

こうした理由から辞書は本来、その使命（わからない言葉をわかるようにさせること）からすると、単語の意匠が大きく変化したり、単語自体が死語となる前に完成させなくてはならない。すなわち辞書の使命をまっとうするためには

- ①単語の意味自体が変化する前に作成を完了させる必要がある
- ②読んだ人なら誰でも理解できるものである
- ③改訂が可能であること

以上の3点が守られなくてはならないことになる。

3. 今回の辞書作成の背景

今回「情報マルチメディア辞典」を作成することになった背景として以下のことがあげられる。

- 1 中学校、高等学校における新科目「情報基礎」で使えるような、手軽でやさしい辞典がなかったこと
- 2 年々急増する情報処理技術者のニーズ
- 3 海外で標準となってきたISO9000やJISへの対応

そこで、今回は以下のような手順（出版社との交渉などは除く。）に従って、辞書の作成を行った。

- ①辞書に掲載する単語の選定
- ②メンバーの選定
- ③中間段階のサンプル作成
- ④サンプルの評価と修正
- ⑤項目の見直し（見直しと削除）
- ⑥本格的な執筆原稿の作成

4. 今回の辞書作成の手順

①辞書に掲載する単語の選定

辞書で必要な単語のリストは、個人の主観にとられないようにするため、他の本や辞書（複数）を参考として作成する。まず、いくつかの重要な専門書や白書などの巻末についている索引を、表計算ソフトを使ってデータベース化するところから始める。表計算ソフトを利用するのは以下の理由による。

- ①ソートが行いやすい
- ②カード型データベースはカードの定義が煩わしく、変更の度にフォーマットの修正を行わなくてはならないが、これはきわめて面倒である。
- ③数千からの項目を扱うため、カード型のデータベースではメモリの大きさに不安を感じる。
- ④簡単に全体を見回すことができる。（マネジメントを容易にする。）

候補となる単語の選定に関しては、編集委員の偏見や好き嫌いを極力排除す
るために、他の辞書や参考書に力を入れて、今回の場合、内外の約15冊の辞書を
とらぬ専ら専門書（マル学や高校の情報基礎で登場する。）で掲載
の数冊の専ら専門書（マル学や高校の情報基礎で登場する。）で掲載
とは最近の分野の単語をよって特定の意味を持たせて
さすも、分野によるよって特定の意味を持たせて
る。）このよくなるして集めた単語をさらに一つのファイルに集めてマージし、残
次に頻度の低いものは削除を検討する。さらには重要度が高いと考えられるので極力残
し、本来のこのよくなるして集めた単語をさらに一つのファイルに集めてマージし、残
使ったり本偏った単語の中と性となる。果、約2600語以上の単語が集まり、さらに検討の結果、
さす定るとさると単語の可と性となる。果、約2600語以上の単語が集まり、さらに検討の結果、
も守るこの重複と性となる。果、約2600語以上の単語が集まり、さらに検討の結果、

この中から別項参照の単語をその中から探す。一般的に常識でわかる単語は、見
つ最後に次第組み「***」にあたる語）は、自分だったらどちらの項目で引くこ
の項目（前述の「***」にあたる語）は、自分だったらどちらの項目で引くこ
とが多いかの判断した。この作業により約300語が消えた。同時に、登録され
た項目の英語名を別枠で登録する。これは後に英字名でソートしたときに類似

語は同じ英字名かかなり似たような英字名になることで判断できると考えたからである。例えば以下のようなケースもある。

エレメント (→素子) 素子	element element
J C L ジョブ制御 (→ J C L)	job control language job control language

②メンバーの選定

多くの辞典類は数多の執筆者を誇るが、その規模が大きいほど、その内容が専門的になる傾向がある。また、その分野が広いほど、その内容が浅くなる傾向がある。したがって、メンバーの選定には、その分野が専門的であり、かつその分野が広いことが望ましい。また、その分野が専門的であり、かつその分野が広いことが望ましい。また、その分野が専門的であり、かつその分野が広いことが望ましい。

- | | | | |
|---|-------------------------------|-----------|----------|
| 1 | コンピュータ一般
電通大 電気通信学部 電子情報学科 | 学部生
教官 | 3名
2名 |
| 2 | ハードウェア
電通大 情報システム学研究科 | 院生 | 1名 |
| 3 | マルチメディア、法律関係
新潟大 法学研究科 | 院生 | 2名 |
| 4 | 表計算ソフト、ワープロ関係
電通大 ネットワーク一般 | 教官 | 1名 |
| 5 | 電通大 情報システム学研究科 | 院生 | 1名 |
| | (合計) | | 10名 |

③中間段階のサンプル作成

続いて内容の作成にとりかかる。前述①で作成した項目は、全てが②の分類で分担を割り当てることができるわけでもなく、また分類できなかったとしても確実にこなしてくれるかどうか非常に怪しい。そこで項目の分類は特に行わず、執筆担当の項目の選択はもっぱら本人の意志に任せた。(各自の目標数は設定した。)

最初のうちは、元原稿を編集委員サイドでこまめに添削し、それを作って本人と検証しながら、原稿を見出し項目に対する原稿を執筆しても良かった。(時には執筆者の誤解などもあり、本人の合意なしに編集サイドで大幅に修正したものもあるが、その分は編集委員が全責任を持つこととして了解を得ている。)

このようにしてそれぞれに分担を決めて執筆を始め、最初原稿が集まってくると、執筆原稿の内容や説明の詳しさを、何をどこまで書くべきかについて、改めて指針をつくる必要があるとなってきた。それは以下に示す各事項である。

- 1 概念的、抽象的な単語(例:「互換性」「インタラクティブ」など)は

タになりがちであった。

それでも結局、ほぼ予定通りに仕事は進み、9月最後の週に講談社の編集担当者にソート済み、英文索引を付与した完成原稿を渡すことができた。

5. 最後に

今回の辞書の作成を通じて、教訓となったことは

- ① 就職戦線が長引きすぎて、人のアサインをしてもその通りに作業が進まなかったため、計画が大幅に乱れた。
- ② 短期間であったため、若手の執筆者の教育に時間が割けず、その結果、文体や語句の統一、禁句の訂正などによる編集委員の負担が増えた。
- ③ 電通大の学生は15倍以上異なる事柄を計算に入れていなかったこと。

以上の3つが考えられ、また、今回の辞書が最終的にはどれほどの出版部数出たのかは不明であるが、少なくとも今回の辞書作成を通じて、短期間に内容のある辞書を作成する秘訣をまとめることができた。その概略は以下のようである。

- ① 編集委員の中に、実際に辞書を執筆した経験のある人がいること。
- ② 担当する項目は義務でなく権利から生じるものであるから、できないと判断したら速やかに放棄（辞退）して、他の人に任せるようなシステムをつくること。
- ③ 編集委員が絶対的な権力を握る。（単語の割り振りや執筆原稿の訂正は、本人から許可を得なくとも実施できる。「大権威の〇〇氏の書いたのだから、直すことなどおそれ多くて」などというのでは、絶対にいい辞典は出来ない。）
- ④ 執筆メンバーの決定は編集委員の独断と専権で行える。
- ⑤ 表計算ソフトを使っての項目の整理。
- ⑥ 辞書利用者の対象がある程度明確に把握できていること。
- ⑦ マネージャーの役割は第一線の人やりやすい環境を作ることと肝に命じること。（会議が頻りにひらかれてばかりいるのでは、本職をやっている時間がなくなる。）

6. 実績比較

実績比較	情報マルチメディア辞典	化学大辞典 (共立出版)
編集委員	4人	100人
執筆者 (協力者除く)	6人	460人
項目数	2200項目	70000項目
一人当たり	220項目	130項目
作成日数	1/4年	10年
(生産性	65	: 1)