

## 情報検索支援のためのシソーラス管理システムの提案

1 T-6

徳田 圭世<sup>\*</sup>、西川 記史<sup>\*\*</sup>、間瀬 久雄<sup>\*\*</sup>、辻 洋<sup>\*\*</sup>、丹羽 芳樹<sup>\*\*\*</sup><sup>\*</sup>日立西部ソフトウェア（株）、<sup>\*\*</sup>（株）日立製作所 システム開発研究所、<sup>\*\*\*</sup>（株）日立製作所 基礎研究所

## 1. はじめに

近年、ネットワークを核とする情報流通が急激に発展し、多様な情報の中から目的とする情報を効率的に獲得することが特に重要になって来ている。情報検索や情報分類では、用語の体系をとりまとめたシソーラスが大きな役割を果たす。我々は今までに、シソーラスの陳腐化を防ぐ観点から複数の専門家によるグループ保守機能を具備するシソーラス管理システム<sup>[1]</sup>を開発してきた。

今回、これを発展させ、ネットワークからの利用に対応可能かつ、情報流通の支援も可能とする方式を検討した。本稿では、ネットワークを介して利用できる形態のシソーラス管理システムについて、その満たすべき要件および実現のためのアーキテクチャについて述べる。

## 2. ネットワーク対応に伴う要求の分析

我々は、シソーラスについて利用面・保守面の両面から具備すべき機能を分析し、それを実現できる管理システムを構築すべきと考える<sup>[2]</sup>。特に、インターネット環境での情報検索にも利用できる形態を実装することが必須であると考ええる。

ネットワーク情報検索への対応に対し、利用面からは、

- ・インターネットからのアクセス

- ・関連用語の充実

- ・利用形態に柔軟に対応可能な構造

の考慮が新たに必要となる。ネットワーク情報検索は、従来の検索システムと相対的に検索対象分野が広範囲である。シソーラス利用が効果を発揮するためには、検索目的に応じた関連用語の提示が必要であると考ええる。さらに、シソーラス全体はデータが普遍的でありある特定分野にのみ関心がある個人利用者には利用範囲が限定できないと使いづらいといった問題点がある。個人の利用形態に柔軟に対応できなければならない。

保守面からは、

- ・保守時のサービス停止による影響の最小化の考慮が必要である。

以上の分析を元に、以下ではシソーラス管理システムの機能要件と、それを実現するアーキテクチャについて述べる。

## 3. シソーラス管理システムの機能要件

上述の要求を満たす機能要件についてあげる。

- (1) インタネット汎用ブラウザからのアクセス  
シソーラス管理システムへのアクセスを汎用Webブラウザから実現する。
- (2) 共起データの格納を可能とする構造

従来の上位語・下位語・関連語以外に共起単語などの情報の格納を可能とするデータ構造を提供する。ここで、同じ文書に同時に現れる頻度の高い2つの単語を共起単語と呼ぶ。検索した情報の絞り込みに効果的な単語候補として、関連用語を提示できる。

Architecture of thesaurus management system for network era

Tamayo Tokuda\* Norifumi Nishikawa\*\*

Hisao Mase\*\* Hiroshi Tsuji\*\*

Yoshiki Niwa\*\*\*

\* Hitachi Seibu Software Co.Ltd

\*\* Systems Development Laboratory, Hitachi, Ltd.

\*\*\* Advanced Research Laboratory, Hitachi, Ltd.

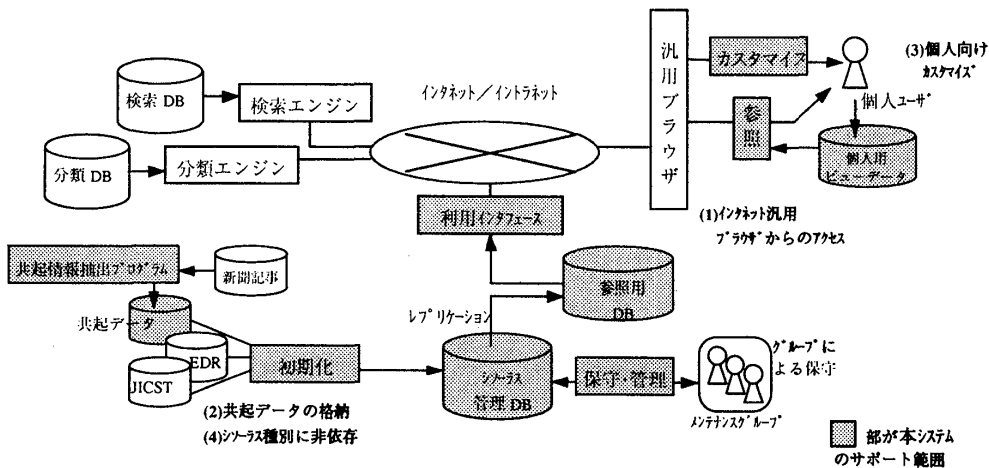


図1 アーキテクチャ

(3) 個人でカスタマイズを可能とする構造

利用目的により個人がシソーラスをカスタマイズ利用できる構造とする。

(4) シソーラス種別に依存しない構造

汎用利用を可能とするため、シソーラスの種類に依存しない構造を提供する。別種類のシソーラスを一元管理でき、シソーラスの拡張に容易に対応可能である。具体的には、単語間の関係で表現されるシソーラス<sup>[3][4]</sup>と、概念記述の関係と各概念と単語の対応関係で表現されるシソーラス<sup>[5]</sup>の双方を格納できる構造とする。

4. アーキテクチャ

図1に我々の提案するシソーラス管理システムを実現するアーキテクチャを示す。

利用面では、汎用ブラウザから個人ユーザに対し参照機能、カスタマイズ機能を提供する。カスタマイズを可能とするため、汎用ブラウザで対話的に行い、個人用ビューデータとして保存する。保守面では、複数の専門家によるグループ保守機能<sup>[1]</sup>、更に、シソーラス辞書および共起情報抽出プログラムにより抽出した共起データを初期化しシソーラス管理DBを構築する機能を設ける。また、保守時に検索サービスを停止させないためシソーラス管理DBと参照用DBを設け、参照用DBへの利用インタフェースを提供する。両DBの整合性はDBのレプリカ機能で維持する。

6. まとめ

我々は、ネットワーク経由で利用できるシソーラス管理システムの構築が必要と考え、具備すべき要件および実現のアーキテクチャを提案した。まず、ネットワーク対応に伴い発生する要求を分析した。特に、利用形態として最も多いと考えられる検索支援の要求を満足する汎用的シソーラス管理システムの機能をあげた。アーキテクチャとして、インターネット/イントラネットからの利用インタフェースを設け、利用面から必要とされる実装構造を示した。

これらを満足するシステムの実装により、特に情報検索においてシソーラスの有効活用が可能と考えられる。

7. 参考文献

[1] 西川、他：オブジェクト・データベースを用いたシソーラス保守システムの試作，情報処理学会第50回全国大会講演論文集(4)，PP33-34，1995  
 [2] 内藤衛亮他訳，シソーラス構築法，丸善株式会社，1989  
 [3] JICST 科学技術用語シソーラス 1993年版，日本科学技術情報センター，1993  
 [4] 日経シソーラス，日本経済新聞社，1994  
 [5] EDR 電子化辞書仕様説明書  
 (http://www.iiijnet.or.jp/edTG-All.tar.gz)