

1. はじめに

日本で1975年に竣工したスパン100m以上の、PC橋のうち、橋面積当り鋼材使用量が最大のものはどれか。その数値はいくつか。その橋に関する資料名を求む。

設計・工事の実務のなかで、上記のように各種の条件を組合せた結果選択された実績情報を必要とする場合は非常に多い。

このような場合、現行の情報検索システムで対処することを考えると、十進分類法あるいはシソーラスに頼って、プレストレスト・コンクリート橋に関する資料名を知り、一つ一つその内容を自からチェックしなければならない。

本来、情報検索システムの種類として、①データ検索(数値はいくつか) ②事項検索(該当するのはどれか) ③文献検索(資料名を含む)がある。しかし、現行の情報検索システムは、図書整理から発展した文献検索が中心であり、データ検索・事項検索は欠落している。したがって、文献検索自体も上例のように、各種の条件を組合せた検索要求に対しては、効果的な対応が困難である。

加えて、上例を遠隔地からの検索要求と仮定すれば、文献を介在しなくてもデータ・事項・名称などの情報を入手すれば満足できることがわかる。すなわち、設計・施工あるいは研究ごとに発生する図面・設計書・報告書の管理を考えるには、文献そのものの他、そこに記録された主要な情報(データ・事項)を対象とする汎用情報検索システムが必要である。

筆者は、建設部門における汎用情報検索システムを確立するために、IBMのプログラムSTAIRS (STorage And Information Retrieval System)を用いてアプローチを試み、実績データ・ベース、文献データ・ベースを作成したので、その内容を報告するとともに、オンライン対話式の検索を実例により紹介するものである。

2. STAIRSの選択

(1) 現行文献検索システムの問題点

現在情報検索として行われているのは、前述のように文献検索である。方式としてはカードソーターによるもの、コンピューターによるものがあるが、いずれも文献検索共通の問題点が指摘されている。

a) コード化やシソーラスが必要

現行方式では内容分類コードやシソーラスが前提となる。内容分類には情報内容・情報源・情報形態などによる分類形式があるが、ノイズ(不要な情報)を防ぐには、詳細かつ多角的な分類が不可欠である。

十進分類法あるいはそれに準ずる分類法については、固有名詞の取り扱いと、ある概念に到達する技術者の思考過程への天下りの制約とをあげておく。

また、シソーラスを用いると、ディスクリプタへの言い換えが検索途中で発生する。あるシソーラスでは、歩道橋→人道橋、沈埋トンネル→沈埋工法へ言い換えないと検索できない。

b) データ作成・質問作業に膨大な工数が必要

データ作成作業を見ると①文献収集→②内容分析→③キーワード集照合→④キーワード決定→⑤コード化→⑥カードパンチ となる。質問作業も①質問文書からはじまり、同様の手順を経て、情報読取りに到達する。またキーワード集の作成・更新作業も大きな負担となる。

これら作業は専門家による長時間を要する作業であり、情報検索システムのためのコストのうち80%を占めると言われる。したがって、このコストの問題を解決しないかぎり情報検索システムの普及はありえないであろう。

c) オンライン対話式の検索ができない

遠隔地からQ & Aを繰り返して、所要の情報に近づくという対話式検索は考慮されていない。

d) 数値のグループ化が必要

数値で表現される設計データを取り扱おうとすると、ある段階ごとにまとめたうえで、そのグループにコードを与えねばならず、ノイズの原因となる。

(2) STAIRSの特長

STAIRSはIBM技術情報センターにおける利用を当初の目標に、情報検索システムの集大成として開発されたプログラムであるが、現在社内外を問わず世界中で使用されている。1973年にOS/VS版、1976年にDOS/VS版が使用可能となった。主な特長を列挙すれば次のとおりである。

- 1) オンライン対話機能による検索
- 2) 自然文・自然語による検索（コード化不要）
- 3) 定型データ・非定型データの取り扱いが可能
- 4) データ・ベースの作成機能
- 5) 索引作成の自動化
- 6) 一群の単語に対し、同意語の指定が可能
- 7) 機密保護機能

したがって、STAIRS の設計思想は文献収集からデータ・ベースの作成、質問から情報読取りまでの中間過程をすべてコンピューターに負担させようとするものである。

3. データ・ベースの作成

STAIRS のデータ・ベース構成を次に示す。

データ・ベース……例・工事実績データ・ベース

ドキュメント……例・工事単位に作成される

パラグラフ……例・工事概要（項目に対応）

センテンス…点で区切られたワードの集合

ワード……単語または数字

データ・ベース作成の準備作業としては、データ

・ ペース名称の決定とパラグラフコード (3 ケタ)

- ・定型データフィールドの指定が必要となる。

今回作成したデータ・ベースの内容を次に掲げる。

1) 工場土木工事概要データ・ベース

土木学会誌 '74 増刊・'75 増刊掲載の工事概要のうち「工場土木」の内容（データ，概要，特色）を情報源とした。原資料の記入項目は必ずしも統一がとれていなかったが，自然文・自然語で入力できるので，単に書き写す程度のコーディングで済んだ。

項目として、工事名、所在地、発注者、施工者、工事費、工期、進捗率、記入日（この4項目は定型データ）のほか、工事概要、工法、材料、特徴などを設けた。記入例を図-1に示す。冒頭の8ケタがデータ・ベース名称に対応し、次の4ケタがドキュメント番号（0007）を表わす。続く3ケタがパラグラフコードであり、例えば100は工事名称、120は発注者のように対応している。パラグラフコードの指定により、同一ワードが別項目に存在する場合のノイズを解消できる。

2) 土木建築工事実績データ・ベース

業界紙上の主要受注発表をそのままデータとした。項目は工事名、工事内容、発注者、設計者、施工者、工期、工事費などで、建築の場合は面積、階数（地上・地下）を持つ。後4者が定型データである。この場合も順次パラグラフコードを定め、対応するデータを記入していった。

3) 土木建築技術文献データ・ベース

JICST (日本科学技術情報センター) 発行の科学技術文献磁気テープを利用した。記事項目にパラグラフコードを与え、変換プログラムによってカー

[illegible]

図-1 工場土木工事概要データ・ベースの記入例

ドのイメージに置き直した。内容は土木建築工學部門の内外雑誌・論文集などの文献を収録し、JICST選定のキーワードを追加している。

項目は、和文標題、欧文標題、著者名、出典名、使用言語、発行国、UDCコードおよびキーワードなどである。

変換プログラムの作成は短時間で可能であり、その利用によってデータ件数が膨大であるにもかかわらず、データ・ベース作成の工数は最小であった。

設計・工事の実績収集を日常業務の一環として行なう場合には、設計概要や工事報告書の用紙に、前もってパラグラフコードやカラム番号を印刷することが推奨される。記入者が情報検索システムを意に介さずに記入した内容が即入力データとなる。

4. オンライン対話式情報検索

STAIRSの検索対象はデータ・ベース中のすべての定型データ、非定型データである。データ相互間に優劣はなく、人間が検索の都度そのアプローチに応じてデータの階層構造を組立てることになる。冒頭の例を引けば、PC橋に到達する場合、PC→橋梁、橋梁→PCの2つのアプローチが可能である。

人間は思考に沿って質問を入力し、出力の結果を判断して次の質問に移ればよい。入出力は情報表示装置(キャラクタ・ディスプレイ)により行なう。

検索命令には非定形データ検索のSEARCH、定型データ検索のSELECT、アウトプット指示のBROWSEなどがある。SEARCHのオペレータとしては、AND、OR、NOTなどの論理演算子のほか、熟語指定ADJ、語尾変化指定(¥を使用：後述)などがある。

例えば、ドボク ガツカイ と2語で記入されているとき「土木学会」という熟語で検索するには、ドボク ADJ ガツカイ と指定する。

SELECTではEQ(=)、NE(≠)、大小関係(>、≥、<、≤)および範囲の指定が可能である。

以下に対話式検索の実例を紹介する。

土木建築工事実績データ・ベースを使用して、「工事費が1億円以上で、昭和51年に竣工した工場で、"東京〇〇"に関する事例」を検索するとした。繰

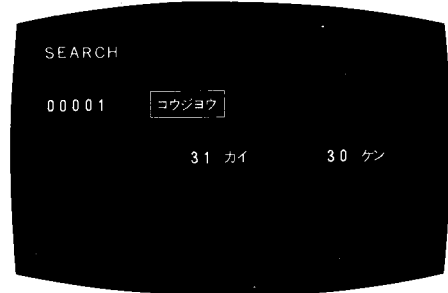
り返すが、手順に定石はなく、以下も一例に過ぎぬ。

1) コウジョウを入力し、「工場」の件数を探す。

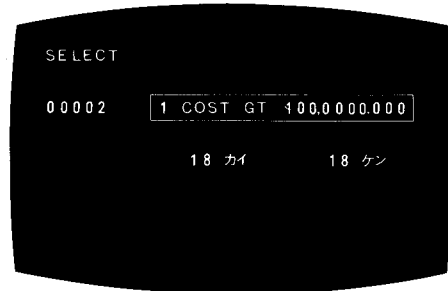
応答は30件のドキュメント(工事)があり、コウジョウという単語が31回存在することを示す。

はキーインを示すが、実画面には枠はない。

キーイン以外はすべて自動的に表示される。



2) 質問1で得られた30件のうち工事費(定型フィールド名COST)が1億円以上の工事件数を探す。



3) 質問2の結果からさらに竣工時期(定型フィールド名FINISH)の年が51で、月日は任意のものを選択する。この場合は語尾変化指定の¥記号を定型データに使用した。



4) 前の質問と関係なく、「東京〇〇」を求めるため トウキヨウ¥ と語尾変化指定を行なう。(ただし、ここで直ちに 3 AND トウキヨウ¥ とすればより速く必要情報に到達できる)

表示から、次に トウキヨウテ¥ とすれば、トウキヨウデンリヨクが一義的に指定できること

が判り、検索入力字数の節減に寄与できる。

SEARCH		
00004	トウキョウ	
トウキョウ	3 カイ	2 ケン
トウキョウカス	2 カイ	1 ケン
トウキョウコウ	1 カイ	1 ケン
トウキョウデンリョク	1 カイ	1 ケン
トウキョウト	16 カイ	13 ケン
RESULT	23 カイ	18 ケン

5) 質問3と質問4の結果をAND条件で結べば、所要の事例数が求められる。このときSORT命令を用いれば任意の定型データの大小順に揃えられる。

ドキュメント内容の画面表示はBROWSE 命令印刷はCOPY命令によるが、命令は専用鍵盤を1打鍵すればよい。印刷例を図-2.3に示す。

SEARCH		
00005	3 AND 4	
3	9 カイ	9 ケン
4	23 カイ	18 ケン
RESULT	4 カイ	3 ケン
SORT COST		

実例で見たように、データ・ベース中のすべてが検索対象となるので、技術文献データ・ベースの場合

合はタイトル中のワード、出典名、著者名などで検索でき、キーワードのみの検索と比較してはるかに高い使用効率を実現する。

5. おわりに ——適用業務の動向——

STAIRSを前提として、現在建設業界では、情報検索システムを単なる文献探しから脱皮させ、総合的な技術情報システムの主要部分とすべく検討を開始し、対象として下記のような適用業務に注目している。

- 法律検索 増大・複雑化する法的規制に対処するため関連法規の主要語・制限事項をデータとする。
- 特許管理 社内審査時間の短縮、重複開発の防止。
- 人事スキル検索 工事実績情報の利用のほか、自然文でなければ表現不可能な技術や趣味の把握。
- 社外情報の検索 JICSTその他市販磁気テープの効果的利用のほか、マスコミ情報の蓄積により業界の動向分析、材料カタログの利用がある。
- オンライン図面管理 工事実績情報に図面保管場所情報を持たせ、保有図面の有効利用をはかる。

参考文献：“仮設構造物総合エンジニアリングシステム”(1975 日本IBM・製造IS)
“情報検索の実演(建設版)”(同上)

AREA	=	1633
BASE	=	
COST	=	373,000,000
FINISH	=	51.03.
FLOOR	=	2
コウシ	タマカ"コセイソウコウシ"ヨウ シ"コウキンソ"ク シ"ヨキフコウ オスシヨリセツヒ" コウシ"ヨウ	
ハツチヨウ	トウキョウト	
セツグイ	シミス"カンセツツクマ	
セコウ	シミス"カンセツツクマJV	
シヨク"イチ	オオタ	
コウキ	51.03.	
タイヨウ	ソ"ウ セツヒ"ハ"ツト コウシ"	

図-2 土木建築工事実績データ・ベースのアウト・プット

COST	=	5,340,000,000
FINISH	=	49.12.10.
NOTE	=	49.12.10.
PARCENT	=	100
START	=	46.10.01.
コウシ"メイシヨウ	コウカンカイテイトンネル	
シヨク"イチ	カワサキシ ミス"イチヨウ オオキ"シマ	
ハツチヨウ	ニホンコウカン	
セコウ	タイセイケンセツ	
キニコウ	49.12.10.	
コウシ"カ"イヨウ	チンマイトンネル L=660M オフクミ セ"ンチヨウ 1540M オウフク4シヤセツフ"フリ ト"ウロ	
コウソ"ウフ"ツ	エレメント B=21 6M X H=7 05M X L=110M 6コ	
トウチヨウ	セツグイ タイセイケンセツ	

図-3 工場土木工事概要データ・ベースのアウト・プット(図-1と比較)