データベース構築促進および技術開発に関する報告書

# 自治体議事録のSGMLデータベース化と 情報検索ブラウザ機能開発

平成9年3月

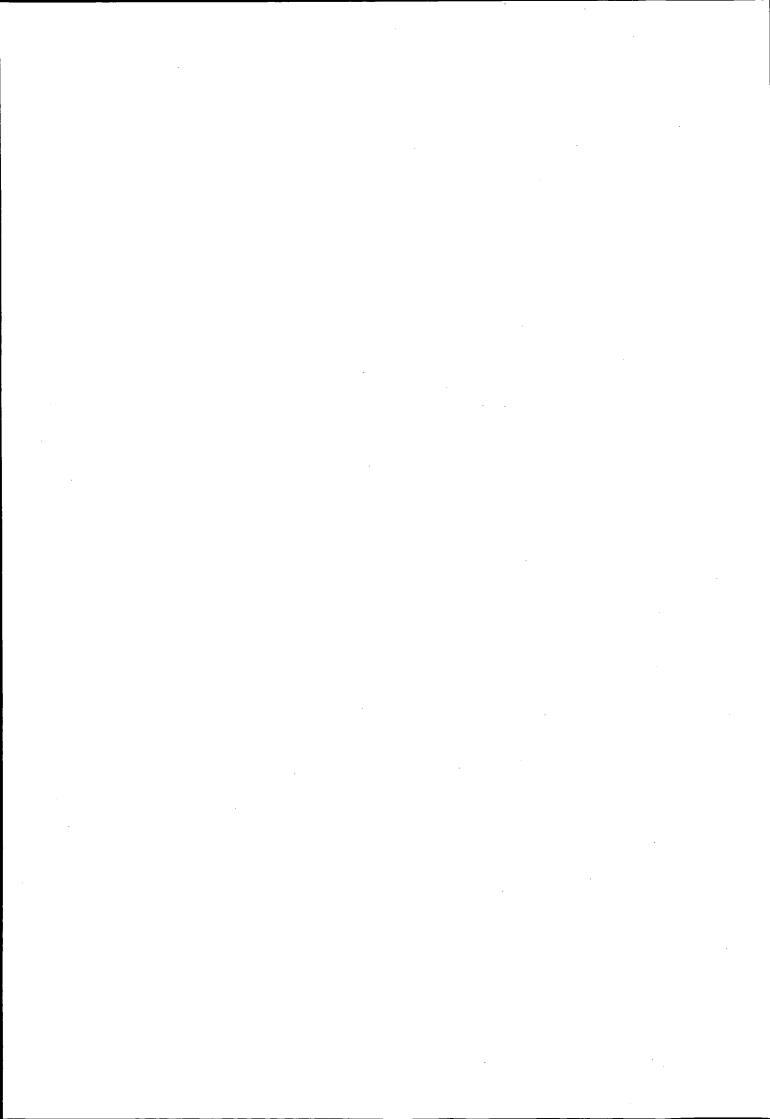
財団法人 データベース振興センター

委託先 株式会社会議録研究所



KEIRIN O

この事業は、競輪の補助金を受けて実施したものである。



	·	•		
		·		
			•	

データベースは、我が国の情報化の進展上、重要な役割を果たすものと期待されている。 今後、データベースの普及により、わが国において健全な高度情報化社会の形成が期待される。

さらに海外に対して提供可能なデータベースの整備は、国際的な情報化への貢献および自由な情報流通の確保の観点からも必要である。しかしながら、現在わが国で流通しているデータベースの中でわが国独自のものは1/3にしかすぎないのが現状であり、わが国独自のデータベースの構築およびデータベース関連技術の研究開発を強力に促進し、データベースの拡充を図る必要がある。

このような要請に応えるため、(財)データベース振興センターでは日本自転車振興会から機械工業振興資金の交付を受けて、データベースの構築および技術開発について民間企業、団体等に対して委託事業を実施している。委託事業の内容は、社会的、経済的、国際的に重要で、また地域および産業の発展の促進に寄与すると考えられているデータベースの構築とデータベース作成の効率化、流通の促進、利用の円滑化、容易化などに関係したソフトウェア技術・ハードウェア技術である。

本事業の促進に当たって、当財団に学識経験者の方々で構成されるデータベース構築・ 技術開発促進委員会 (委員長 東海大学教授 上條史彦氏) を設置している。

この「自治体議事録のSGMLデータベース化と情報検索ブラウザ機能開発」は平成8年度のデータベースの構築促進および技術開発促進事業として、当財団が株式会社会議録研究所に対して委託実施した課題の一つである。この成果が、データベースに興味をお持ちの方々や諸分野の皆様方のお役に立てば幸いである。

なお、平成8年度データベースの構築促進および技術開発促進事業で実施した課題は次表のとおりである。

平成9年3月

財団法人 データベース振興センター

### 平成8年度 データベース構築・技術開発促進委託課題

分 野	課 題 名	委 託 先
社 会	1 報道写真を中心とした商用デジタル写真タベース構築 2 WWWによる医薬情報全文検索データベーの構築と利用者Q&Aデータベースシスの構築 3 生活博物史データベース・プロトタイプの	ース 日本電子計算(株) テム
	4 信濃毎日新聞記事データベースの構築	信濃毎日新聞㈱
	5 患者投薬時に交付する服薬等指導文書の タベース構築	デー (株)YS企画
中小企業振興 地域活性化	6 自治体議事録のSGMLデータベース化 報検索ブラウザ機能開発	と情(株会議録研究所
	7 全国ベンチャー支援機関のネットワーク よる起業化支援データベースのプロトタ 構築	
	8 インターネットを用いたイベント情報サース   ス	ービ (社) 日本イベント産業振 興協会
	9 データベースクリアリングサーバのプロ イプの作成	トタ セントラル開発㈱ 情報図書館RUKIT
技術	10 中堅・中小企業向け顧客データベース利。 フトのプロトタイプ作成	用ソ (株)日経リサーチ
74U	11 写真データベースへの感性からの接近に る調査研究	関す (株)中日新聞社
	12 インターネット上の情報オブジェクトを した高信頼アプリケーション開発技術に る調査研究	

## 目 次

1		概要		1
	1	. 1 目的		1
		1. 1. 1	目的	1
		1. 1. 2	地方自治体における議会議事録の現状	2
		1. 1. 3	地方自治体における議会議事録及び文書情報化に関するアンケート調査結果	3
	1	. 2 地方	自治体における文書情報化のあり方	9
		1. 2. 1	行政改革と文書情報化	9
		1. 2. 2	自治体文書情報システム化のあり方	1 1
		1. 2. 3	議会議事録における情報システム化のあり方	12
2		文書情報	と技術の動向	1 4
	2		ML文書データベース	14
		2. 1. 1	SQLデータベースからSGMLデータベースへの変遷	14
		2. 1. 2	SGMLとは何か	15
		2. 1. 3	SGML文書データベースの有効性	1 7
	2	. 2 文書	<b>青報ネットワーク</b>	19
		2. 2. 1	自治体文書情報ネットワークの現状と今後	19
		2. 2. 2	インターネットを利用した情報公開との関連	2 0
		2. 2. 3	バーチャルガバメント化の可能性	2 1
3		会議録検	索システム	2 3
	3	. 1 会議	録検索システムの検証	2 3
		3. 1. 1	会議録検索システムの意義	2 3
		3. 1. 2	会議録検索システムの特徴と要目	2 4
		3. 1. 3	弊社会議録検索システムの検証「VOICES」	$2\ 4$
	3	. 2 SG	ML化の対応	28
		3. 2. 1	SGMLとの本質比較	28
		3. 2. 2	SGML化への課題検証	2 8

4 自治体議事	事録のSGMLデータベース	3 3
4. 1 自治(	体議事録の普遍性と特殊性	3 3
4.1.1	自治体議事録の現状と傾向・	3 3
4.1.2	自治体議事録SGML化の課題	3 5
4.2 自治(	本議事録DTD設計	3 7
4.2.1	自治体議事録の文書構造分析	3 7
4. 2. 2	自治体議事録のデータベース仕様文書(DTD)設計	3 9
4.2.3	自治体議事録SGMLデータベースの構築	48
4.2.4	情報検索ブラウザ	5 0
4. 2. 5	自治体議事録SGMLデータベースの運用	5 2
5 SGML	データベースの各種情報メディア出力と文書情報運用の可能性	5 5
5.1 印刷:	メディア	5 5
5. 1. 1	自治体議事録の自動組版化	5 5
5.2 イン	ターネット/WWW	5 6
5. 2. 1	SGML⇒HTML変換	5 6
5. 2. 2	インターネットによる自治体議事録情報の発信	5 6
5.3 CD-	-ROM	5 7
5. 3. 1	自治体議事録のCD-ROMによる展開	5 7
5.4 情報道	<b>重用</b>	5 8
5.4.1	文書管理機能の活用	5 8
5. 4. 2	他自治体議会議事録との横断的データ検索・活用	5 8
6 今後の展 <sup>室</sup>		60
6.1 今後日	目指す技術開発	60
6 1 1	SGMIデータベースによる自治体議事録滞用システムルの展開	6.0

#### 1 概要

#### 1. 1目的

#### 1.1.1 目的

ペーパーレス化とインターネットの活用をきっかけとして、全世界的に情報ネットワークによって業務を進めるヴァーチャル組織化が進展してきた。そんな中で「CALS (Continuous Acquisition and Life-cycle Support/Commerce At Light Speed)」が提唱され、様々な分野においてCALS導入の検討が進められている。その中のキーワードが文書情報管理体系である「SGML (Standard Generalized Mark-up Language)」である。

これまで文書データのデータ交換基準は、MS-DOSテキストファイル等プレーンなデータファイルでしかなかった。つまりその文書データの中身・構造を体系化してデータを管理する手法はなく、データベース的発想での文書データ交換基準は存在していなかった。そこで生まれてきたものがこの「SGML」であり、文書データ構造定義まで含まれたこのSGMLファイルにより、異機種間のデータ交換や文書データのデータベース化・マルチメディア展開を容易にすることとなり得た。

SGMLはこのCALSの進展とともにその実用化が進み、航空機・軍事分野の作業マニュアルや学会論文など大量文書の管理運用にその利点があり、米国・欧州を中心に多分野にわたって、その導入は進展している。

国内においてもこのCALS・SGML推進のために、民間分野では通産省が今年度の実証実験として26業種のプロジェクトに補助金を交付、またCALS技術組合の活動によって、CALSの業界標準づくりが進展しており、SGML活用のための業界標準DTD (Document Type Definition:文書型定義)の設計も急がれている状況にある。

これに対して中央省庁や地方自治体など行政機関においては、平成6年12月に閣議決定された「行政情報化推進基本計画」に基づき、行政改革・情報公開に対応すべく、平成9年度内検討完了を目標に作業が進められている。(平成9年1月スタートした中央省庁ネットワークである「霞が関WAN」は、そのためのインフラ整備の第1ステップである。)

しかし現状は文書管理においてSGMLを導入するという方向性は示されながらも、各 行政文書情報分野毎のDTD検討などは具体化しておらず、そのDTD等の提唱が待たれ ている状況にある。 弊社はこれまで各自治体・団体等の議会議事録の速記・データおこし等の編集から、議事録印刷や議事録検索システムの提供を通じて、議会議事録の分野においては蓄積されたノウハウがあり、行政機関のSGML化の流れに対して、議会議事録文書データのSGML化のいて、今回その研究開発を行い、議会議事録分野の標準DTDやその文書管理運用に関して提唱することを目的と考える。

#### 1.1.2 地方自治体における議会議事録の現状

議会議事録とは、地方自治法等の法律に基づき、議会運営内容のありのままを記録として残して公表するという目的から存在するものであり、ほとんどの議会においてその調製を行っている。(全国の都道府県および市町村等の地方自治体おける総議会数:2571 議会の中で、全文記録によって調製しているところは2446議会(95.1%)である。)

また従来、議会議事録とは速記者による速記法による調製が主体であったが、現在はテープ録音した結果を逐語的に全文を議事録調製するものが主体へと変化してきた。

表1-1. 地方自治体における議事録調製方法の状況

速記法による議事録調製	59議会	2.41%
テープ録音による議事録調製	2387議会	97.59%

これらの結果は、行政機関のリストラ策の結果の一環であると考えられ、特殊技術者である速記者を自治体自らが配さず、議会議事録のコンテンツに関する速記・翻訳作業の外部委託が進展していることと考えられる。現に外部委託の比率は次のような結果が示されている。

表1-2. テープ録音による議事録調製議会(2387議会)における内部対応および外部委託の比率

テープ速記および反訳作業の内部対応	844議会	35.36%
テープ速記および反訳作業の外部委託	1543議会	64.64%

また制作された議事録について、自治体の議会事務局に保管されることはもちろんであるが、対外的に配布されている状況は次のとおりである。

表1-3. 議事録の対外配布の状況

対外配布する	議会議員等の関係者に配布している	867議会	33.72%
	地区公民館等へ住民閲覧用にも配布し	199議会	7.74%
	ている		
しない	議会議員等の関係者に配布していない	1704議会	66.28%

このような結果から、現状の議会議事録とは、記録を残して議会事務局によって保管をするという目的で捉えられている自治体が2/3を占めているという結果が見受けられる。しかし、自治体の情報公開ニーズの高まりの中で、市民オンブズマン活動などが各地で発足していく状況において、議会議事録の目的も記録に残すものから、一般に公表するものへと変化することが求められてくると考えられる。その一環として一部の議会では、議会として独立したインターネットホームページを開設している自治体も生まれはじめ(埼玉県議会・北海道標茶町議会など)、マルチメディア時代の流れの中で、多様な形態での議会議事録の公表が進んでいくことと考えられる。

また議事録のコンテンツ制作に関する外部委託がこれだけの比率で進展していくということは、議会議事録の文書情報について、標準となるその文書管理体系を示しておかなくては、各々の議会毎に異なった形態・体系での議事録制作が進んでしまい、議会議事録のあり方が不明確となってしまう恐れもある。他の行政文書が統一した形態での文書管理が進展していく中で、議会議事録のみこれとは逆行した形態で進展しまうことは、行政情報化の目的に反することであり、今こそ議会議事録の標準仕様を誰かが提唱しなくてはならない状況にある。

これらの背景を検討すると、議会議事録の印刷+マルチメディアによる情報発信、議会 議事録の文書構造標準体系の整備とともに、SGMLによる文書管理体系整備が間違いの ない方向性であり、今回の事業の重要性が改めて確認されたこととなる。

[※データ出典:「地方議会人」(平成8年3月号)]

1.1.3 地方自治体における議会議事録および文書情報化に関するアンケート調査 結果

今回の事業を進めるにあたって、次の点を検証するために全国の地方自治体の議会事務 局および企画・情報システム部門に対してアンケート調査を弊社では実施した。

- ●「地方自治体における行政情報化がどのように進展しているのか。」
- ●「地方自治体における行政情報化の捉え方と興味はどの程度なのか。」
- ●「マルチメディア化に対してどの程度着手しているのか。」
- ●「SGML化についてどのように考えているのか。」
- (A) 議会事務局に対するアンケート調査結果(協力:117自治体)
  - ①議会事務局においてインターネットに接続していますか?

表1-4. インターネット接続状況

接続している	30自治体	25.64%
現在検討している	57自治体	48.72%
まったく考えていない	30自治体	25.64%

#### ②インターネットホームページ作成についてどのように考えていますか?

表1-5. ホームページ作成状況

ホームページはすでに完成している	7 自治体	5.98%
これから作成する予定	23自治体	19.66%
庁内職員で作成する予定	9 自治体	7.69%
まったく考えていない	78自治体	66.69%

#### ③議会情報をインターネットに掲載することを考えていますか?

表1-6. 議会情報公開の意識状況

議会の仕組みや活動内容について	0 自治体	0.00%
4ヶ月に1回程度の更新で掲載したい		
将来的に議会情報を掲載していきたい	22自治体	18.80%
そこまで考えていない	9 5 自治体	81.20%

#### ④公文書についてその文書データを保存していますか?

表1-7. 文書の保存状況

文書管理体系を整備して全庁的に保管している	11自治体	9.40%
各セクション毎に保管は任せている	100自治体	85.47%
今後文書データ保存を指示していく考えである	6 自治体	5.13%

#### ⑤SGMLという文書管理体系を知っていますか?

表1-8. SGMLの認識状況

まったく知らない	101自治体	86.32%
研修会を通じて話は聞いている	14自治体	11.97%
通達文書により知らされている	2 自治体	1.71%

議会事務局における情報化進展の状況を図る指針として、インターネットの活用方法と 文書データの保存にみる文書管理体系の整備状況について検証を行った。。

まずインターネット活用状況は、住民からの情報公開のニーズの高まりもあって、①の結果にあるように75%弱の自治体において、インターネット接続について実施もしくは検討中という回答が得られたが、②のホームページの開設については約2/3が開設を考えていないという消極的な回答であった。これはホームページを開設するためのコンテンツ編集作業が手間であると考え、新たな原稿作成やデータ作成等の多くの作業工数が発生し、それらを予算化しての内部制作もしくは外部委託が難しいという判断の結果ではないかと推測される。また③の議会情報のホームページ開設についても、すぐに対応したいという回答が0件であったことから、定期的な情報更新の手間が発生するという懸念からの結論であると推測される。

そして文書管理体系についても、④の結果にあるように各セクション・担当者に任せて あるという回答が85%強であることから、全庁的な文書管理基準はほとんどの自治体に おいて定められていないと考えられる。これにより多くの自治体において、部門間の文書 交換や担当者交代時の引継業務などが非効率的状況であると推測される。

これらの結果は、⑤のSGMLの知識に関する回答が意味するところであり、86%強の自治体でSGMLがまったく知られていない状況であることから、SGML導入による

文書管理の効率化、議会議事録データの様々な情報メディアへの展開などそのメリットが 自治体の議会事務局の現場において理解されていない現状では、これらアンケート各項目 の結果もいたしかたないことであると考えられる。

#### (B) 企画・情報システム部門に対するアンケート調査結果(協力:77自治体)

#### ①庁内LAN構築はなされていますか?

表1-9. LAN構築状況

LAN構築をされている	25自治体	32.47%
LAN構築を計画している	38自治体	49.35%
LAN構築は考えていない	14自治体	18.18%

#### ②イントラネット構築についてどのように考えていますか?

表1-10. イントラネット構築状況

現在構築検討中である	13自治体	16.88%
構築はまだ考えていない	22自治体	28.57%
勉強中である	42自治体	54.55%

#### ③インターネットに接続していますか?

表1-11. インターネット接続状況

接続している	37自治体	48.05%
現在検討している	3 9 自治体	50.65%
まったく考えていない	1 自治体	1.30%

#### ④インターネットホームページ作成についてどのように考えていますか?

表1-12. インターネット接続状況

ホームページはすでに完成している	35自治体	45.45%
業者委託して作成する予定	8自治体	10.39%
庁内職員で作成する予定	8 自治体	10.39%
まったく考えていない	1自治体	1.30%
必要性はあると考えている	25自治体	32.47%

#### ⑤公文書についてその文書データを保存していますか?

表1-13.公文書の保存状況

文書管理体系を整備して全庁的に保管している	3 自治体	3.90%
各セクション毎に保管は任せている	68自治体	88.31%
今後文書データ保存を指示していく考えである	6 自治体	7.79%

#### ⑥SGMLという文書管理体系を知っていますか?

表1-13.公文書の保存状況

まったく知らない	59自治体	76.62%
研修会を通じて話は聞いている	18自治体	23.38%
通達文書により知らされている	0 自治体	0.00%

自治体の企画部門・情報システム部門における情報化進展の状況を図る指針として、議会事務局において確認した2点以外に庁内LAN(Local Area Network)およびイントラネットの構築状況についてを判断基準に加えてみた。

まず庁内LANおよびイントラネットの構築状況については、①の結果にもあるように情報インフラの整備が各自治体において進展していることから、庁内LANは約1/3の自治体において整備済みのようである。しかし②の結果を見るとLANは構築されながらも、イントラネット整備については明確に整備済みであるとの回答は得ることができず、現在検討中という回答も17%弱という状況である。

本来LANの整備とイントラネットによる情報システムによる業務運用は同様のステップで進行していくべきものであるが、自治体の現状において情報インフラ整備のみ先行し、運用体制の整備がそれに追いついていないという状況にあると推測される。

これは⑤の文書管理体系に関する回答にも表れており、全庁的な文書管理体系が整備されている自治体が4%弱という結果から、イントラネットを整備するために必要となる要件が未整備であることから、その実施が進行していないことと考えられる。

イントラネットを整備していくためには、次の2点がポイントとなると考えられる。1つはこの文書管理体系整備である。文書のフォーマットを標準化して、庁内において文書交換が容易になり、その文書が他の部門においても自ら作成した文書と同様に活用できることが目標となる。またもう1点は業務進行の規則・基準策定である。ある業務はどのような担当者を通じて最終決裁に至るのか、誰が責任者で誰の内容確認を得る必要があるのかという決まり事を整備することである。

自治体はその組織的な業務運営から後者については、規則基準が民間企業以上にきちんと整備されている。しかし前者の文書管理についてはパソコンによる電子データ化は進んではいるものの、文書作成基準が明確ではないことから、各々が自分なりの文書作成を行ってしまっていると推測される。そこでこの文書作成基準の策定ということからも、文書構造を明確に定義するSGMLによる文書データ保存の有効性はあると考えられる。

#### 1.2 地方自治体における文書情報化のあり方

#### 1.2.1 行政改革と文書情報化

ここ数年、行政機関の効率化とスリム化による行政改革の実現が、ニュージーランドやアメリカなどの先行事例と対比して、日本国内においてもその具体的な方策が議論され始めている。これに対して政府では、平成6年12月25日の閣議決定によって「行政情報化推進基本計画」を提唱し、中央省庁および全国の地方自治体に対してその早期実現を求める指針を示した。

この基本計画では「行政の情報化とは、行政の事務・事業および組織に通ずるシステムを改革するための重要な手段として位置づける。」と定義されており、国民等に対する行政サービスの高度化と行政事務の高度化・効率化を情報・通信システム基盤の整備によって実現するため、情報化推進のための組織的・人的基盤の充実、予算および調達の改善を行うという内容のものである。

これまでの行政改革の方向性とは、行政機関内部の効率化のみを議論していれば良い状況であったが、インターネットによる情報公開や申請業務の電子化・オンライン化など情報化インフラの発展整備に伴って、我々国民等とのやりとりまでも考慮に加えなければ、総合的な情報化に至らないことから、今回の行政情報化推進基本計画の中にもこの点が盛り込まれている。

具体的な「行政情報推進基本計画」の内容は次のような点にある。

#### (A) 国民等に対する行政サービスの高度化

- ①情報提供サービス等の高度化
  - ・ページものの情報提供の便宜性
  - ・白書等のCD-ROM等による民間への提供
  - ・情報提供窓口の一元化等のための制度の見直し
- ②行政手続等の迅速性・高度化
  - ・行政手続について、事案審査等行政機関内部の事務処理を効率化する情報システム の整備に合わせ、申請・届出・報告・相談等の電子化・オンライン化を推進。電子 的縦覧・閲覧の推進

- ・各種許認可等の事務手続について、「紙」による事務処理を代替する技術の進展等 を踏まえつつ、情報・通信技術を活用した申請・届出等を行えるようにするための 各種許認可等に係る諸規程類の見直しの推進
- ・ワンストップサービス、24時間役所を目指す調査研究

#### (B) 行政事務の高度化・効率化

- ①意思伝達、事務処理の高度化
  - ・行政事務処理に伴って発生する情報の広範な電子化を行うとともに、一般行政事務 における文書の作成・保管・伝達等の事務処理について、情報システム化を推進
  - ・各省庁文書管理規則について、施行文書の公印の省略、電子的決裁方式の導入、セキュリティの確保の検討を進めつつ、逐次見直しを進める。
  - ・官庁会計事務、人事・給与関係事務等に係る文書等、「紙」による保管・提出が義 務づけられている文書について、これら事務の情報システム化に合わせて、その電 子化を図ることとし、これに必要な制度面の見直しを進める。
  - ・省庁間の情報の共有化のための標準化、省庁間電子文書交換システムの整備等

#### ②行政情報の総合利用

- ・データベースの省庁間利用
- ・省際行政事務に係る情報について、関係省庁間の調整を図りつつ、共同利用できる データベースとしての整備の推進

#### (C) 情報・通信システム基盤の整備

- ①情報・通信システムと執務環境の高度化
  - ・パソコン等1人1台環境の整備
  - ・施設のインテリジェント化など執務環境の整備
  - ・分散処理、マルチメディアなど新技術の普及、積極的活用およびオープンシステム 化の推進
  - ・情報の総合利用を行うための標準化

#### ②通信ネットワークの高度化

- ・LAN/WAN(Wide Area Network)の整備
- ・行政機関のネットワーク化と地方公共団体、民間等との情報交換手段の確保

これらの点を踏まえ平成8年6月18日には、行政情報システム各省庁連絡会議幹事会において「白書等データベースの統一的な仕様」が了承され、行政情報の中で「白書・年次報告書・調査研究報告書およびこれに類する文書」を対象に、文書データのSGML化の提唱が行われ、これに関する標準DTDがまとめられ示された。

これにより行政情報の文書データは他分野に関しても、体制がされ次第SGML化を推進していくことと考えられ、今回の事業テーマである「議会議事録」に関しても、その文書構造が定型化できることから、SGML文書化の対象となるとも判断される。

またこれらの情報化事業は各自治体において早急に対応しなければならないことが「高度情報通信社会推進本部 制度見直し作業部会報告書(平成8年6月)」からみてとることができる。

この報告によると、書類の電子データによる保存および申告・申請手続の電子化・ペーパーレス化の双方ともに、「平成8年7月から速やかに着手し、民間事業者等からの要望の強いもの等優先順位を考慮しつつ、できる限り速やかに実施するものとし、特に周到な検討を要する場合も、平成9年度末までには検討を完了し、できる限り速やかに所要の措置(法改正が必要なものについては、法案を提出)をとる。」とまとめられている。

これによると平成10年度以降、具体的な電子文書による行政事務の推進が始まることと考えられることから、これに合わせて今回議会議事録のSGML化に関する提唱は、その有効性が高いことと考えられる。

以上のような点から、中央省庁および地方自治体ともに、行政改革、つまり業務のリ・エンジニアリングを実施するために、SGML化による行政情報化は間違いのない方向性であると考えられる。

#### 1.2.2 自治体文書情報システム化のあり方

自治体をはじめとした文書情報システム構築および運用のコンセプトは「One Sourse Multi Use」である。自治体内部においてもせっかくパソコン等の情報化機器を導入しなが ら、活かしきれていない場面が多々見受けられる。

例えば住民から申請された申請案件について、ワープロで制作された申請書を「紙」にて受領して、それを担当課員がその内容を見ながら、あらためて自らのパソコンにその申請内容を入力し直している。また庁内LANで配信された文書をすべての人が端末を持っていないことから、「紙」へ印字出力し、それを課内担当者すべてに配布している。

これらの点は情報システム化によって期待する利点である、ペーパーレス化・データ共 有化に反することとなるはずである。

そこで自治体文書情報システムの構築にあたっては、各自治体において定型化された汎用情報管理システムに業務を合わせて利用していくのではなく、自らの業務分析を充分に行い、担当者は誰であって、決裁権は誰にあって、その文書がどのような経路を辿っていくのか、最終的にその文書がどのような形になっていくのかなどを知った上でのシステム設計でなければ本当に活用される情報システムにはなり得ないはずである。

これらの作業を行うことによって、その情報の中身である各種文書情報の構造、つまり SGML・DTDも明確になってくるはずである。これにより SGML導入が容易になってくるはずなのである。また自治体業務とは、各自治体ともに同様の業務形態をとっていることと考えられることから、DTDも共通利用できるケースが多いことと考えられる。そこで自治体文書情報の SGML化の指針となる分野毎のDTD設計をある機関が全国の自治体業務を分析して行うことによって、自治体文書の SGML化が容易に早く進展することと考えられる。今回の議会議事録 DTD設計もその一助となればと弊社では考えているが、他分野に関しても同様に進展することを望むものである。

SGMLとは文書共有化が目的である。もし自治体間において文書交換が必要となる際に、各々の自治体においてDTDが異なるケースが発生した場合、これでは情報交換が難しく、全国規模での行政情報化というテーマもまったく意味をなさなくなってしまう。この点からも自治体文書DTDの統一仕様整備は重要であると考えられる。

#### 1.2.3 議会議事録における情報システム化のあり方

議会議事録の現状は前述のとおり、テープ録音した結果の反訳作業によって議事録という印刷物が生まれ、その議事録という印刷物によって、保管・管理・検索などが行われているのが一般的である。

積極的な自治体においては、議事録のコンテンツデータを印刷物とともに活用し、議事録用文書検索システムを導入して、そのシステムの全文検索エンジンによって、キーワード検索を行って、以前の議事録内容から必要な部位を抽出するという作業にそれを活用している。(弊社においても会議録検索システム「VOICES」を全国の自治体に対して提案し、導入実績も多数あり。)

しかしこれらのシステム化で想定されているものは、議会事務局による議事録の検索という点に的が絞られており、現在叫ばれている情報公開に対応した一般市民が議事録内容を活用できるシステム化やインターネットによる議事録データの公開などの面については、検討が行われていない状況にある。

一般市民の議会議事録検索の場面を想定すると、例えばある地域開発事業があった場合に、そのテーマに関する議会における議論の進展状況を見るのと同時に、その事業に関する今年度の予算状況や開発事業の事業概要など、会計部門や担当部門に所蔵される他の関連文書もこれにリンクさせて情報検索したいというニーズが高いはずである。これらはインターネットの技術によって、自治体側においてこのような情報公開と関連情報のリンクを貼っていれば実現は可能なはずである。

このような中で他の一般文書がSGML化されていく流れの中で、議会議事録もこれに合わせていかなくてはならないことからも、行政情報化の進展によるSGML化は議会議事録分野においても同様に考えなければならないのである。

またSGMLとは文書データの文書構造を定義することから、議事進行の見出し部分など、文書構造で定義された各項目をキーワードとして検索抽出することも可能であり、これによって、議事録の概要版の制作なども容易になることと考えられる。これまで議事録の内容を検討しながら、議会広報を議会事務局において編集していたと考えられるが、DTDの設計によって、議事録全文情報から議事要旨概要を自動的に抽出し、議会広報を自動編集できる仕組みづくりも可能となるはずである。またこれはデータの再入力なく、SGML to HTML(Hyper Text Mark-up Language)変換によって、インターネットホームページへ自動アップロードを行い、議事録ホームページとして市民に公開することも可能になるはずである。

議事録作成の際に電子データ化がなされていることから、前述の「One Sourse Multi Use」というコンセプトに従い、議会議事録の多面にわたる活用を検討し、議会事務局内部のクローズドシステムではなく、自治体全体および市民に対してのオープンシステム化をインターネットおよびイントラネットを活用して志向していくことが重要であり、そのために文書情報のSGML化は有効であると考えられる。

#### 2 文書情報化技術の動向

#### 2. 1 SGML文書データベース

#### 2.1.1 SQLデータベースからSGMLデータベースへの変遷

これまでの文書データ管理技術は、各アプリケーション毎の閉鎖的環境の中で進展してきた。各々のアプリケーションがその機能によって勝負し、自らのアプリケーションへの囲い込みを目的にファイル形式を異なる形態として相互互換が不可能な状況にあった。これは各文書関連アプリケーションの創世期には、自らのシェア獲得とマーケット寡占化のために仕方のないことではあるのだが、これで苦労をしたのがユーザーサイドであったことはいうまでもない。

そこで様々なアプリケーション間のデータ変換プログラムが作られてきたわけだが、アプリケーション毎にレイアウト機能が異なることから、100%完全なデータ変換が実現できるものはほとんどなく、イメージどおりのデータ互換は不可能であった。

そこで多くのケースとして、MS-DOSテキストファイルによるデータ交換を行っていたが、これはあくまでレイアウト情報を持たないプレーンテキストであり、データ交換は容易であるが、レイアウトやデータ構造を定義する内容はそのファイルに含まれていないことから、多くの課題を持っている。

その後RTF(Rich Text Format)ファイルのような書式付き文書ファイルによって異なるアプリケーション間でのデータ交換が可能なファイルも生まれたのだが、ワープロ専用機との互換の問題など、これにおいても完全なものとはいえなかった。

そこでソフト固有のレイアウト情報を省き、文書構造を示す識別情報を「タグ」という印で示し、文書構造が理解できるプレーンテキストとして生まれたのが「SGML」である。データ交換を行う際にも、そのSGML文書を作成する際の約束ごとである「SGML宣言」とその文書の文書構造を示す「DTD」をそのSGML文書ファイル「文書インスタンス」に付しておけば、互いにその文書データを取り扱うことが可能となるのである。確かに文書を「紙」に印刷するためのレイアウト情報が付加されていないため、紙に印刷するためには別の仕組みが必要とはなるが、このデータ形式によって文書情報データベース、印刷用データ、マルチメディア出力用データともに、同一のファイル形式でストックしていくことができることから、データ加工のアプリケーションにより様々な文書データの展開が可能となり、非常に自由度の高いファイル形式であると考えられる。

これによりこれまでSQLデータベースでは難しいと考えられていた文書データのデータベース化もこのSGMLによって容易なものとなった。SQLデータベースもデータ項目の非固定長化や並列SQL処理による高速検索性などにより、そのアプリケーション上においては文書データも高速に取り扱うことも可能ではあるが、アプリケーション間データ互換やエディター等により文書データ形式としてデータ閲覧を行う際など、課題点は残されている。

情報システムのオープン化の流れの中で、容易に文書データ交換が可能であり、世界的な文書管理体系として位置づけられている「SGML」は今後の文書管理の標準として、 CALSの流れも受けて、国内においても標準となり得るものと考えられる。

#### 2.1.2 SGMLとは何か?

SGML文書とは「SGML宣言」「DTD」「文書インスタンス」の3つの要素から構成されている。簡単に説明するとSGML宣言とはどのようなSGML文書であるかを指示する宣言文であり、DTDとは文書構造を示しておりタグ付けの規則になっており、文書インスタンスは文書構造を示すタグ付き文書ファイルである。本項では各々の概要を説明し、SGMLの概要についてその解説を行う。

#### (A) SGML宣言

SGML宣言とは、そのSGML文書が書かれている内容の基準について示したものであり、次の6項目で構成されている。

- ① CHARSET:使用する文字の指定
- ② CAPACITY:文書の容量指定
  - ③ SCOPE: SYNTAX (具象構文) の適用され範囲
  - ④ SYNTAX: 文書中の構文 (SGML文書内において意味のあるキーワードや記号) の指定
  - ⑤ FEATURES: SGML機構の使用指定
  - ⑥ APPINFO:アプリケーション固有情報

これらによって使用されるコード種や要素名などの最大文字数、タグマーク(ex.< >)、SGMLが提供する機構をどの程度活用して文書インスタンスが書かれているのかなどが

示されている。つまり文書交換を行う際に使用しているコード系などを認識させることが このSGML宣言の目的である。

なおISOでは、標準SGML宣言として「ISO8879」の中で定義されたものも 存在する。多くの場合はこの標準的なSGML宣言を利用している。

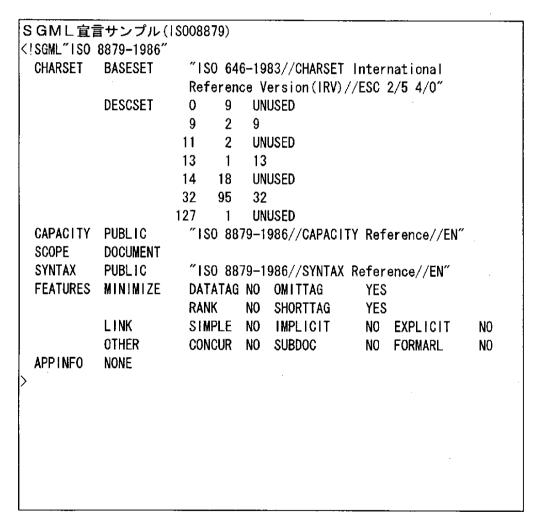


図2-1. SGML宣言

#### (B) DTD (Document Type Definition/文書型定義)

DTDによって定義されるものが文書構造である。文書インスタンスの中で構成される 文書タグを定義し、その文書がどのような文書構造によって構成されるものであるのかが これによって示されるのである。

そのためDTDとは、異なる業界や文書内容によって、各々のDTDが定義されている。 SGMLの目的が文書互換にあることから、目的とする文書分野や業界において標準DT Dが定義されている場合にはそれを採用するべきであるし、DTDが存在しない文書分野 においては、自ら文書構造を分析して、DTD設計を行わなくてはならない。

#### (C) 文書インスタンス

一般的にSGML文書と呼んだ場合、この文書インスタンスを示している。つまりDT Dによって定義された文書型に従って作成されたタグ付き文書が文書インスタンス(実現値)である。

#### 2.1.3 SGML文書データベースの有効性

これまでの文書情報の管理とは「紙」による管理が主流であった。機器の操作マニュアル等も単純な機器であれば、少量の紙の取扱説明書で解説が可能ではあろうが、複雑で高機能な機器では、その説明書も多くのページ数が必要となり、紙という形態から保管スペースが必要となってきてしまう。

例えば戦闘機や戦車など複雑な軍事機器においては、その操作および整備マニュアルをもしも紙によって保管しようとした場合、その総量は数トン単位と大量になってしまい、その紙の重さによって戦闘機は離陸不能、戦車は走行不能になってしまうとまで言われている。SGMLもこのような背景もあって、軍事分野から発展を見せていったのだが、これらマニュアル情報をSGMLで体系化された文書形式で、CD-ROMにストックしておけば、その検索性も優れており、軽量で必要なときに必要な情報のみを検索できるという面でその有効性は非常に高いと考えられる。

またSGMLファイルであれば、必要があれば必要な部分を組版アプリケーションによって組版・印刷処理をして紙のマニュアルとして出力することも可能であり、必要なときに、必要な部分を、必要なメディア形態で出力が可能というオンデマンド対応が可能となるのである。

また前述のとおりSGML文書とは、特定のアプリケーションに規定されない文書構造のみが示されたプレーンテキストである。よって様々なアプリケーション上においてその文書データを取り扱うことができる。

文書データベースによってデータベース管理を行うことも、印刷用組版アプリケーションを通じての印刷出力など多くの可能性が考えられる。その中でも特に注目すべき点はマルチメディアへの情報表現である。

インターネット・WWW(World Wide Web)で注目を集めるハイパーテキスト体系である「HTML」があるが、これはSGMLを汎用的定義に置き換えた派生定義であり、SGML to HTMLは下方変換であることから容易にその変換は可能である。この点からWWWのコンテンツ制作に関しても、基となる文書データをSGMLにおいて構築しておけば、ハイパーテキストによるマルチメディア展開も容易であるが言えるのである。

#### 2.2 文書情報ネットワーク

#### 2.2.1 自治体文書情報ネットワークの現状と今後

前述の地方自治体における情報化に関するアンケート結果でもわかるとおり、自治体文書情報ネットワークの現状は、庁内LANの整備など周辺技術である情報化インフラの整備は進展を見せているものの、本質的な業務効率化のためのコンテンツ活用という中身の部分までは着手できていない状況にある。

このように国内においては暗中模索の状況にあるが、情報化先進国アメリカに眼を向けてみると、情報化により成功した事例は幾つも見ることができる。

例えばカリフォルニア・マーシド郡の「福祉業務情報化」がその1つの事例である。同 郡では福祉業務において次のような問題点を抱えていた。

- 資格審査の複雑化
- ・膨大なペーパーワーク (業務総量の8割を占める)
- ・職員の志気の低下、スキルレベルの低下
  - →これらにより住民へのサービス低下が発生

この点の解決のため、同郡では最新の情報管理技術を全面的・統合的に採用し、行政内部業務の抜本的な効率化と高度化を実現した。この情報管理技術の特徴は自らの内部的な業務効率化のみにこだわらず、対住民サービスの向上を主眼に、AI(人工知能)を積極的に活用し、事務のスキル・判断要素までをシステム化し、ペーパーワークの改善によって、職員が本来発揮すべき知的生産性向上に寄与したものと評価される。

日本においては、全体の業務の流れを分析せず、一部分のみのシステム化によって効率化に至ったといっているケースが多く見受けられるが、それは本当の意味の効率化ではなく、住民もしくは他部門へ必ずそのしわ寄せが至っていることと考えられる。しかしこのマーシド郡のケースは業務全体を業務分析し、そのすべての業務をシステム化してしまった点にそのポイントはある。これにより審査に必要な質問項目の特定、該当規則の照会、資格の有無の決定、給付額の計算、必要書類の作成などが、システムとの対話形式でオペレーションができ、福祉業務が申請者と職員協力のもと自動的に行われるに至ったのである。

これにより申請者は、職員からの質問が終了すると自動的に給付額が算定され、即時的にそれを知ることができ、大幅な時間短縮に至った。行政内部においても、必要知識が組織全体で共有化できることや、新入職員の研修期間の大幅短縮、給付額計算の精度向上、時間短縮による人件費削減、そして何よりこれまでペーパーワーク主体の業務から住民サービス主体の業務へと業務内容が180°転換し、サービス向上につながったのである。

もちろんこれらを支える文書管理体系もSGMLであったし、SGMLによる文書共有 化と抜本的な行政事務の情報化改革とは行政サービスの向上につながり、公務員がよりク リエイティブなプランナー的要素の業務までその活動の幅を広げることが可能となると考 えられる。

このようなことから考えると、今回の事業テーマは議会議事録のデータベース化に絞ったSGML検証であったが、今後は議会運営に関わる全体業務の情報システム化について検討を行うべきであると考える。 (議員経歴管理・議会日程管理・議員報酬計算などの議会事務局の業務全般にわたる検証)

#### 2.2.2 インターネットを利用した情報公開との関連

自治体における情報公開の気運が市民オンブズマンなどの活動によって高まり、情報公開制度による行政情報公開が各自治体において行われている。しかし現状の情報公開とは、文書により必要とする情報・文書を明記の上で申請し、その審査および文書検索(所在部署を確認して探し出す)に1週間から10日間という長期間がかかっている状況にある。

これは自治体内部の文書が一元的・体系的に管理されていないために発生する弊害であり、住民サービスの向上に務めるのであれば、この点は早急に対処しなくてはならない課題である。

インターネットの普及によって、求める情報を自らが探して(検索)、即時的に情報閲覧を行うことに慣れてきている。この流れに対して、現状の自治体情報公開のタイムラグというのはこの流れに反していることと考えられる。そのためにもSGMLによる行政文書の抜本的な情報管理体系の整備が重要なのである。

またもちろんSGMLによる文書管理ということは、前述のとおりSGML to HTML によってハイパーテキスト化することも容易であり、インターネットによる情報公開の実現も可能となる。

また担当部門における各事業の事業概要、議会での議論内容、それに関わる予算案など、 関連する情報に関してハイパーリンクを貼ることによって、住民はもっと自治体の行政活動に近づくことができ、本当の意味での住民参加型行政活動の実現にもつながることと考えられ、これは結果として行政事務の効率化にもつながるはずであるし、住民の声がインターネットを通じて自治体にも寄せられることから、より良い行政活動が遂行できるはずである。

#### 2. 2. 3 バーチャルガバメント化の可能性

地方自治体の活動は、中央省庁の系列に従って良い点悪い点ともに考えられるが、縦割り化が進んでいる。その弊害として複数の分野に関わる省際行政事務に関しては、住民としてその担当窓口が不明確であったり、複数の窓口への問い合わせを行わなければその結果を得ることができないなど、そこには課題点が潜んでいる。

そこでインターネットによる情報化技術を活用とともに、自治体内部の行政業務分析によるあらゆる住民の問い合わせに対する対応窓口の明確化とそれに関する文書情報の体系化管理の実現によって、インターネットによる申請・届出等行政手続の電子化、行政業務問い合わせの電子化など窓口業務がわざわざ市役所を訪ねなくとも可能となる電子政府・電子役所(バーチャルガバメント)も実現は可能なのである。

住民の行政機関に対する不満は、求める案件についてどこが窓口となるのかわからず、 その相談・申請・届出に戸惑ってしまうという点にある。行政サイドにおいてあらゆる問 い合わせに対して体系化して整理し、どの内容の問い合わせはどこの部門において応える べきかという決まり事を整備しておけば、住民としてはインターネットを通じてあたかも 1つの窓口に問い合わせるだけで、すべての問い合わせに応えてくれるように感じとれる はずである。

またこれらは住民のみにメリットがあるのみではなく、前述のマーシド郡の福祉業務情報化の事例と同様に、行政内部の効率化にも寄与し、クリエイティブワークに職員が注力できるようになると考えられる。

もちろんこれらを実現するためにも文書情報のSGML化は必要条件であり、住民からの問い合わせに対する対応についても、SGML文書化を進める業務分析の過程において明確化できるものであると考えられることから、行政改革を進めていくことが、業務分析となり、それにより文書管理体系が明確化することによりSGML化が自然に進行するも

のと考えられる。

また議会議事録のコンテンツをこのバーチャルガバメントの中で活用することを考えると、議会において議論されているテーマ・内容が他の行政文書のどの部分にどのように関わってくるのかという検証を行い、そのハイパーリンクを定義づけることにあると考えられる。そこでバーチャルガバメントにおける議会事務局の役割とは、行政活動全般を見回し、今議論されている議題がどの部門に関わり、どの事業計画に関わり、どの予算分野に関わるのかなどを検証するという、行政機関全体のスタッフ部門として位置づけが変遷していくのではないかと考えられる。

自治体におけるインターネットの活用とは、とかくホームページの開設が目標のように 考えられている。しかしインターネットの有効性とは双方向コミュニケーションである。 ホームページを開設するだけであれば、紙メディアが単に画面上に代替えされたのみの意 味でしかない。そこでこの有効性を充分に活用するためにも、住民とのコミュニケーショ ンも考慮した双方向性での自治体におけるインターネットの活用が今こそ求められている のである。

- 3 会議録検索システム
- 3.1 会議録検索システムの検証
- 3.1.1 会議録検索システムの意義

会議録検索システムとは、各地方自治体の議会で討議される内容を記録した議事録をデータベース化し、その内容を検索・参照するシステムである。そもそも自治体行政とは首長を始めとする執行部が行政実務を担当遂行し、議会がこれを審議・承認する関係により成り立っている。このしくみの中で議会の議事内容である会議録(議事録)は、自治体行政の結果と方向を集約する会議の記録として必然的に生み出されてくるものである。しかし現在から未来へいたる自治体行政と議会運営の直接的な担い手たる議員や自治体職員にとって会議録とは、自らの活動の結果出てくる単なる副産物、としてあるだけでは決してとどまらない。

議会の会議録は法律によりこれを記録し永年保存する義務が定められていることからも明らかであるとおり、きわめて公式性の高いものであるが、この公式性の高さを議会という公的な場で記録されたものだからだとして平板に解釈するだけでは不足である。議会における過去の討議内容は、常に何らかの形で自治体行政の現在に反映されていなければならず、また未来へと進行する自治体行政の方向を決定的に規定する重みを持っていなければならないからこそ公式性が高いのであり、記録として保存の重要性があるのである。すなわち行政執行部は過去の会議録の内容を現在に投影する形で仕事をしなければならないという責任を負っており、議員は過去の会議録の内容に基づいて政治活動を組み立てなければならない立場にあるわけである。つまり行政関係者にとって、過去の記録である会議録とは、現在の自らの立脚点であり死活点である、といって言い過ぎではないであろう。

こうしたことから会議録の検索が行政にとってきわめて重要な問題であることが明らかであろう。現実にほとんどの自治体で、製本された会議録がぼろぼろになるまで参照されているのが実態である。会議録の検索をシステム化することは行政の効率化・円滑化をし、 生産性を高める上できわめて重要な意味をもつものである。

#### 3.1.2 会議録検索システムの特徴と要目

会議録検索システムはおおむね以下の要件を満たす必要がある。

#### (1).会議・発言の管理

検索結果がいついかなる会議で誰が発言したものであるか、を知る必要がある。すなわ ち、ある発言に着目したときに、その発言が、

いつ(開催年月日)のもので、

何の会議(定例会、臨時会、各個別委員会、協議会、・・・)の発言で、

誰の(発言者)発言なのか、

が何らかの形で分からなければならない。

#### (2).全文検索機能

会議録は会話を記録したものであるため、その内容は基本的に曖昧である。したがって数値や定型的なデータでの検索とは異なり、検索者は述べられているであろうテーマを想定することで検索のキーとなる言葉を特定し、その言葉の存在する位置を対象となる会議録の中から探し出す方法、すなわち全文検索の手法を用いなければならない。

#### 3.1.3 弊社会議録検索システムの検証「VOICES」

弊社では数年前から全国に先駆けて会議録検索を研究し、そのシステム化に取り組んできた。そして会議録検索の汎用パッケージとして商品名「VOICES」という製品を発売し、幾多の納入実績を積んできている。ここでは弊社会議録検索システム「VOICES/Win」の構造を検討する。

図に示すとおり、「VOICES/Win」のアプリケーション構造は基本的に、以下の3つのモジュールからなる。

#### (1) アプリケーションモジュール

ユーザインターフェース、印刷、モジュール制御などの機能を持つ。

#### (2).全文検索エンジンモジュール

文字成分分析表を用いた検索エンジンであり、学習ファイルを参照する。また、検索の際には、与えられたキーワードから同義語・類義語を派生させ、それらの派生語も元のキーワードとともに検索キーとするシソーラス展開機能も持つ。

#### (3).文書(発言)管理モジュール

このモジュールが会議録検索というアプリケーションとしてのさまざまな機能を提供する元になっている。すなわち、発言者ジャンプ、日程ジャンプ、文書間のジャンプ、個人の全発言の抽出や、議長次第書の作成などなど、会議録検索システムとしての「VOIC ES/Win」の豊富な機能は、この文書管理によるところが大きい。

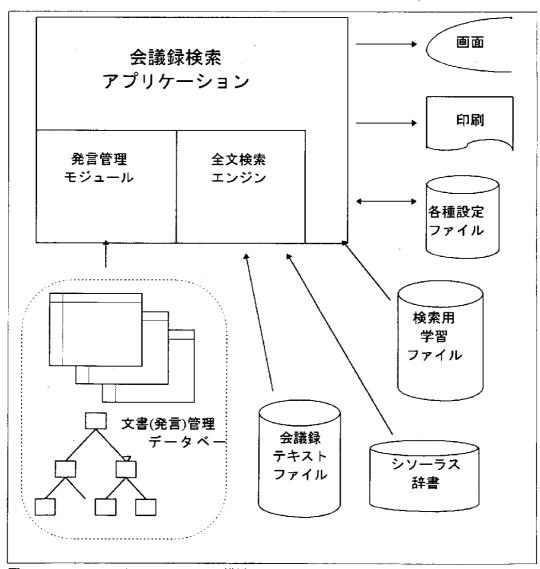


図3-1. VOICES/Win のシステム構造

そこでさらにこの「VOICES/Win」の文書管理を詳細に検証してみる。

「VOICES/Win」に会議録データをセットアップするために、データの加工を 行っているが、その内容は以下のとおりである。

- (1).会議名を入力する。
- (2).ページ番号を入力する。
- (3).発言者の名前の様式を統一する。
- (4).各発言に簡易タグを付ける。
- (5).改行コードの追加と削除。

上記の中の(4).の詳細は次のとおり。

△会議録本文中に記載されている日程に「△」文字を付ける。

- ○議長・副議長、その他の議事進行者の発言の冒頭に「○」文字を付ける。
- ◆質問者の発言の冒頭に「◆」文字を付ける。
- ◎答弁者の発言の冒頭に「◎」文字を付ける。

これらの簡易タグは、会議録という文書の構造を大まかにではあるが反映したものになっている。後に「自治体議事録DTD設計」の章で詳しく述べるが、地方議会の会議録は、大きく以下の構造を有している。

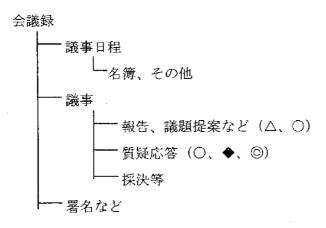


図3-3. 自治体議事録の概要構造

このなかで討議の実質的な内容は、「議事」とりわけ「質疑応答」にあり、また量的に も会議録の大部分のページがこの「議事」しかも「質疑応答」で形成されているのが通例 である。上述の簡易タグは主としてこの部分に対して付けられるものである。

この作業はテキストデータすなわち文字列を、その意味的構造において把握し分類する作業である(分割ではない)。この作業により会議録のテキストデータは、文字コードが羅列しているだけの原始的状態から初めて抜け出し、そこに構造的で意味的な付加価値が与えられることになる。これは原初的とはいえハイパーテキストの思想を持ったものであり、少々おおげさな言い方をすれば「会議録マークドアップランゲージ」とも言えるものである。

ここでその作業が文字列の分割ではないことは、重要な意味を持つものと思われる。なぜならば、分割を行う場合、分割という行為には暗黙で1つの意味しか与えることができず、さらに詳細な情報を必要とする場合は、他の何らかの方法によって与えることが余儀なくされるからである。たとえば、分割されたものの順序や分割されたものの詳細な分類は、ディレクトリ構造やファイル名によりするなどしなくてはならない。しかし、ディレクトリ構造やファイル名によりするなどしなくてはならない。しかし、ディレクトリ構造やファイル名によって情報を付加する方法論は、現在時点では有効であるかもしれないが、コンピュータのオペレーティングシステムに依存した方法論であり、本質的には永久保存が必要な会議録の性質になじんだものとは言い難いのである。

また、(5) は、文書中で明らかに改行が必要と思われる個所に改行を入れ、逆にそれ以外の個所の改行コードを除去することにより、パラグラフを明確化するための作業となっている。

#### 3.2 SGML化の対応

#### 3. 2. 1 SGMLとの本質比較

前節(3.1.3)で詳しく見たように、弊社会議録検索システム「VOICES/Win」は会議録データを単なるテキストデータとしてだけでなく、そこにその意味的把握を含めた簡易タグを付加することで本格的な検索機能を実現している。そしてその簡易タグは、検索機能の実現を目的とするものではありながらも、その実は文書構造に根ざした内容で付けられているものである。

したがって、単純に現状のいくつかの簡易タグに対応するSGMLタグを定めたDTDを定義すれば、その時点で一応SGML化はほとんど終了することになる。残すところは簡易タグをSGMLタグに変換する過程のみである。

こうした観点に立てば、弊社会議録検索システム「VOICES/Win」は、本質的にはもう既にSGMLに対応していると言ってもいいのかもしれないようにも思われる。

ただしかし、それはあくまでも本質のみを論じているに過ぎず、現実には幾多の問題があるだろうことは、「VOICES/Win」にはわずか数個の簡易タグしか扱っていないことをとってみても容易に想像がつくことである。

そこでまず、このタグの数に関して考察をすすめることにする。

#### 3.2.2 SGML化への課題検証

#### (1).データの入力

「VOICES/Win」で定義している簡易タグは前述のとおり、 $\triangle$  (日程)、 $\bigcirc$  (議長)、 $\spadesuit$  (質問)、 $\bigcirc$  (答弁)の4つであり、質疑応答を中心にした議事そのものしかタグ付けの対象にしていない。しかし文書としての会議録には、その他の要素もたくさんある。会議録のどういった要素をどう捉えていくかについては、DTDの設計として後述するとして、いずれにしろ、SGML化するためには、現在簡易タグがカバーする部分を含め、会議録全体をくまなく構造的に分解しなければならない。

本論本文部分は量的には多いが構造は単純、それに前後する部分は量的には少ないにも 関わらず構造は複雑、という他の文書でも一般にいえるであろうことが自治体会議録にも 同様に当てはまる。すなわち、データ量としては圧倒的に多い議事部分は構造的には単純 であり、むしろ議事以外の部分にさまざまな要素が多く存在し、形式的にも複雑な構造を 持っている。とすれば当然のことながらSGML化しようとすればタグの数は飛躍的に増 えることになるのは避けられない。

しかし、タグの数が増えるということはまず何よりもデータの入力が煩雑になるという ことを意味する。データの入力には次のいくつかの方法が考えられる。

a. SGMLのDTDで定義されたタグをそのまま入力する。

現時点ではもっとも現実的な方法であるが、次の問題点がある。

- ・入力を行う作業者がよくDTDを理解し、SGMLタグを記憶する必要があり、 教育に時間がかかる。
- ・入力されたものにミスがないかどうかデバッグ作業が必須となる。
- b. SGML入力エディタで入力する。

詳しくは後述するが、次の問題点がある。

- ・入力作業者はSGMLタグを記憶する必要はないが、ある程度DTDを理解していなければならない。
- ・エディタの操作が作業に占める割合が大きく、便利ではあるが生産性は必ずしも 高いとはいえない。
- c. SGMLコンバートの可能なWYSIWYGツールで入力する。 直接SGMLコンバートが可能なツールは現時点では存在しない。
- d. SGMLのDTDで定義されたタグに対応する簡易タグを定義し、簡易タグで入力を行い、入力されたデータをコンバータにて、本当のSGML文書に変換する。

この方法はSGMLタグを直接入力する場合と比較すると、打鍵数が少ないなどで 生産性はある程度高いが、以下の問題点がある。

- ・a. と同様に入力作業者はよくDTDを理解している必要がある。SGMLタグを記憶する必要はないが簡易タグを記憶する必要がある。
- ・入力されたもののデバッグ作業はある程度軽減されるが、デバッグ作業の度にコ ンバータを通す必要があるため、その分オーバーヘッドになる。

以上、たとえどのような方法を取っても一長一短であり、いずれにしろタグ数の増大は データ入力作業の生産性、すなわちコストに非常に大きな悪影響をもたらす。この点の解 決に関しては4.2.5に後述する。

# (2).表示出力系

次に表示・印刷に関してであるが、ディスプレイ画面への表示と紙への印刷では大きく事情が異なることをまず念頭に置かなければならない。また、紙への印刷も会議録中の一部分、すなわち画面に表示された部分のみを印刷する場合と、会議録全体を印刷する場合ではやはり事情は大きく異なる。画面に評した内容を印刷するだけなら、極端にいえば画面のハードコピーの延長線上で考えて差し支えないであろう。

画面へ表示出力する場合、多くの場合ユーザはその表示形式にさほど拘らないため、製本された会議録と形式が大きく異なっていようともさほど問題にはならない。むしろ、画面には画面の特性を生かした表示形式を取る方が好ましい場合が多い。現状の会議録検索システムで行っている方法で、SGML化された会議録データを画面表示するには、次のいくつかの方法が考えられる。

a. SGMLデータをクライアント側で処理して画面表示する。

この方法ではクライアントにSGMLデータを現状の会議録検索システムの画面 表示モジュールが理解できる形式に変換するコンバータを付け加えてやればいい。

b. SGMLデータをクライアント側の要求する形式にサーバ側で変換処理して画面表示する。

この方法ではクライアントにSGMLデータを現状の会議録検索システムの画面 表示モジュールが理解できる形式に変換するコンバータを付け加えてやればいい。

どちらが良いかはシステムの設計によって変化する。より具体的に言えば、クライアント側で動くアプリケーションが何であると想定するかによって決まってくる。しかし、クライアント側のアプリケーションが何であろうと動作できるようにするならば、b. のタイプが好いということになろう。

ただ、いずれにしろ今まで会議録検索システムでは、簡易タグ (△、○、◆、◎) は画面上に表示しているのだが、SGMLタグのついた文章はタグの存在が煩わしくたいへん読みにくいものなので、SGMLデータを表示する場合は、SGMLタグはユーザの目に触れてはいけないということになる。

# (2).印刷出力系

既に述べたが、文書をSGML化する目的は、「One Sourse Multi Use」に寄るところが大である。そう考えるとせっかくSGML化するのであるから、会議録SGMLデータから必要になればいつでも正式な会議録の版下印刷ができて当然、ということになろう。しかし、これには現状では相当困難な問題が存在する。しかし、差し当たり会議録検索システムは会議録の版下印刷機能を持たないので、これに関しては第4章で後述することにする。

# (3).全文検索エンジンとの融和性

会議録検索システムは、その機能を主として全文検索エンジンから生み出している。全 文検索エンジンとは読んで字のごとく、文章を全面的に検索するためのエンジン、すなわ ちテキストデータを対象とした検索ソフトウェアである。一方、SGMLデータは完全に テキストデータであるから、そのまま全文検索エンジンの検索対象になりうる。その結果 も疑う余地なく保証されるところのものである。

敢えて問題点を探れば、検索の対象として、ユーザがSGMLタグと同じ、もしくは部分的に一致する綴りの検索語を与えたときどのような動きをするか、ということが挙げられよう。ただし、これには2つの解釈が成立する。

その第一の解釈は、ユーザはSGMLタグの存在を知らないという前提に立って論ずる解釈である。仮にユーザがたまたま偶然にアルファベットで "kaigiroku" という検索語を与えて検索をかけたくなったとし、かたやSGMLタグの中にその綴りと同じ綴りのタグ <kaigiroku> があったと仮定すると、そのタグは検索結果としてヒットしてくることになるわけだが、SGMLタグのことなど知りもしないユーザにとっては何やら意味不明の検索結果としてしか理解不能である事になってしまう、ということである。実際の会議録検索の現場でこのようなアルファベットの綴りで検索をかける例があるとは思えないのだが、しかし実際やってみればこれは現実に起こりうる。それが現実味の無い物でも、理論上起こりうる不具合は解決しなくてはならないのがコンピュータの世界である。

第2の解釈とは、SGML文書の利点はそれ自体が一つのデータベースを形成しているのであるから、そのインデックスであるSGMLタグが検索できなければ逆に意味がない、との解釈である。これもまた当然のことである。

しかし、SGMLデータベースを構築するに当たっては、一見矛盾するかにみえるこの 2つの問題を解決するのは必ずしも全文検索エンジンに依らなくともよい。データベース の構成上でこの問題を解決するというつもりで研究・設計を行った。

# (4).文書管理データベースとの融和性

文書管理はSGMLの真価を発揮するにはうってつけの土俵である。現在の文書管理データベースはSGMLデータベースの構築に当たっては、その具体的な実現方法は大きく変わることになる。しかし、結論的にはデータベースに入るときにデータがSGMLの文書インスタンスになっているかいないかだけの問題であり、文書管理自体の機能も原理的な仕組みも基本的には何ら変わりはない。すなわち、そのデータがいつの、どんな会議の何日目のデータであり、その中で誰と誰がどういう順で発言したのか、が管理できており、その結果、あるデータを特定したときにそのデータを説明するそれらの情報が得られ、また逆にそれらの条件を指定したときにデータを特定することができるようになっていればよいので、会議録という文書をツリー状に分類する構造は基本的に変わらない。

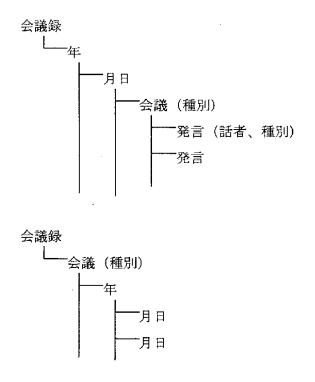


図3-4. 自治体議事録の文書管理構造

- 4. 自治体議事録のSGMLデータベース
- 4.1 自治体議事録の普遍性と特殊性
- 4.1.1 自治体議事録の現状と傾向

# (1).現状

自治体議事録がおよそどんなものであるかについては既に述べたが、自治体議事録をSGML化するという視点からその具体的状況を掴んでおかなくてはならない。

弊社においては東日本を主として全国で数百自治体の会議録の調整を行ってきた実績があり、その実例に多く接しているという点において、自治体議事録の現状と傾向を最も良く知る立場にあると自負するものがある。

まず、議事録に記録される内容であるが、これについては議会という法により定められた組織である性格を反映してか、ほとんど基本的にはどの自治体も差はみられない。自治体毎に大きく違いの出てくる部分を敢えてあげるとすれば、議事録に関連する文書であるところの、議案書や請願書・陳情書、あるいは報告書・意見書などの扱いについてである。これらの文書を積極的に会議録に添付する自治体もあれば、まったく添付しない自治体もあり、議案の審議状況一覧は添付・掲載するが、議案書は添付しない自治体もある、といった違いである。しかし、これらの文書はあくまで関連資料なのであり、議事の記録としての会議録とは本来別個のものである。したがって、やはり基本的には、会議録に記載される事項は全国的に共通しているといえるだろう。

記載される事柄が同じだとすれば、次の問題はその載せ方である。まず事柄の記載順序であるが、順序に関しても大まかな枠組みは変わらない。ただし、前述の関連文書の掲載のされかたはその順序においても当然変化が多く見られる。

次に、文書体裁であるが、これは各自治体においてまったくまちまちである。用紙サイズはB5判、A4判のいづれもあり、書き下し方向も縦書き、横書きいずれも存在する。 段組しないところもあれば、2段組みの自治体もある。字詰めやフォントサイズ、タイトルの修飾、罫線の種類などまで考慮すれば、数え切れないほどのレイアウトスタイルが存在する。

以上の現状を、総じていえば、文書構造はほぼ等しいが、その出力形態はさまざまである、ということになろう。SGML化を考える場合、前半の「文書構造はほぼ等しい」ことはたいへん都合がよい。なぜならば、文書構造を定義するDTDが設計しやすく、また

そのDTDは汎用的なものになりうるからである。だが後半の「出力形態はさまざま」であることはたいへん不都合である。なぜならば、会議録汎用DTDに基づいて作成されることになる会議録SGMLデータは、会議録を製本する能力を持たないことにより、現状のユーザニーズを満足させないからである。

# (2).傾向

ではこうした現状にある自治体議事録は、どのように変化してゆく傾向を持っているだろうか。また、SGML化するに際して不都合な側面である「出力形態はさまざま」な現状は改善される見込みがあるのであろうか。

まず関連文書の掲載範囲、およびその出現順序と製本の際の組み版レイアウトに関して はその仕様の統一はほとんど絶望的ともいえる難点を有している。

それは、会議録の掲載内容や調製仕様の変更は基本的には当該自治体の議会がこれを承認しなければならない、という点にある。議会の承認が必要だというのは、もちろん法的にそういう規制があるわけではないが、そうかといって例えば会議録を調製する担当部所である議会事務局が勝手に掲載内容や調製仕様の変更を行うわけにはいかないという事情を持つ自治体が多いことによる。それぞれの自治体の会議録は、過去さまざまな変遷を経るなかで、それなりの理由や根拠を獲得しつつ現在の姿になっているのであろうから、それらを全て捨て去るのは容易なことではないことは想像に難くない。

だが一方、例えば紙サイズは漸次A4判に移行しつつある。この点では数年後の近い将来には、ほぼ全国的にA4判に統一されるであろうと思われるし、また従来の主流であった縦書きを横書きに変更する自治体が漸増してきているのもここ数年の傾向である。もちろんこのような変更を施す際にも多くの自治体では議会の承認をもって改めるわけだが、このような機会に組み版上の細かなレイアウトについても統一した書式に切り替えることを考えれば、事はスムーズに運ぶであろう。また、会議録調製の民間委託を始めるに際して書式を変更する自治体も多い。

# 4.1.2 自治体議事録SGML化の課題

以下に自治体議事録をSGML化するにあたっての課題をまとめる。

# (1). S G M L 化の効果

本来文書をSGML化することにより期待される効果としては、

- ア. 特定のコンピュータアーキテクチュアやアプリケーションに依存しないことによる保存性の向上、
- イ、文書を構造的に整理することによる検索性の向上、
- ウ. 文書を部品化することによる再利用性の向上、

がある。これらのうち、ア、の保存、イ、の検索については大いに効果が期待されるが、ウ、の再利用については、会議録という文書の性質上あまり、そっくりそのままでの再利用は期待されるところがない。しかし、議事内容の要約に関しては必ずしもその限りではなく、現状でも議会だよりなどに議事内容が要約されたりしている。また、自治体議事録に関連する多くの文書、請願書・陳情書、議案書、議決文書、そしてそれらの集約としての条例文などが等しくSGML化され、それらの文書の間で一連の繋がりが持てるようになったとき、今以上に議事内容の参照と要約が求められてくる可能性は十分にあるといえよう。

#### (2).製本調製との関係

会議録は当然のことながら製本されなければならないため、SGML化を行う際に、体裁のとれた印刷出力を強く意識しなくてはならない。しかも問題なのは、その印刷の組み版レイアウトはただ単に体裁が取れていればよいというものではなく、先に見たとおり、現状では各自治体により固有のスタイルを実現しなければならないことである。もちろん、SGMLの文書インスタンスを出力するアプリケーションに、どのような体裁で出力させるかは各自治体でそれぞれに決めてもよいわけであるが、現状現実の会議録の印刷体裁を一切変えないという前提で、SGML化を考えるとするならば、文書の構造だけを定義したDTDに基づくSGMLタグではあまりにも情報が少なすぎる。すなわち、そのためには印刷体裁の上では意味があるが、文書構造上は何の意味も持たない、レイアウトのためだけのSGMLタグを多く定義しなければならないことになってしまうわけである。出力レイアウトのためのタグ付けなどはSGMLの本来意図するところとは無縁な行為であるばかりでなく、タグが増えることによりコストもまた増加するということを意味し、これ

はSGML化にとって大きな障壁となってしまう。

しかし、本当に問題なのは、コンピュータとネットワークのテクノロジーが進歩した結果、文書という存在の意義が根本から変化しているということなのである。すなわち、これまでは文書という存在の価値を高めるためには、紙の上での読みやすさ美しさを追求する意外に方法がなかったわけだが、現在では文書主体を紙の上の存在から電子的な存在へと変えることで紙という場所的物理的制約から解放され、交換性の高さ検索性の高さを追求することこそが文書の存在価値を高める方法になったということである。したがって自治体議事録も印刷という方法論から一度離れ、データとして(SGML文書として)その価値を飛躍的に高めた後で、再び印刷というメディアに立ち戻る、という発想の転換がぜひとも必要なのである。

したがって各自治体での固有の組み版レイアウトは、自治体議事録の標準DTDを確立 した後、その文書インスタンスを出力する際に読みやすさ美しさを追求する、というプロ セスを踏んでいってこそ真に意義のあるものになるのである。

# (3).図表

会議録においてはあまり図は出現しないが、表は往々にして出現する。しかし、SGM Lで定義している表組みのタグは大変複雑であり、入力は困難である。この点に関しては 特にデータ入力のための優秀なツールが必要になる部分であろう。

また、図表に関しては、会議録そのものではなく、その関連文書にて多く用いられるものである。したがって、請願書、陳情書、意見書などがそれぞれでSGML化されることが、理想的な道筋であろうと思われる。

# 4. 2 自治体議事録DTD設計

# 4.2.1 自治体議事録の文書構造分析

4.1.2 で自治体議事録をSGML化するにあたっての課題をまとめたが、結論的には印刷出力を意識してしまうと現状の自治体議事録の形では汎用的なDTDの設計は不可能であるという他はない。しかし逆に印刷出力にとらわれず文書構造のみを分析するならば、汎用的なDTDの設計は十分可能である。

そこで今回は、紙の上でのレイアウトを意識せずに、純粋に会議録の文書構造を分析してみた。

以下に自治体議事録(会議録)の構造を示す。字下げされたものは入れ子になっている ことを表す。各項目名の後ろに()で括られたものはその具体例である。

# 会議録

前段

表紙

題(東西市会議録) 開会日(平成7年2月15日開会) 閉会日(平成7年3月6日閉会)

目次

本文

題(平成7年第1回定例会)

開会日時(平成7年2月15日水曜日午前10時開議)

議事日程

日程一覧

副題 (議事日程)

日程次第(日程第n)

議案(会期決定の件)

会議付託案件

副題(本日の会議に付した案件)

案件、または案件名

名簿

# 出席議員

肩書き (議長)

議員番号 (110番)

議員名 (小池好子君)

# 欠席議員

肩書き

議員番号 (120番)

議員名(保田修君)

欠員(5人)

議員番号

# 列席者

肩書き(市長)

人名 (常岡稔君)

# 説明員

肩書き (総務部長)

人名 (平原実雄君)

# 事務局職員出席者

肩書き (事務局長)

人名 (丸山直美君)

# 議事

開会時間(午前10時21分開会)

# 発言者

肩書き (議長)

議員番号(110番)

議員名 (小池好子君)

#### 発言内容

文 (これより、本日をもって招集されました・・・)

不規則発言(異議なし、と呼ぶ声あり)

閉会時間(午後4時41分閉会)

# 議員署名

添付資料

# 4. 2. 2 自治体議事録のデータベース仕様文書 (DTD) 設計

前項で自治体議事録の大まかな文書構造を載せたが、その分析を更に進め、今回自治体議事録のDTDを製作した。DTDを作るにあたっては、以下の点を注意して設計を行った。

- (1).文書インスタンスが蓄積されたときにその分類管理がしやすいDTDとする。 この点については、3.2.2 で述べたが、時系列の管理と会議種別での管理が可能になるようにした。
- (2).検索性・索引性の高いDTDとする。

先に検討した弊社会議録検索システムですでに実現しているところの質問者発言の 抽出や特定者の発言抽出などの機能を実現できるのはもちろんのこと、議題と発言の関 連などが容易に取れるような設計とした。

以下にそのDTDを掲載する。

</th <th>会議録し</th> <th>DTD</th> <th>(案)</th> <th>• .</th> <th>&gt;</th> <th></th>	会議録し	DTD	(案)	• .	>	
/td <td>kaigirok [</td> <td></td> <td></td> <td></td> <td></td> <td></td>	kaigirok [					
ELEMENT</td <td>kaigirok</td> <td></td> <td>(front?, body)</td> <td>+(referto other </td> <td>quote)</td> <td>&gt;.</td>	kaigirok		(front?, body)	+(referto other	quote)	>.
ATTLIST</td <td>kaigirok</td> <td>kid</td> <td>ID #IM</td> <td>PLIED</td> <td></td> <td>&gt;</td>	kaigirok	kid	ID #IM	PLIED		>
ELEMENT</td <td>referto</td> <td></td> <td>(#PCDATA)</td> <td></td> <td>&gt;</td> <td></td>	referto		(#PCDATA)		>	
ATTLIST</td <td>referto</td> <td>idret</td> <td>f IDREF #RE</td> <td>QUIRED</td> <td>&gt;</td> <td></td>	referto	idret	f IDREF #RE	QUIRED	>	
ELEMENT</td <td>other</td> <td></td> <td>(#PCDATA)</td> <td></td> <td>•</td> <td> &gt;</td>	other		(#PCDATA)		•	>
ELEMENT</td <td>quote</td> <td></td> <td>(#PCDATA)</td> <td></td> <td></td> <td>&gt;</td>	quote		(#PCDATA)			>
ENTITY</td <td>% persons</td> <td>"(roi</td> <td>le, no, na)+"</td> <td></td> <td>&gt;</td> <td></td>	% persons	"(roi	le, no, na)+"		>	
ELEMENT</td <td>role</td> <td>- 0</td> <td>(#PCDATA EMPTY</td> <td>)</td> <td>&gt;</td> <td></td>	role	- 0	(#PCDATA EMPTY	)	>	
ELEMENT</td <td>no</td> <td>- 0</td> <td>(#PCDATA EMPTY</td> <td>) .</td> <td>&gt;</td> <td></td>	no	- 0	(#PCDATA EMPTY	) .	>	
ELEMENT</td <td>na</td> <td>- 0</td> <td>(#PCDATA EMPTY</td> <td>)</td> <td>&gt;</td> <td></td>	na	- 0	(#PCDATA EMPTY	)	>	
ATTLIST</td <td>na</td> <td>party</td> <td>/ ID #IMPLIED</td> <td></td> <td>&gt;</td> <td></td>	na	party	/ ID #IMPLIED		>	
</th <th>書誌情報</th> <th>R</th> <th></th> <th></th> <th></th> <th>ı —&gt;</th>	書誌情報	R				ı —>
ELEMENT</td <td>front</td> <td>- 0</td> <td>(titlesec, toc</td> <td>, conven?, sign?)</td> <td></td> <td>&gt;</td>	front	- 0	(titlesec, toc	, conven?, sign?)		>
ELEMENT</td <td>titlesec</td> <td>- 0</td> <td>(title &amp; date)</td> <td></td> <td>&gt;</td> <td></td>	titlesec	- 0	(title & date)		>	

```
<!ELEMENT title
                       -- (#PCDATA)
                                                                     >
<!ELEMENT date
                            (open, close)
                                                             >
<!ELEMENT open
                       - 0 (#PCDATA)
                                                             >
<!ELEMENT close
                       - 0 (#PCDATA)
<!ELEMENT toc
                       - 0 EMPTY
                                                             >
<!ELEMENT conven
                       - 0 (title, table)
<!ELEMENT table
                       -- (hi, youbi, naiyou)
<!ELEMENT hi
                       - 0 (#PCDATA)
<!ELEMENT youbi
                       - 0 (#PCDATA)
<!ELEMENT naiyou
                       - 0 (#PCDATA)
                                                             >
<!ELEMENT sign</pre>
                       - 0 (title, %persons;)
<!--
               本文
<!ELEMENT body
                      - 0 (council+, que?, omd?)
                                                            . >
<!ELEMENT council
                      - 0 (titlesup, title, data, conf)
                                                             >
<!ATTLIST council
                      id
                            1 D
                                      #REQUIRED
<!ELEMENT titlesup
                       -- (kyear & kmonth & knumb & ksyub)
<!ELEMENT kyear</pre>
                       - 0 (#PCDATA)
<!ELEMENT kmonth
                      - 0 (#PCDATA|EMPTY)
                                                             >
<!ELEMENT knum
                      - 0 (#PCDATA|EMPTY)
<!ELEMENT ksyub
                      - 0 (#PCDATA)
<!ELEMENT data
                      - 0
       ((day, begt),
          ((count?, attend, absent, blank, resseki?, expl, clerk)&(sched))) >
<!ELEMENT day
                      - 0 (#PCDATA)
<!ELEMENT begt
                      - 0 (#PCDATA)
                                                             >
<!ELEMENT count
                      - 0 (#PCDATA)
<!ELEMENT attend</pre>
                      - 0 (title, %persons:)
                                                             >
<!ATTLIST attend
                              ID
                       i d
                                      #REQUIRED
                                                             >
<!ELEMENT absent</pre>
                      - 0 (title, %persons;)
<!ATTLIST absent
                      id
                            - ID
                                      #REQUIRED
                                                             >
<!ELEMENT blank
                      - 0 (title, %persons;)
```

```
<!ATTLIST blank
                      id
                           ۱D
                                    #REQUIRED
<!ELEMENT resseki
                    - 0 (title, %persons;)
                                                          \geq
<!ATTLIST resseki
                     id ID
                                    #REQUIRED
<!ELEMENT expl
                     - 0 (title, %persons:)
<!ATTLIST expl
                     id
                           ID
                                    #REQUIRED
                     - 0 (title, %persons:)
<!ELEMENT clerk
<!ATTLIST clerk
                     id
                             ID
                                    #REQUIRED
<!ELEMENT sched
                     - 0 (title, no, time?, billsec+, sup?, clr?)
<!ATTLIST sched
                      id ID
                                    #REQUIRED
<!ELEMENT time
                     - 0 (#PCDATA)
                                                          >
<!ELEMENT billsec
                    - 0 (no?, billb|k)+
                                                          \rangle
<!ELEMENT billblk
                    - 0 (ban, bill)+
                                                          >
<!ELEMENT ban
                     - 0 (#PCDATA)
                                                          >
<!ELEMENT bill
                     - 0 (#PCDATA)
                    billid ID
<!ATTLIST bill
                                    #REQUIRED
<!ELEMENT sup
                     - 0 (title, p+)
                                                          >
                     - 0 (#PCDATA)
<!ELEMENT cir
                                                          >
<!ELEMENT conf
                     - 0 (time?, confsec+)
                    - 0 (title, (speaker, speech)+)
<!ELEMENT confsec
                                                                  >
<!ATTLIST confsec
                     id ID
                                                          >
                                    #REQUIRED
                    - 0 %persons;
<!ELEMENT speaker
                                                          >
<!ELEMENT speech
                     -0 (p+)
<!ATTLIST speech
                     htyp (chairman|question|answer) chairman
                     billid ID
                                    #IMPLIED
                                                                  >
<!ELEMENT p
                     - 0 (#PCDATA)
                                                          >
<!ATTLIST D
                     billid ID
                                    #IMPLIED
                                                                  >
<!ELEMENT signature - 0 (title?, (%persons:)+)</pre>
                                                          >
<!ELEMENT que
                    - 0 (title, figgrp+)
                    -- (fig+, figcap?)
<!ELEMENT figgrp
<!ATTLIST figgrp
                    id ID
                                    #REQUIRED
<!ELEMENT fig</pre>
                      - 0 EMPTY
<!ATTLIST fig
                             NAME
                                    #REQUIRED
                      name
```

```
        <!ELEMENT figcap</td>
        - 0 (#PCDATA)
        >

        <!ELEMENT omd</td>
        - 0 (title, toc, omdsec+)
        >

        <!ELEMENT omdsec</td>
        - 0 (title, figgrp+)
        >

        ]>
```

このDTDの中で定義されているタグを次に一覧する。

表 4-1. DTDタグ一覧

No.	タグ	意味
1	kaigirok	文書タイプ(=自治体会議録)
2	referto	相互参照
3	other	その他(拍手など)
4	quote	固定情報
5	%persons	個人名
6.	role	役職
7	no	各種番号
8	na	発言者名
9	front	書誌情報
10	titlesec	本の表紙
11	title	各種題名
12	date	開会日と閉会日のブロック
13	open	開会日
14	close	閉会日
15	toc	目次
16	conven	定例会日割り表
17	table	日割り表内の表
18	hi	日割り表の日付
19	youb i	日割り表内の曜日
20	na i you	日割り表の会議内容
21	sign	署名議員
22	body	本文
23	council	会議録第?号

24	titlesup	会議の基礎データ
25	kyear	招集年
26	kmonth	招集月
27	knum	招集回数
28	ksyub	会議種別
29	data	その会議の情報(のブロック)=出欠議員、議案名など
30	day	日付
31	begt	会議の開始時刻
32	count	出席者数
33	attend	出席議員
34	absent	欠席議員
35	blank	欠員
36	resseki	列席者
37	expl	説明のための出席者
38	clerk	議会事務局からの出席者
39	sched	議事日程
40	time	議事日程の中の開会時刻
41	billsec	議案のセクション番号と議案番号と議案名を含むブロック
42	billblk	議案番号と議案名を含むブロック
43	ban	議案番号
44	bill	議案
45	sup	追加日程
46	clr	議会事務局からの報告
47	conf	議事
48	speaker	発言者名
49	confsec	議事の中で目次に出さなければならない区切り
50	speech	発言内容
51	р	文章(段落)
52	signatue	署名
53	omdsec	朗読を省略された文書の部の節
54	omd	朗読を省略された文書の部
55	figgrp	図のグループ
56	figcap	図のキャプション
57	figgrp	図=ファイル名

さらに、先に挙げた自治体議事録(会議録)の構造図にそのタグを当てはめると以下の ようになる。

```
<kaigiroku>会議録
                    <front>前段
                                         <titlesec>表紙
                                                              <title>題(東西市会議録)
                                                               <open>開会日 (平成7年2月15日開会)
                                                              <close>閉会日(平成7年3月6日閉会)
                                         <toc>目次
                    <body>本文
                                                <data>
                                                                             <title>題(平成7年第1回定例会)
                                                                             <day>,<begt>開会日時(平成7年2月15日水曜日午前10時開議)
                                                                             <shed>議事日程
                                                                                                  <br/><br/>billsec>日程一覧
                                                                                                                        <title>副題(議事日程)
                                                                                                                        <ban>日程次第(日程第n)
                                                                                                                        <br/>
<br/
                                                                                                  <sup>会議付託案件
                                                                                                                        <title>副題(本日の会議に付した案件)
                                                                                                                        <定体、または案件名
                                                                             <shed>
                                                                             名簿
                                                                             <attend>出席議員
                                                                                                  <role>肩書き(議長)
                                                                                                  <no>議員番号(110番)
                                                                                                  <na>議員名(小池好子君)
                                                                             <absent>欠席議員
                                                                                                  <role>肩書き
                                                                                                  <no>議員番号(120番)
```

<na>議員名(保田修君)

<br/>
<br/>
dank>欠員(5 人)

<no>議員番号

<resseki>列席者

<role>肩書き(市長)

<na>人名 (常岡稔君)

<expl>説明員

<role>肩書き (総務部長)

<na>人名(平原実雄君)

<clerk>事務局職員出席者

<role>肩書き(事務局長)

<na>人名(丸山直美君)

</data>

議事

<time>開会時間(午前10時21分開会)

<speaker>発言者

<role>肩書き (議長)

<no>議員番号(110番)

<na>議員名(小池好子君)

<speech>発言内容

文(これより、本日をもって招集されました・・・)

<other>不規則発言(異議なし、と呼ぶ声あり)

<time>閉会時間(午後4時41分閉会)

<signature>議員署名

<omdsec>,<que>,<figgrp>添付資料

</body>

</kaigiroku>

時系列の文書管理については、

タグ<kyear>、<date>、<knum>、によって、いつの会議か、何回目の会議かが分かるようにした。また、会議種別による管理については、<kaigirok>タグの属性により、会議の種別を指定しておけるようにした。

例えば、

- 1:定例会
- 2:臨時会
- 3:委員会
- 4:全員協議会
- 5:その他の会議

と定義したとすれば、文書インスタンスの方は、

<kaigiroku kid="1">

などと記述できるようにした。なお、極力柔軟性を持たせるためにこの属性は省略可能とした。また、<ksyub>タグによっても会議種別は拾えるようになっている。

通常の場合、政党や会派は会議録上には出現しないが、場合によって会派などでのひも付けもありうると思われる。そこで<na>タグは会議録に記載される個人名を示すタグであるが、このタグにも属性を付けることによって会派などを表現できるようにした。

例).

<na party="会議録研究所">小池好子君

発言については、属性を

<!ATTLIST speech

htyp (chairman | question | answer) chairman

billid ID

#IMPLIED

と宣言したが、chairman | question | answer は、議長発言、質問、答弁のいずれかを意味し、省略時は議長発言になるようにした。また、もう一つの属性名 "billid" は議題を表し、属性値が < billblk > タグの属性値と一致させることにより議題との関連を表現できるようにした。

例).

<bill billid="10">議案第31号</bill>

<speech htyp="10">

<酶 3 1 号について質問させていただきます。まず、・・・

議会の通常の議事進行は、一つの議案についていくつかの発言があり、それが終わると次の議案についての発言が行われる形である。この議事進行の原則に照らせば、<speech>(発言)タグよりも大きいくくりで議題と関連させるタグを入れるのが理想的である。

# <議案についての討議 billid="100">

<speech htyp="question">質問します。・・・

<speech htyp="answer"> お答えします。・・・

<speech htyp="question">さらに質問します。・・・

<speech htyp="answer"> お答えします。・・・

<speech htyp="chairman">それでは、この議案については終わります。

</議案についての討議>

しかし、実際には一発言で複数の議案についての質問が行われたり、いったん終結した はずの議案について発言が突如遡及したりすることも往々にして発生する。こうしたケースに対応するため各発言ごと、さらに各段落ごとに議題を表現する属性を付加せざるをえなかった。

#### 例).

<bill billid="10">議案第31号</bill>

<speech htyp="10">

<この件について質問させていただきます。まず、・・・

あと先ほどまで議論になっていた件に関してですが・・・

</speech>

# 4.2.3 自治体議事録SGMLデータベースの構築

このようなDTDに基づいてSGML文書インスタンスが作成された後、データベースを以下のように構築することにした。それを説明するに際してはさしあたり理解しやすいようにリレーショナルデータベースのテーブルイメージで表現することにする。さらに実際の文書インスタンスをデータベースに入力するには、SGMLパーサを使用する。

まず、文書インスタンスから文書管理情報として、会議ごとの管理テーブルが作成できる。

会議D	<day></day>	<kyear></kyear>	⟨kaigiroku kid=⟩	<a href="mailto:line"></a>	(knum)	<ksyub></ksyub>
1	1995/3/1	平成 7年度	1	3月	第 1回	定例会
2	1995/6/12	平成 7年度	1	6月	第 2回	定例会
3	1995/9/4	平成 7年度	1	9月	第 3回	定例会
4	1996/10/29	平成 8年度	3	10月		決算特別委員会
*************************	******************************	/				1. Maria de Caración (1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1
				41144144144144	***************************************	***************************************

# 図 4-1. 会議管理テーブル

各カラムはDTDで定義したタグと対応させており、基本的にはタグがつけられているデータをそのまま取り込み、必要に応じてデータ型を変換している(日付や年号など)。 各会議は文書に分かれるため、さらにその詳細である文書情報に分解できる。これは同時に会議内の日程情報でもある。このテーブルにも対応するSGMLタグがつけられたデータが取り込まれる。

先頭の会議IDはレコードのユニークキーである。

<kaigiroku kid=>の列のデータは会議の種別を表しているが、これは前節で挙げた例のようなタグの属性値である。

会議ID	文書ID	<open></open>	<title>&lt;/th&gt;&lt;/tr&gt;&lt;tr&gt;&lt;th&gt;1&lt;/th&gt;&lt;th&gt;1&lt;/th&gt;&lt;th&gt;1995/3/1&lt;/th&gt;&lt;th&gt;03月01日-00号&lt;/th&gt;&lt;/tr&gt;&lt;tr&gt;&lt;th&gt;1&lt;/th&gt;&lt;th&gt;2&lt;/th&gt;&lt;th&gt;1995/3/1&lt;/th&gt;&lt;th&gt;03月01日-01号&lt;/th&gt;&lt;/tr&gt;&lt;tr&gt;&lt;th&gt;1&lt;/th&gt;&lt;th&gt;3&lt;/th&gt;&lt;th&gt;1995/3/6&lt;/th&gt;&lt;th&gt;03月06日-02号&lt;/th&gt;&lt;/tr&gt;&lt;tr&gt;&lt;th&gt;1&lt;/th&gt;&lt;th&gt;4&lt;/th&gt;&lt;th&gt;1&lt;b&gt;99&lt;/b&gt;5/3/8&lt;/th&gt;&lt;th&gt;03月08日-03&lt;del&gt;号&lt;/del&gt;&lt;/th&gt;&lt;/tr&gt;&lt;tr&gt;&lt;th&gt;1&lt;/th&gt;&lt;th&gt;5&lt;/th&gt;&lt;th&gt;1995/3/9&lt;/th&gt;&lt;th&gt;03月09日-04号&lt;/th&gt;&lt;/tr&gt;&lt;tr&gt;&lt;th&gt;1&lt;/th&gt;&lt;td&gt;6&lt;/td&gt;&lt;td&gt;1995/3/10&lt;/td&gt;&lt;td&gt;03月10日-05号&lt;/td&gt;&lt;/tr&gt;&lt;tr&gt;&lt;th&gt;1&lt;/th&gt;&lt;th&gt;7&lt;/th&gt;&lt;th&gt;1995/3/24&lt;/th&gt;&lt;th&gt;03月24日-06号&lt;/th&gt;&lt;/tr&gt;&lt;tr&gt;&lt;th&gt;2&lt;/th&gt;&lt;th&gt;8&lt;/th&gt;&lt;th&gt;1995/6/12&lt;/th&gt;&lt;th&gt;06月12日-00号&lt;/th&gt;&lt;/tr&gt;&lt;tr&gt;&lt;th&gt;2&lt;/th&gt;&lt;th&gt;9&lt;/th&gt;&lt;th&gt;1995/6/12&lt;/th&gt;&lt;th&gt;06月12日-01号&lt;/th&gt;&lt;/tr&gt;&lt;tr&gt;&lt;th&gt;2&lt;/th&gt;&lt;th&gt;10&lt;/th&gt;&lt;th&gt;1995/6/15&lt;/th&gt;&lt;th&gt;06月15日-02号&lt;/th&gt;&lt;/tr&gt;&lt;tr&gt;&lt;th&gt;2&lt;/th&gt;&lt;th&gt;11&lt;/th&gt;&lt;th&gt;1995/6/16&lt;/th&gt;&lt;th&gt;06月16日-03号&lt;/th&gt;&lt;/tr&gt;&lt;tr&gt;&lt;th&gt;&lt;/th&gt;&lt;td&gt;&lt;/td&gt;&lt;td&gt;,&lt;/td&gt;&lt;td&gt;&lt;/td&gt;&lt;/tr&gt;&lt;tr&gt;&lt;th&gt;&lt;/th&gt;&lt;td&gt;,&lt;/td&gt;&lt;td&gt;,&lt;/td&gt;&lt;td&gt;•&lt;/td&gt;&lt;/tr&gt;&lt;/tbody&gt;&lt;/table&gt;</title>
------	------	---------------	--

図4-2. 会議録第 n号管理テーブル

一方、発言の管理は、一日程(文書)毎に次のようにテーブル化される。発言の種類として、<speech>タグの属性を取り込んでいるほか、名前や役職も文書インスタンスからタグにより識別して取り込むことができる。

個人ID	<na></na>	<role></role>	<pre><speech htyp=""></speech></pre>
59	小池好子君	環境事業部長	3
********************	常岡稔君	教育次長	3
61	石倉一美君	7番	. 2
62	平原実雄君	助役	3
	丸山直美君	42番	2
	浅野正明君	税務部長	3
65	富澤智君	教育次長	3
66	保田修君	18番	2
			PARTITION OF THE PARTIT

図4-3. 発言管理テーブル

このようにSGMLタグで囲まれた部分のデータは部品として切り出すことができるため、データベースの内部にあっては、必ずしもデータはテキストでなければならないわけではなく、最も効率のよいデータ形式や構造でデータベースをインプリメントし、必要な文書管理を実現することができる。

もちろんこのようにSGMLタグに忠実な形でデータを部品化してデータベースを構築 した場合、必要に応じて入力時とまったく同じSGML文書インスタンスをデータベース から出力できるのは言うまでもない。 さらにこのようにSGMLデータベースを構築すれば、3.2.2で少し触れた、全文 検索時にSGMLタグをヒットしてしまう問題とデータベースとして機能させるためには SGMLタグを検索しなければならない、という2つの表裏する問題を同時に解消できる ことも明らかとなった。

ここで一つ問題点を報告しておかねばならない。それは前節で議会議事録の汎用DTD の定義を試みたのだが、実はその設計時にも現実の議会運営のあり方や会議録への記載内 容をまったく無視するわけにはいかない面があったことである。たとえば前節で述べた議題と発言の関係が必ずしも原則的でない点などが典型例だが、その外にも招集回数(タグ <knum>)の数えかたが自治体により異なる点などがある。この招集回数などはデータとして持っていなくとも、会議録データを時系列に管理できていれば本来必要に応じて計算により自動的に求まるものであるが、自治体によって数えかたが異なるためにそれができずに現時点ではデータとして定義せざるを得ない結果となっている。こうした意味で理想的な汎用標準のDTDでは恐らく必要でないタグがまだまだ多分に含まれているわけである。表4-2. に将来不要になりそうなタグの主立ったものを列挙する。

表4-2. 再考の余地が残る主な要素の一覧

15	toc	目次	会議録記載形式の標準化により省略の可能性あり
16	conven	定例会日割り表	11
21	sign	署名議員	II .
27	knum	招集回数	会議管理方式の標準化により省略の可能性あり
.30	day	日付	他のデータとの重複を整理できる可能性あり
32	count	出席者数	自動生成できる余地あり
35	blank	欠員	II .

タグの数は必要最少限に抑えられるべきであることを考えれば、会議録記載形式はもち ろんのこと議会の運営や管理といった面からもCALSの標準化を進めれば、SGML化 は現在想像するよりもずっと低コストで実現できるであろうし、データベースもまた効率 がよくなるに違いない。

# 4. 2. 4 情報検索ブラウザ

上記のように構築されたSGMLデータベースを実際に検索して、表示を行うブラウザ については、特別なアプリケーションを開発するのではなく、WWWブラウザを利用する

のがよいと結論した。

その根拠は、

- (1). CALS/SGMLは、ネットワーク環境が前提になった上での考え方であること、
- (2).さらに、その利用は小規模のLANは想定されておらず、大規模なLANやWAN) が想定されていることから、コストの高いクランアント/サーバ型のアプリケーションでは現実味が薄いこと、
- (3).自治体に対して行ったアンケートの結果においても、LANの構築が終わっている、 もしくは構築予定中の自治体が相当な割合に達しており、環境的に問題がないのは もちろんのこと、むしろネットワークで動作するアプリケーションが不足している 状況にあること、
- (4).ネットワーク環境はともかく、インターネットに対する関心は非常に高く、また接続している自治体がほとんどであり、WWWブラウザにある程度の慣れがあるものと思われること、

などである。

検索結果や会議録データの表示については、SGMLをHTMLに変換する必要があるが、これはタグの変換テーブルを用意すればそのコンバータの作成については問題がない。むしろ、その変換テーブルを各自治体の要望に応じて変化させることができる、すなわち表示出力の形式はそれぞれの現場のニーズに合わせることが可能なので、その意味でも柔軟性が高い。

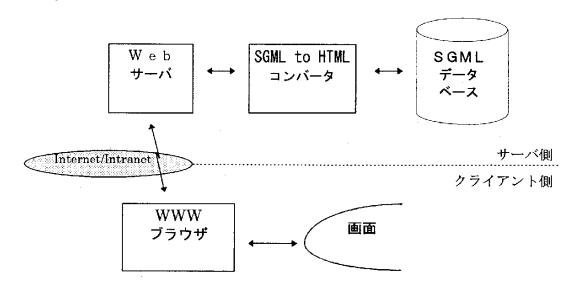


図4-4. SGML データベースとWWWブラウザ

# 4.2.5 自治体議事録SGMLデータベースの運用

本来、文書のSGML化は特定の種類の文書のみをSGML化すればそれで良いというものではない。文書のSGML化は、CALS実現の1ステップとして位置づけられてこそ、その価値を増すものである。すなわち、他のいろいろな文書も同様にSGML化し、それらを相互にリンクさせることにより、効率を向上し、個々の文書の価値をより多く引き出さなければならないのである。会議録には先に述べたとおり、議事に付随、もしくは関連する多数の文書が存在する。したがってこれらの文書を押し並べてSGML化したうえでSGMLデータベースを運用すれば、大きく現在の行政事務を効率化し、文書情報を高付加価値化することができるであろうと思われる。

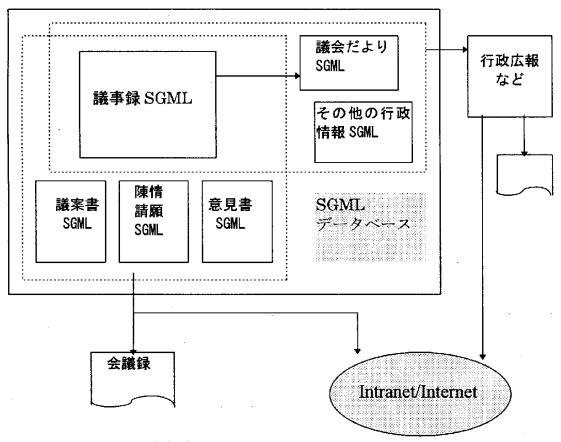


図4-4. SGML統合データベースの運用例

とりわけ現状では、例えばある自治体条例をとったときに、その条例が施行される背景 は何だったのか、その条例はいかなる議論を経て生まれたものなのか、そもそもその条例 が制定されるに至った契機はいかなるものであったのか、などを知るにはたいへんな情報 収集努力をしなくてはならないであろう。しかし、背景や議論について知るには会議録が あり、最初の契機については請願・陳情書があるはずであり、これらの文書を参照すれば よいのであるが、膨大な会議録のどのページを見ればよいのか、いつのどの請願書を読め ばよいのか、当の自治体職員ですらしばしば途方にくれるであろうことが想像に難くない。。しかし、これらの文書が総合してSGML化されて管理されていれば情報の流れは実にスムーズであり、それぞれの文書の存在価値が大きく高まるであろうことは疑う余地がない。 また議会日程などに管理しようとするとき、これまでは議会事務局においてワープロ等において予定表を作成していると考えられる。ただこのようなSGML文書体系以外で作成された文書は他とのリンクを図ることができず、再活用に至らない無駄なデータとなってしまう。しかしSGMLデータベースにおいて日程管理に関わる項目をDTD上で定義付けておけば、SGMLファイルにおいて他の文書インスタンスとともに文書管理情報も押さえておくことは可能となる。これによって効率的な文書管理も実現できることであろう。

しかし、1つの自治体でSGMLデータベースを構築し運用するだけでなく、複数のあるいは全国の自治体がそれぞれにSGMLデータベースを構築・運用することになれば、さらに大きなメリットを生むであろうと思われる。先の条例文から会議録や請願書をひも解く例を再び引けば、ある条例案を検討する際に、すでに先行して施行している自治体では施行に至る過程がどのようなものであったか、施行した後の状況がどのようであるか、などがほとんど時間をかけずに調べることができるようになるであろう。

しかしそのためには、各自治体で等しい文書DTDを使うことが重要である。各自治体でそれぞれ別個の文書DTDを定義していたのでは、文書交換が効率的に行えず、せっかくのSGML化が意味のないものになってしまう。前節では各自治体間での会議録の出力形式の差異をとりあえず無視してDTDの設計を行ったが、このような汎用的で標準的なDTDを各自治体間で共有し、それに基づいて各文書をSGML化すべきであろう。そしてその出力に関しては規定を持たせず、各自治体で思い思いのローカルなスタイルで行うのもよし、標準を踏襲するもよし、というかたちにするのが理想の形態ではないだろうか。しかし今回の地方自治体向けアンケートからもわかるように、SGMLという文書データベース手法をあまり聞いたことがないという現実から考慮すると、SGMLデータ化を主眼において自治体に対して文書管理を提案していくことよりも、本事業を契機にしてまずは議会議事録分野について、その文書データおよびその文書データの管理データまで収めることができるというSGMLデータの利点とそのデータを取り扱う弊社の会議録検索シ

ステム「VOICES」を通じて、まずは体験いただき、その結果SGMLの利点を具体的に理解いただいた上で、他の自治体文書データ分野に関するSGML化を具現化いただく展開を見せていかなければ、本当に意味のある自治体へのSGML普及は見られないことであろう。

自治体における目的はSGML化することではなく、行政事務および文書の取扱いという煩雑な業務の効率化により、行政改革を文書分野においても進めていくことである。現状SGML文書データを取り扱う環境が整っていない状況では、煩雑さが増すばかりである。そこで自治体におけるSGMLの採用は間違いのない方向性であることから、今回の事業をきっかけに各種自治体文書分野におけるSGMLを活かすことができ文書管理AP等のSGMLツールが出現してくることを期待する。

- 5 SGMLデータベースの各種情報メディア出力と文書情報運用の可能性
- 5. 1 印刷メディア

# 5.1.1 自治体議事録の自動組版化

SGMLとは文書構造は規定されているものの、そこには組版レイアウト情報は規定されていない。そこで印刷用組版アプリケーションを利用して版下組版を行う必要があるのだが、文書構造によって規定される文書要素(見出し・段落など)とその文書要素に対応するレイアウト情報(フォントの種類[ゴシック・明朝等]やフォントの大きさ)をリンクさせれば、レイアウト要素である組版コマンドを手作業で配さなくとも自動的に組版を行うことは可能である。

つまりSGML文書を受け付けるフォーマッタにレイアウト情報を付与すれば、そこに 与えられたSGMLファイルが自動的に組版・印刷されることとなるわけである。

また具体的なSGMLデータの印刷への展開に関しては、現状次の2種の系統が考えられる。

# (A) SGML→TeX→PS (ポストスクリプト/ページ記述言語)

・SGML創世期においては、このTeXによる組版出力が主流であると考えられていたが、TeXには組版出力に至るまでに情報処理技術が必要となる。そのためプログラム的工夫と編集作業による人手などを考慮すると、TeXの運用に慣れたオペレーターが存在しない場合には実用化に時間がかかってしまうことと考えられる。

#### (B) SGML→各種DTPソフト→PS(ポストスクリプト/ページ記述言語)

・米国製品等多数SGML文書を取り扱うことができるDTPソフトが国内でも入手できるが、縦組など日本語独特の組版に対応しきれていないソフトが多いのが現状である。そんな中、日本語SGML対応と日本語組版技術の優秀性を併せ持ったDTPソフトが「ページコンプ(大日本スクリーン製造)」である。SGMLタグをページコンプのネイティブタグに変換することにより容易にSGMLデータの流し込みが可能である。

# 5. 2 インターネット/WWW

# 5. 2. 1 SGML→HTML変換

前述のとおり、HTMLとはSGMLから派生したものであり、ハイパーテキスト用の 汎用タグを付したものである。そこでDTDによって定義された文書要素毎にどのHTM Lの汎用タグに置換するのかを定義すれば、容易にバッチ処理によってSGML to HTM L変換を行うことが可能である。

これまでの文書データとは、ワープロソフトによって、そのレイアウトを確認しながら データ構築がなされるという文書作成方法であったが、SGMLデータは容易にHTML 変換できることから、文書作成時には文書構造のみ明確にすることに専念し、文書レイア ウトについてはHTML変換後に検討するという考え方で文書作成するべきである。

# 5. 2. 2 インターネットによる自治体議事録情報の発信

全国的にあまり事例は多くはないが、議会のホームページを開設している自治体もある。 (埼玉県・北海道標茶町等) しかしそれらは、議会の仕組みや議員紹介などのPRページ がほとんどであり、議会議事録について本格的に公表しているケースはまだ少ない。

そこで議会議事録コンテンツのSGML化は、前述のとおり容易にHTML変換によってインターネットによる情報発信が可能になることから、議事録データの情報発信も容易に可能となる。しかしそれのみでは意味がない。関連行政文書とのリンクにより、よりダイナミックな情報検索が可能となる仕組みづくりが必要である。また写真等の画像データも表現方法として取り込むことが可能であることから、議会で議論されている案件に関連した画像情報も含めて発信すれば、印刷物による文字のみの議会議事録以上に理解しやすく具体的な情報検索が住民において可能となるはずである。

#### 5.3 CD-ROM

# 5. 3. 1 自治体議事録のCD-ROMによる展開

議会議事録に求められる機能として、その検索性という面もあげられる。そのために弊社においても議事録検索システムを各自治体議会事務局へ提案しているが、それが自治体他部門および住民サイドで実現することができるのが、CD-ROMという大容量の情報メディアである。

MAX500MBという大容量を持ったこのCD-ROMという媒体は、重さ50gの中に紙で換算すると数万ページのデータを収録することができる。他の利点としても、制作費の削減、メンテナンスの即時性、可搬性、デジタルデータの格納(文字・画像・音声・動画など)、コピー単価が安価などがあげられ、大量データの一括管理には適したものである。

自治体議事録も年間4回の定例議会が開催されたと考えた場合、それを印刷物によって保管・活用しようとすると、そのページ数と容量・重量は相当なものとなり、容易に持ち運びできるものではない。この点だけでも有効ではあるが、その他にも検索機能が印刷物のように固定化(目次・索引など)されず自由度が高い、画像と文字データのリンクが可能なことから多面的な情報検索が可能、音声・動画とのリンクにより一層臨場感のある表現が可能などの有効性が考えられる。

各種オーサリングツールを活用して、検索画面と検索方法を設計すれば、確実に印刷物以上の表現はできるはずであることから、たしかに印刷物のようにパソコンがなければいつでもどこでも見ることができるという点には劣るが、保存という機能の他にCD-ROMでは検索性とそのメディア表現の豊かさによって、ニーズは高いものと考えられる。

#### 5.4 情報運用

#### 5.4.1 文書管理機能の活用

前項では各種情報メディアへのSGML出力の可能性に関して検証を行ったが、SGML文書ファイルの意義は、文書の中身であるコンテンツとその文書の管理情報がともに存在するという点にあることから、その文書管理情報の活用も運用段階では考えていきたい。(この点がHTMLや印刷用組版コマンドとの大きな相違点である。)

例えば今回の議会議事録DTDを見ても、開催スケジュールや出席議員などの文書管理情報が含まれている。これまでのデータベース思想では、文書データと管理データは各々別のファイルとして有し、そのリンクによって文書管理が行われてきた。しかし往々にして文書データは更新したにも関わらず管理データは未更新というケースが発生し、双方のデータ間で矛盾が生じてしまうケースがある。

これは複数のファイルが存在することによる弊害であり、SGMLファイルであればその文書構造(DTD)において管理データ項目を明確化しておけば、1つのファイルで文書データ・管理データの両面の把握ができるのである。

また具体的な運用においては、その都度SGMLデータすべてをアプリケーションで引き出してくる必要はなく、文書データ・管理データ各々を取り扱うアプリケーションにおいて、その必要とするデータのタグを認識させ、そのデータのみを取り扱うように考えれば、運用オペレーション上はこれまでのスタイルとなんら変化はなく問題も生じない。

日本におけるSGMLの論議はメディア出力の部分に集中し、まだHTMLの域を超えていない。そこでこの文書管理の面を検証していくことがCALSの実現につながり、しかも日本におけるSGMLの普及に寄与するのではないかと考える。

#### 5.4.2 他自治体議会議事録との横断的データ検索・活用

もう1つのSGML化の意義は、文書構造の標準化、定型化にある。これまでの議会議事録は自治体ごとの印刷出力形態(組版スタイル)の違いから、必然的に文書形式が各々異なっていた。しかしDTD設計段階において多くの自治体の議会議事録を分析したところ、文書構造自身は標準化することが可能であるとの結論に至り、その結果が今回の議会議事録DTDとなった。そこでSGML化を図ることによって文書構造が統一されること

から、他の自治体議事録も含めた横断的な情報検索も容易になることと考えられる。

例えば、議案テーマのタグが付された文書データの中に「老人福祉」という単語が出てくる議案の議事内容データを抽出したいという場合、SGMLデータであれば容易に検索・抽出ができ、しかもどの自治体に関しても統一したフォーマットでのデータ出力が可能である。しかし組版スタイルをベースとしたこれまでの文書データでは、全文検索エンジンの活用によって、キーワード抽出はできるが、その議事内容について統一したフォーマットで出力するには至らないのである。

情報公開と自治体間連携が進んでいく行政改革の時代に、このような情報検索方法のニーズは益々高まるはずである。そのためにも行政文書、そして議会議事録に関して、SGMLによる統一的データベース構造によって文書データを取り扱うことの意義がある。

- 6 今後の展望
- 6.1 今後目指す技術開発
- 6.1.1 SGMLデータベースによる自治体議事録運用システム化の展望

SGMLによる自治体議事録管理の今後を検討していくと、議会そのもののあり方についての再検討が必要であると考えられる。議会とは住民に対してより豊かな生活ができる自治体運営を検討する機関ではあるのだが、現状においてどうしても議会と住民の間には大きな壁があり、住民が議会運営や行政活動の中身について理解しにくい状況にあるように考えられる。

これはこれら議会活動や行政活動が議会報や広報などの印刷物を通じて一方的に住民に報告されるのみという片方向型コミュニケーションメディアしか存在していなかったことに要因があると考えられる。

しかしインターネットの発展により、住民1個人でもこの通信媒体を通じて容易に自らの意見を示すという情報発信が可能な状況に変遷してきたのである。そこで最終的に目指す自治体および議会のあり方とは、住民・議会議員そして行政機関の職員が1人1台のパソコンを有し、インターネット等の情報通信技術によって、議論すべき行政課題について、互いの意見を出し合うバーチャル議会が目標となるのではないだろうか。

またこのような体制を整備することによって、議会組織そのものは住民の代表者が議論する場として残すべきではあるが、次のようなことも可能になる。

- ・議論すべき行政課題について住民が意見を発信する。それらの意見を集約し、住民の代 弁者である議会議員がそれらの意見をベースとして議会において議論を行う。
- ・陳情に出向くこと必要がなく、インターネットを通じて電子的な陳情が容易に行われる。
- ・議員は必要に応じて、自ら行政文書情報データベースから求める情報を検索して、現在 の自治体運営の状況の確認と議会決定された案件がその後行政内部でどのように対応さ れているのかを検索して確認を行う。
- ・同様に住民も議会での議論と行政活動が誤りのない方向で進行しているのかを確認する ことができる。
- ・議会議員を選挙によって選出する際にも、その議員が議会によってどんな活動を行って きたのか容易に検索することができる。

各地で盛んに活動している市民オンブズマンも行政活動のあら捜しが目的ではなく、このような開かれた行政・議会を求める現れであると考えられる。SGMLによる文書管理体制の整備とインターネットを通じた双方向コミュニケーション技術はこのような見えない壁を取り外すことができ、住民・行政機関・議会各々の効率化につながり、一層注力すべきクリエイティブな部分への業務の移行がこれらの展開によって実現することができ、これらの活動が「暮らしやすい自治体」づくりに寄与することを期待するものである。

# 禁無断転載

平成9年3月発行

発 行 財団法人 データベース振興センター 東京都港区浜松町二丁目4番1号 世界貿易センタービル7階

TEL 03-3459-8581

委託先 株式会社 会議録研究所

東京都新宿区市谷八幡町16-209 TEL 03-3267-6051

印刷所 フュージョン 株式会社

札幌市中央区南15条西9丁目 TEL 011-551-8055

