

## 行政情報の標準化に対する調査研究

池田聡史<sup>i</sup>**Research for the Standardization  
of the Governmental Information**Satoshi IKEDA<sup>ii</sup>

*The Technological Research Council on Computer Utilization (RIYOKEN) was established in the Agency of Industrial Science and Technology, MITI in 1968 to conduct multilateral and professional research mainly on computer utilization problems in administrative fields from a technical standpoint by having experts from the various governmental organizations. Nowadays, almost all the kinds of administrative documents in Japanese government offices have become electronic documents. But it is difficult to exchange the documents between each offices because the formats of the electronic documents is not standardized. The Government have also begun to supply its information by electronic documents such as CD-ROMs in white papers. But it is necessary for the users to prepare the software that can be adapted each document because the formats of the electronic documents are not unique. Here we take up the SGML as a standardized electronic document format and research and examine the usefulness and possibility of SGML in administrative fields to achieve the standardization of the governmental information.*

## 1. はじめに

行政情報には、国民が行政組織に対して行う許認可等の申請文書、行政組織内においてそれを処理するための付帯情報、広く国民へ情報を提供する白書等、さらには行政組織内の事務処理情報、行政組織間における連絡事項等、様々なものがある。近年まではこれらの情報のすべては紙による文書という形でのみやり取りされてきた。

しかし、行政事務の合理化をさらに押し進めるためにパソコン一人一台の普及、霞が関 WAN の本格運用等による行政情報の電子化が急務となり、これらの情報が紙だけでなく電子情報として作成、交換、加工されることが必要となった。特に、電子情報の交換を円滑に行うためには各省庁において扱われる電子情報の標準化が必要である。また、これらの行政情報や研究成果等は白書等で CD-ROM 等の電子情報でも提供されているが、各省庁が現在提供する電子情報はデータの格納形式が統一されておらず、利用者が電子情報を読み込み、加工する場合には個別にソフトウェアを用意する必要がある。このため、情報相互利用を進める上での弊害となっている。現在、統一された電子データの格納形式としては SGML が世界的に注目されてきている。そこでこの SGML を取り上げ、行政情報の標準化を進めるため行政分野における SGML の活用の可能性について検討する。

## 2. 電子計算機利用に関する技術研究会(利用研)とは

電子計算機利用に関する技術研究会(略称:利用研)は、行政分野における電子計算機利用上の技術的側面からみた諸問題について多角的かつ専門的な調査研究を行うため、各省庁の意向を受けて、昭和43年5月、通商産業省工業技術院に設置された。各省庁の情報処理部門の担当者が集まって行政情報処理に対する調査研究を行う機関としては利用研が我が国唯一のものであり、現在では29省庁、5オブザーバ機関の参加による3分科会、9研究班での活発な調査研究活動を行っている。利用研の研究成果は、研究会総会での報告、利用研の研究成果と合わせて省庁のシステム等を紹介する研究発表会の開催、利用研だよりの発行、利用研ホームページ(<http://www.riyoken.go.jp/>)等により、広く普及を図っている。

## 3. 行政情報の電子化状況と問題点

現在、多くの省庁においてはLAN環境が整備され、PCの一人一台化も進んでいる。このような状況において行政情報の電子化状況と現在までの問題点を検討する。

**官庁の情報システム調達の特徴** 官庁においてはある一定の金額を超える全ての物品の調達(購入)は基本的に競争入札となる。これは情報システムといえども例外ではない。入札では特定のベンダーによる製品を指定して購入するわけにはいかないため、必要とする機能のみを列挙し、それを満たしている製品であれば全てが落札の対象となりうる。また、情報システム機器のように技術の陳腐化が激しいものに関しては購入ではなく、借用とすることが一般的である。したがって、各省庁においてワープロだけでも数種類の製品が使用されており、しかも製品のバージョンは調達を行った当時のものを数年間も使わざるを得ないためにまちまちである。

このような状況の中で、官庁における行政情報の電子化は急速に進んでは来たものの電子データの標準化は遅れるばかりであり、遅れば遅れるほど各官庁では製品等に依存した独自形式による電子データの蓄積が進み、ますます標準化が難しくなっている。

**霞が関WAN** 政府が行政のあらゆる分野において情報通信技術の成果を普遍的に活用し、行政の質の高度化、国民サービスの質的向上を図ることを目標として平成6年12月に閣議決定を行った「行政情報化推進基本計画」に基づいて、平成9年1月より霞が関WANの運用が開始された。平成9年3月現在で19省庁20機関が接続され、平成10年度には28省庁等31機関が接続される。これにより各省庁とのやり取りはインターネット経由ではなく霞が関WAN経由で実行することが可能となるため、白書等のデータベースや公文書交換システム等が実現可能となる。

**文書取り扱い規則について** 行政事務においては法律文書、契約文書等の作成をはじめとして文書の取り扱いには厳密な規則が存在する。この規則は法律や省令により定められており、おいそれと変更が可能なものではない。例えば、起案(稟議)文書では、紙での決裁が必要であり、また内容に応じてその起案文書を施行後すぐに廃棄してよりものから永久に保存する必要があるものまで保存期間が厳密に決められている。これらの文書は比較的体裁が厳密に決まっている上に、定型業務となっているものも多い。また、省庁間での公文書のやり取りも現在では紙でのみ行われているが、これも将来は霞が関WANを用いて電子的にやり取りすることが検討されている。

現在では公文書の作成は一部の添付資料等を除いてほとんどはすでにワープロによって作成されており、公文書の電子化は比較的早期に実行が可能となる見込みであり、通商産業省においては実際に電子的に決裁を行う電子決裁システムの開発に取り組んでいる。しかしながら、各省庁において作成される電子データの型式は各省庁が現在使用しているワープロに依存するため、公文書の交換様式の仕様の決定が難航している。例えば、現在では標準的なワープロは HTML 型式や RTF 型式により文書の保存が可能となっているものが多いが、未だにそれらの型式での保存が不可能である古いバージョンのワープロを用いている省庁も多い上に、HTML 型式や RTF 型式ではワープロによって文書の再現性が十分ではないこと、項目ごとに記述されるわけではないためデータベースとしての利用が制限されること等の欠点がある。

#### 4. 行政分野における SGML の活用の可能性についての検討

行政情報の標準フォーマットとして SGML が注目されてきている。現在のところ作成した文書を自動的に SGML 文書に変換する機能を持ったものは残念ながらほとんどない。マイクロソフトの Word というワープロに機能を付加する形で自動的に SGML 文書を作成するツール等が出回っているが、未だ市場は小さく、官庁においては実際には全く普及していないのが現状である。このような状況においてなぜ、行政分野において SGML が注目されているのかを検討する。

**標準フォーマットとしての SGML の検討** 日本でも SGML は 1992 年に JIS X4151 として「文書記述言語 SGML」[3]が制定され、具体的な DTD は論文等のブック形式が 1996 年に JIS X0804 として「SGML 文書交換様式 (SDIF)」[4]が制定されている。これは後述する「白書等データベース」においても採用され、今後はこの形式での電子データの整備が進められるものと思われる。したがって各省庁はデータベースシステムの入札仕様書においては SGML 型式でデータの読み書きができることと明記する必要がある。ただし、SGML 型式をサポートしないワープロベンダーからのクレームを避けるためには SGML 型式を採用することの必要性を整理するとともに、ベンダーに対しても SGML 型式への対応を積極的に働きかけていかなければならない。

**特許庁における HTML と SGML** 特許庁では特許出願を一般的なパソコンから ISDN 回線で行うことが可能になる制度を平成 10 年度より開始することを決めたが、これは発明者や弁理士が特許出願書類の作成を一般的なパソコン (Windows95) で動作するワープロにより HTML 形式で作成した後に、それを特許庁が無償配布を行うソフトウェアにより特許庁での出願文献格納形式 (X フォーマット) に変換し、パソコンから ISDN 回線の特許庁に電子出願を行うというものである。従来までの専用端末ではなくパソコンから ISDN 回線を通じて特許出願を可能とするこの制度は電子出願用の専用端末を持っていない出願人の金銭的負担を減らすことが可能であるだけでなく、一般のワープロからも特許出願書類の作成が可能になるなど、メリットは大きい。一方で特許庁は特許文献を特許公報として CD-ROM により電子データで提供しているが、この電子データは HTML 形式ではなく、SGML 形式である。特許文献は各項目等のフォーマットが厳密に決まっている上にデータベースとしての有用性が高いので、SGML を採用しているわけだ。特許出願書類の作成では一般的なワープロが SGML 型式をサポートしていないために HTML 形式を採用したのであるが、データベースとしての情報の有効利用という面では残念なことであった。

**SGML 研修制度について** 前述のとおり現在では作成した文書を作成者が意識せずに SGML 文書として保存する機能は一般的なワープロには備わっていない。また、SGML 文書をデータベースとして有効に活用していくためには文書構造の設計等においてある程度専門的な知識が必要となる。そのような状況の中で行政情報の SGML 化を進めるための一つの政策として SGML 資格制度を設け、これを行政組織において積極的に採用することが有効である。すなわち、庶務係から政策立案部署にいたるまで文書作成を行う必要がある者は全て、この SGML 資格制度を取ることを推奨するのである。将来的には義務付けることも検討する。現時点では、SGML 資格制度自体はまだまだ検討の段階でしかないが、今後は情報処理資格と同等の資格制度として制度化されることが期待される。

**行政情報検索提供実験** 情報処理振興事業協会 (IPA) の新産業創造データベースセンターによる「行政情報検索提供実験」[7]が 1995 年から開始されている。本実験では、「目次による検索」及び「キーワードによる検索」ではあらかじめ HTML 形式に変換されたデータに対して検索・表示を行っているが、データの保存形式として SGML 形式を使用し、「シソーラスによる検索」では SGML 形式のデータに対して検索を行い、表示の際に HTML 形式に変換を行っている。本実験がパイロットシステムとなり、「白書等データベースの統一的な仕様」[6]が決定されたところであるが、データの保存形式として SGML 形式だけではなく、便宜上 HTML 形式との併用であるため、データ量の増大、変更に対する処理の増大等の課題があるといえる。今後は SGML 型式のみ格納することが望ましいわけだが、そのためには文書検索機能、HTML への自動変換機能の更なる強化が必要となる。

**白書等データベース** 総務庁が中心となって推進する「白書等データベースの統一的な仕様」[6]により、今後は各省庁において白書等のデータベースが整備されることとなる。対象は白書だけではなく、年次報告書や調査研究報告書等のほかにも、将来的には各省庁が保有する行政情報全般をデータベース化する。このデータベースの特徴としては霞が関 WAN を用いて各省庁での相互利用がなされること、データの格納型式を SGML 型式として、JIS X4151、JIS X0804 に準拠していること、全文検索や目次検索、SGML 型式での標準的なブラウザがまだ存在しないために HTML 変換機能を持つこと等が挙げられる。ここで課題となるのが実際にデータベースを構築し、行政情報の蓄積、運営管理を行うのは各省庁であり、各省庁は自ら保有する文書を SGML 化しなければならないという点である。その作業自体は大蔵省印刷局あるいは民間業者に委託することにより可能であるが、このことが各省庁内での定常の文書作成に SGML 型式の採用の検討を促す結果となった。

## 5. 利用研報告書の SGML 化に向けての実現可能性と課題の検討

実際に現在利用されている行政情報の一例として利用研報告書を取り上げ、これを SGML 化して公開する際の留意点、問題点について検討する。利用研では過去の報告書は過去数年分まではワープロ一太郎を用いて文書の作成を行ってきた。また図表については作成者が自ら一太郎等で作成することもあれば切り張りによって行うこともある。ここでは、テキスト文書だけは一太郎型式で保存されているものとする、SGML 化する手順は図 1 のようになる。

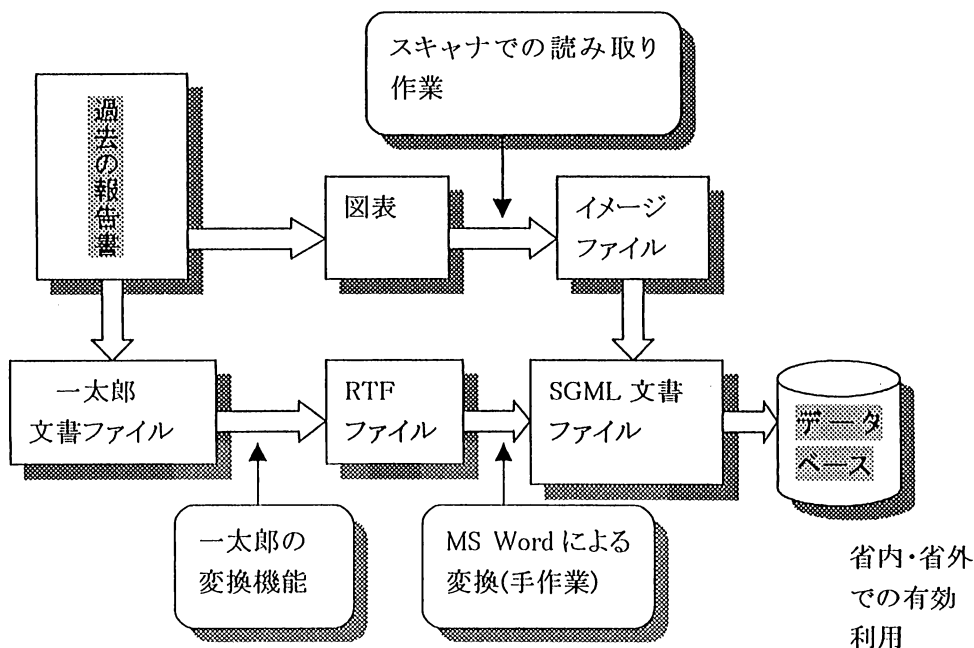


図 1 利用研報告書の SGML 化のイメージ

利用研の報告書を 1 年間分、SGML 化する場合の経費を算出するために実際にある出版業者に見積もりを依頼した結果を表 1 に示す。ここでは、報告書は全部で 20 冊、1 冊あたり平均 50 頁で図表が 1 頁あたり一つとしている。

表 1 SGML 化作業経費

項目	数量(人日)	単価(千円)	金額(千円)	備考
1)システムデザイン	2	100	200	
2)RTF 変換ソフト	1 式	30	30	ワープロの種類数
3)タグ付けエディタ	1 式	50	50	文書形式数
4)文書の SGML 化	(約)1,000 頁	3	3,000	
5)図版スキャン作業	(約)1,000 点	0.5	500	
			3,780	

ここでテキストが電子データとして残されていない場合には報告書のテキストを図表と同じようにスキヤナで読みとって OCR により電子化する方法と、キーパンチャーにより電子化する方法があるが、校正までも含めた作業時間を比べるとキーパンチャーにより電子化する方法が現実的である。その場合には費用は校正まで含めて 1 文字あたり 2 円である。

この結果、利用研の報告書を 1 年分 SGML 化するには約 400 万円かかることが分かったわけだが、この金額は現在利用研で報告書を 1 冊あたり約 200 部印刷するのに要する費用とほぼ同額である。簡単に言ってしまうと報告書を印刷せずに SGML 化すればかかる費用は今までと変わらないということになるが、実際にそうするにはいくつかの問題点が存在する。

まず、SGML 文書では現在表示ソフトウェアを持っている省庁はほとんどないために報告書を仮に電子データで配布しても研究員はそれを読み込むことができないので、前述の白書等データ

ースのように HTML 変換機能を作る必要がある。次に HTML 変換を行ったとしても報告書がまるまる HTML 型式でのみ配布された場合には現在のパソコンと HTML ブラウザでは紙で参照する方が見やすいという問題がある。これは情報の電子化には常について回る問題ではあるが紙文化からの移行にはユーザに対する細心の配慮が必要である。

## 6. 今後の課題

行政分野における事務処理等の文書の電子化は今後もより一層推進されていくであろう。しかしながら、電子データの標準化はまだまだこれからというところである。規格だけが先行して決められて実現がなかなか伴わないのが実状で、電子データの標準化を実際に進めるためには情報処理部門の職員の問題意識を高めることはもちろん、行政事務に関わる職員一人一人の意識改革まで進めていかなければならない。

## 7. まとめ

本稿では利用研の概要を説明した後行政情報とは何かを考察し、行政情報を電子データとして作成、交換、加工するためには電子データの標準化が必要であることを述べた。また実際に利用研の報告書を例として取り上げてその SGML 化を行う費用を算出するとともに行政情報の標準化に対する今後の課題を考察した。

## 参考文献

- [1] 根岸正光, 石塚英弘, “SGML の活用”, オーム社, 1994
- [2] 斉藤伸雄, “SGML アプリケーションの最新動向とその事例”, 情報知識学会誌 6巻2号 (SGML 特別号), 情報知識学会, 1997
- [3] JIS X 4151-1992, “文書記述言語 SGML”, 日本規格協会
- [4] JIS X 0804-1996, “SGML 文書交換様式 (SDIF)”, 日本規格協会
- [5] “行政情報化推進基本計画”, [http://www.somucho.go.jp/gyoukan/kanri/b\\_01.htm](http://www.somucho.go.jp/gyoukan/kanri/b_01.htm), 総務庁, 1994
- [6] “白書等データベースの統一的な仕様”, [http://www.somucho.go.jp/gyoukan/kanri/c\\_02.htm](http://www.somucho.go.jp/gyoukan/kanri/c_02.htm), 総務庁, 1996
- [7] “行政情報検索提供実験”, [http://www.cii.ipa.go.jp/sangyo/search/sch\\_j.html](http://www.cii.ipa.go.jp/sangyo/search/sch_j.html), 情報処理振興事業協会 (IPA)

---

<sup>1</sup> 通商産業省 工業技術院 総務部 総務課 電子計算機利用技術開発室  
電子計算機利用に関する技術研究会 (利用研) 事務局

<sup>2</sup> *Secretariat for the Technological Research Council on Computer Utilization (RIYOKEN),  
Computer Utilization Technology Office, General Coordination Department, Agency of Industrial  
Science and Technology, Ministry of International Trade and Industry*  
E-mail: [ikeda@riyoken.go.jp](mailto:ikeda@riyoken.go.jp)  
<http://www.riyoken.go.jp/ikeda/>