

CALS 時代の電子文書に向けて

Electronic Documents toward CALS era

若 鳥 陸 夫

要 約 紙による文書（線形文書）と電子文書（非線形文書，動的文書）との違いを述べ，その実現の形態の一つである SGML 文書の作成・改版・他系統（STEP，EDI，WWW）との関係についても触れる．更に，日本での CALS の概念の運用には，計算機側だけでなく，計算機を使う人間側（データ作成側）の進化が必要であり，日本語教育が鍵であることを強調する．

Abstract Some differences between paper documents and electronic documents in their properties are described in this paper. A linkage methodology among SGML documents, EDI document and STEP document is also discussed. Finally, it is emphasized that education of Japanese language is a key factor to succeed in CALS in Japan.

1. はじめに

計算機（computer）の援用によって，文書（document）の構成要素である文字，画像，映像，音響などをデジタル化して構成した文書を“電子文書”と定義することとする．この定義は，非常に広い範囲の応用を包含する．例えば，仕様書，書籍（絵本，小説，解説書など），報告書，学術文献，取扱説明書，プログラム，会議録，商品目録，小冊子，映画などが対象であり，記録媒体の種類及び様態を問わない．ただし，ここで言及する電子文書は，事業用に用いられる範囲だけに限定する．

電子文書は，その作り方から次の二つに分類できる．

a) CALS 以前（BC, Before CALS）の文書

CALS（Continuous Acquisition and Life-cycle Support）の概念を導入しない文書．閉じた社会の間だけで，文書を共有できる．在来の多くの文書がこれに分類される．

b) CALS 以後（AC, After CALS）の文書

CALS の概念を導入して，情報交換の相手の処理系が未知であっても，再生の確立が高い文書．この再生には，計算機種別，操作系（operating system）種別及び応用系（application system）種別を問わない．これを CALS 風と略す．

CALS 風と呼ぼうが呼ぶまいが電子文書の進化が起きている．それは，自己の管理範囲の中だけで処理できればよかった電子文書が，自己の管理範囲を越えた計算機との間の文書の共有・交換が計算機の戦略的な利用にとって有用になり出したからである．これは，計算機の小型化，低価格化，機能向上などによって利用者の爆発的拡大が基盤の一つにある．しかし，利用者の作成する文書は，今のところ，紙文明（パピルス文明）の形跡を随所に残しており，記録媒体がパルプから電子，光，磁気などの

電子媒体に変化したただけであって、格納空間の減少だけに留まっている。逆に万人が読める紙上の状態が誰も直接読めない密度に格納され、可読性という面だけは後退しているかも知れない。

ここでは、a) 紙の文書と電子文書のもつ特性の違い、b) 電子文書ならば具現化できる機能及びc) 人が意識して入力しなければ真の電子文書にならない3点を説く。

2. 電子文書の特徴

2.1 非線形な文書 (non-linear document, 投錨など)

紙の文書は、文字、絵、写真などを複数の平面(2次元)へ順序付けて収容している。この形態は、表現内容の群れが連なりとして配置されることから、“線形”と呼ばれる。この線形な文書は、前から後ろへ一本の糸をたどる風に読まれることを想定して書かれる(図1)。

前書き 複数の章(複数の節(複数の項(複数の号))) 後書き
備考 括弧内は、入れ子状態を示す。

図1 線形文書の想定された読み順

一方、紙の文書中の脚注、参考文献、索引などのように、平常、読み飛ばすかどうか読み手の選択になっている仕掛けを電子文書では積極的に使うことができる。このような文書は、文の途中の参照指定位置から別の参照内容へ飛んだり、用語解説などを本文から外して別位置に解説しておき、読者の選択によって、それらが提示される風に構成できる。このように、読み方が直線的でなく複数本ある(くもの巣状)文書を、“非線形文書”(non-linear document)と呼ぶ(図2)。

字句 ... 参照指定 ... 字句	1本目の線
参照内容(字句 ... 参照指定 ... 字句)	2本目の線
参照内容	3本目の線

図2 非線形文書の読み順

例えば、インターネットウェブの参照機構(投錨, アンカリング)を使用した文書がこの非線形文書の一つの実現例である。その投錨機構では、任意の文面(text)中に錨(anchor)の綱を結び付け、その綱の先にそれに関連する文面をおくという風に文書を作成し、読者が希望する位置の錨を上げることによって、その指定部分の説明が引っ張られて浮上して、見える状態となる(図3)。

非線形文書の基本概念は、数学で言う有向グラフ(directed graphs, 図4)を基盤としている。ある単位概念を“節”(node)とし、複数の節の間を方向性のある矢印

船(文面A)の錨の綱[操作：投錨(作成時), 抜錨(閲覧時)]
 水面(錨の周辺の隠蔽)
 錨 水草(文面B)へ食い込む]

図 3 錨の概念

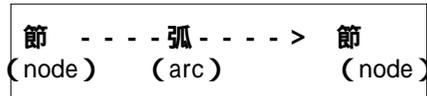


図 4 有向グラフ

(弧, arc)で結合して, 節間の関係を表現したものである.

弧を付けた状態は, 結合(link)又は超結合(hyper link)と呼ばれる. 超文書(hyper document)は, この超結合を応用した文書の意味で, 在来型文書と区分するために呼んでいることが多い. 非線形文書は, この有向グラフによる弧を網状に張り巡らした文書という意味あいでもある.

非線形文書の構成は, 線形文書と同じこともあるが, 錨を多用して, 詳細な事項, 補足説明, 用語定義, 引用文献などをその提示面から隠し, 閲覧者が指示した場合だけ, それらを提示するようにしてくれた方が多様な閲覧者の要望を矛盾なく入れやすい.

この形式を追及すると, 文書の読み方(経路)を複数想定して構成し, その各々の枠を埋めるという“下向き文書作法”(downward composition)になる. 非線形文書の場合の読む経路の内容群は, “文書断片”(document fragments)となる. これらの文書断片は, 文書全体のどこからでも参照される内容であるときには, その内容記述も参照元の文脈に依存しないように書かれなければ閲覧者にとって意味不明になる. すなわち, 最初に最上位の文書(輪郭)を作成し, その輪郭を想定読者に合わせて分割し, 分割された部分の子文書の共通項をくくり(因数化し), 必要な文書断片を作成する.

多数の文書断片を先に作成して, 後で枠組を作る方法は, “上向き文書作成法”(upward composition)と呼ばれる. この方法は, 悪い表現をすれば, “泥縄式”であって, 泥棒(悪文)を捕え(作成)てから縄どころか刑務所(上位文書)を建設する手順に例えられる. ただし, この方法も, 文書断片間関係を連想して新しい発想をもたらす場合などの無計画な文書に利用されるのは, 悪いことではない.

2.2 表示面の性質

電子文書を提示する表示面は, 紙の文書とは大きく異なる. 電子文書の表示面の利点と欠点は, 次のa)及びb)が支配的である.

- a) 利 点
 - 1) 消去して、再描画できる。
 - 2) 遠隔地でも瞬時に複製して見ることができる。
- b) 欠 点
 - 1) 表示面が固定寸法で、小さい
 - 2) 一覽性に限界がある。
 - 3) 解像度が低い。

したがって、電子文書は、これらの利点を伸ばし、欠点を補う方向を折り込んで作成される。例えば、紙であればページに表現したものは、一般的に削除できないが、電子的表示面は、同じ面を消去して、再描画できる。したがって、多数の絵を早く再表示すると動画として見せることもできるし、文書断片の関係を結合によって表現すれば、閲覧者との対話によって、次々と興味ある事項を提示することもできる。電子文書は、遠隔地に容易に複製できる性質から、同じ文書を地域を超越してその内容を共有できる。表示面の小ささによる一覽性の限界は、見出し部を熟考して作成することで、欠点の一部を補うことができる。

これらの表示面の性質から、電子文書向きの文書は、紙での交換を目的にしたワードプロセッサの対象文書とは異なる。例えば、表は、その縦横は画面で見える程度の項目に押さえる。スクロールしなければ見えない程の細胞 (cell) 数の表は、表本来の目的の一つである、“一覽性”を損ねる。絵の外枠の寸法が大き過ぎ、内部の情報粒度が細か過ぎてつぶれた線、主張が不明になった絵、原色を使い過ぎてお世辞にもきれいな配色といえない絵など身近に多々見ることができる。何のために、図・表を提示しているのかを再考した方がよいこともある。

工業系に多い図表の多用は、本文中に正確な記述がなく図表に語らせるという習慣があるがよくない。それは、複雑な絵は多義性があり、幾くともおりにも解釈できるからである。したがって、作者と読者とが同じ認識を有していないと、絵が本文の理解の妨害をすることも発生する。

これは、電子文書 (入れ物) の特質ではなく、紙文明にもあったが、紙による出版では、幾段階かの工程の途中で、清書又は書き直しが実施されるので、著者の原稿が直接読者にわたることの多い電子文書の方だけに著者の甘さ (アマチュア性) が露呈する性質がある。

3. 電子文書での様式

3.1 論理構造と割付け構造

電子文書は、編集・抜粋・合成・割付けなどが容易だから、多目的に使うことができる。例えば、論理構造及び割付け構造に分離すると次の利点ができる。

電子文書は、論理的な文書構造及び割付けの分れる単位の構造としておき、再生のときに提示面に合わせて割付けるというように、論理構造と割付け構造とを分離しておく、電子媒体が、紙、表示装置、CD ROMなどに变化する場合の元文書の一つで管理できるようになる。

ある種類の論理構造と割付け構造とを予め国際規格・機能標準によって限定して、

国際水準の再処理性をねらった規格に“開放型文書体系（ODA）及びその文書交換様式⁶¹⁾がある。この一番機能の低い水準は、ITU T 勧告の“ミクスドモードのファクシミリ”に例がある。ファクシミリだから、国、組織、団体、製造者などを問わずに地球規模での再処理性がなければならないので、厳重な規格になっている。他方、印刷応用とデータ倉庫への応用から着手された文書の表現様式に“文書記述言語 SGML⁶²⁾¹³⁾がある。この SGML は、言語を生成する超言語の性質をもち、SGML で記述された言語の規定を文書型定義（DTD, Document Type Definition）又は SGML 文書の頭部に前置した文書型宣言（DTD, Document Type Declaration）で表される札（付箋、タグ）を文書中に混ぜて記述する。

文書型を定義する毎に違う言語が生成され、共通化したタグを設計すれば、生成言語に非固有の処理系によって再生できるようにすることもできる^[617]。

文書の割付け構造だけを主体に表現する言語に、HTML（Hyper Text Markup Language に由来）がある。この様式の文書は、割付け、表示体裁などを記述して、利用者への提示を目的とした文書に加工できる。この HTML は、SGML によって、文書型定義として定義した言語であって、現在、インターネット（網間網）を介した電子文書での普遍化した記述様式になっている（図 5）。

```
<HTML >
  <HEAD >
    <TITLE >英字見出し</TITLE >
  </HEAD >
  <BODY >
    <H3 >文面本体</H3 >
    <A HREF = URL >投錨</A >
    第1図<IMG SRC = URL >
  </BODY >
</HTML >
```

図 5 HTML の構文例

HTML は、SGML 構文のうち、図 6 のような一部の構文を使って定義している。図 5 の“HTML ”、“HEAD ”、“BODY ”、“H3 ”、“A xxxx ”及び“IMG ”は図 6 中の開始タグ、“/HEAD ”、“/TITLE ”、“/H3 ”、“/BODY ”、“/A ”及び“/HTML ”は終了タグである。さらに分析すると、“HTML ”、“HEAD ”、“TITLE ”、“BODY ”、“H3 ”、“A ”及び“IMG ”は、図 6 中の共通識別子、“HREF ”及び“SRC ”は、属性の名前、“URL ”は属性値にそれぞれ対応している。HTML の共通識別子の機能（意味）は、HTML 規定によって定められている。HTML では、単純化などのため図 6 の文書型指定は使用されていない。この文書型 HTML の定義のように、SGML の応用側が自由に文書要素を定め、それを形式記述した文書型定義を作成することができる。

この HTML は、ワードプロセッサで出力形式を指定するだけでも生成することはできるが、見出し（目次）、用語索引などに関連付けるまで人手で実施するには、楽

文書要素	=	要素
要素	=	開始タグ?, 内容, 終了タグ?
開始タグ	=	< 文書型指定, 共通識別子, 属性指定並び >
終了タグ	=	</ 文書型指定, 共通識別子 ">"
属性指定並び	=	属性指定*
属性指定	=	(名前 "=")? 属性値指定
属性値指定	=	属性値 属性値表記
内容	=	(ここには, 構文解析可能な文面が記述される, 詳細省略)

備考 1. 複数の文書型の混用を許すCONCUR YESの場合だけ文書型指定がある。
 2. 記号は, 次の意味をもつ。
 , : すべてが式の構成の順でなければならない。
 : そのうちの一つが現われなければならない。
 ? : 任意選択(1回以下)を示す。
 * : 任意選択の反復(0回以上)を示す。
 : 引用符で囲まれた字句は, 終端記号を示す。
 = : 等号の右辺が等号の左辺を生成する。
 3. 挿入可能な空白などの位置は, 省略している。

図 6 HTML 定義に関係する SGML 構文 (抜粋)³⁾

ではない。そこで、見出しの自動作成、用語の自動投錨などができる処理系が求められている⁸⁾。

HTML で書けない論理構造を SGML で書けるように策定中の言語に XML (eXtensible Markup Language に由来) がある。XML は、SGML から派生した言語ではあるが、その利用目的のインターネットを介しての文書閲覧・再利用を主体とした“生まれ代わり言語”の様相である。SGML 文書は可視化の保証はないが、XML 文書は具体性をもつので、可視化の可能性が高い。

電子文書にとって一番重要な視点は、対象とする文書が形式性をもたない場合は、2進数表現 (ODA)、文字列表現 (SGML) のいずれにも適さない点である。このため、自由に制約なく作成された既存文書・遺産文書 (legacy document) は、その文書構造を分析して SGML タグを付けても、SGML 化の利点はほとんどない。すなわち、SGML 文書を作成する前に、できるだけ広い範囲の人と文書の様式を統一し、その合意こそ SGML のような形式言語を使う最大の利点である。その合意がえられた後、それを形式的に表現する手段の一つが SGML (XML, HTML) である。この人間側の標準化作業は、電子文書化の要求段階に相当する。文書の要件とは、デジタル化するだけでなく、その文書の役割、処理工程、転送対象部分、改定処理工程など、文書周辺の要因を十分勘案して、決めなければならないものである。

別な表現をすれば、広い適用対象の人間側の文書標準化の結果を反映した電子文書だけが、“CALS 以降の電子文書である”とも言える。電子文書の共有性を高めるために、文書型のひな型を用意して、それから派生させることによって、目的の派生文書型間の共通性を維持する研究も NCALS で行なわれた⁹⁾¹⁰⁾。

3.2 絵・写真・映像の扱い

絵・写真・映像の類は、従来、それらの符号化に共通性がなく、処理系依存というのが実態であったが、WWW での閲覧文書のように、絵 (点画): GIF, 写真: JPEG, 動画: MPEG などと符号化方式を絞ることによって、どの計算機でも再生でき、そ

の有用性が実証された。NCALS, DoD などでも、絵の符号化として、GIF のような私的様式を排除して、次の様式が採用されている^{[9][10]}。

- a) 3次元図形：IGES (Initial Graphic Exchange Standard に由来，米国規格)，
- b) 2次元図形：CGM (JIS X 4132，計算機図形メタファイル)，
- c) 動画：MPEG，
- d) 音響：Audio (JIS X 4109，開放型文書体系，音響内容体系)
- e) ソフトウェア開発情報：CDIF (CASE data Interchange Format に由来)
- f) 数式：TeX

3.3 文字の扱い

3.3.1 交換の哲学

CALS 時代の日本語文書といったら操作系の違いを越えて共通化されているものを指すという理想をもつ。これは、SGML 表現の文書では符号の違いは顕著であり、受け取った文書の前部 (SGML 宣言) をその操作系に合わせて書き換える必要まであるからである (図 7)。このように操作系の内部符号のままの SGML 文書を、筆者は、CALS 以前 (BC, Before CALS) の SGML 文書と呼ぶ。

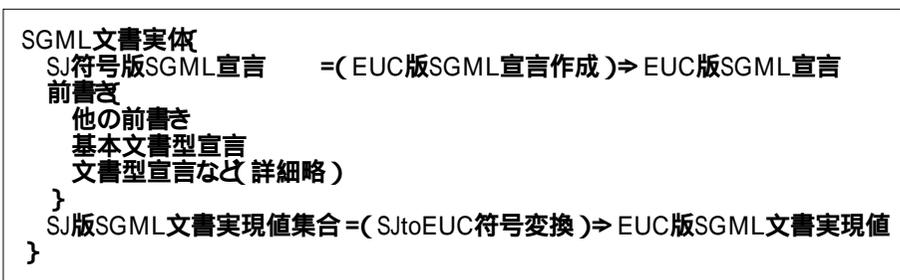


図 7 CALS 以前 (BC) の SGML 文書交換方法

3.3.2 文字集合と文字符号

NCALS では、複数組織 (操作系を制約できない) 間での電子文書の交換・共有を重要視したので、NCALS 共通符号を規定して、交換時点 (ファイル状態又は通信状態) では、それに従うこととされた。その符号とは、次のような構成である。

- a) 符号表左図形文字領域：ASCII 1 バイト英数字，又は SS 3 を頭に付けた補助漢字 (JIS X 0212 1990) の一文字。

備考 1. 漢字符号集合中の英数字の内，ASCII と同じ図形文字は，漢字符号によらず，この 1 バイト英数字とする。

- b) 符号表右図形文字領域：符号化漢字集合 (JIS X 0208 1997)

備考 1. 1 バイト片仮名文字は，使ってはならない。必ず，2 バイトの漢字符号領域のものに置き換える。

3.3.3 共通符号による SGML 文書交換

このような共通符号によって，文書を作成・処理することに仮定すれば文書の複写・転送・格納において，共通の計算機の操作命令群を使うことができる。

この NCALS 共通符号を使う場合の交換の模式を図 8 に示す。この方式によれば、SGML 宣言を含めて、操作系間での交換によって、SGML 宣言を取り替えなくてもよい。この性質をもつ SGML 文書を筆者は、CALS 以後 (AC) 時代の SGML 文書と呼ぶ。



図 8 CALS 以後 (AC) の SGML 文書交換方法

3.3.4 NCALS 共通符号の制定時の考え

半角片仮名及び全角英数字が、使用禁止になっていることが NCALS 共通符号の特徴である。これらの印字の寸法は、文字符号の属性ではなく、本来、表示体裁特にフォントの指定の問題として符号系から分離する最近の趨勢に従っている。在来文書をどうするかは課題があるが、それを NCALS 適合文書へ変換する時点で、表示体裁希望などを入れておくなどの解決法がある。いずれにせよ、輪郭線フォントを扱う時代では、全角とか半角というあいまいな表現でなく、ポイント制での指定に変わってきている。FORM が文字数だけしか扱わない COBOL, Fortran などでは、輪郭線フォントを制しきれない時代になっている。

SGML 文書などでは、フォントは後で印刷媒体が指定された時点で与え、一文字ずつ文字送り幅を計算するので、問題はない。

この文字の扱いで重要なのは、符号値ではなく、文字集合の方である。その理由は、文字符号値は符号変換によって処理できるが、文字集合を越えた文字 (外字) が使われた場合、符号変換の対象文字が存在しないからである。

NCALS では、UCS も考えられたが、一旦、文字集合を広げた文書は、狭い在来の文書系では、再生の可能性が低いこと、入力系がすべての操作系に備わっていないことなどから採用されなかった。NCALS としては、UCS の適用を実証するのが目的ではなく、CALS 流の文書方式を適用する応用域の実証だったからである。もう 5 年もすれば、UCS の搭載計算機が増加して、環境が UCS に有利になるかも知れないが、1998 年現在、多国語を支援する文字変換系は市販されていない。更に、世界中の郵便系が 2 進数の宛先を解釈するように進化するまで、2 バイト又は 4 バイトの文字符号を文字として扱ってはくれない。広域での符号の問題は、基盤として、すべてが揃わないと使えないのである。閉域系が UCS に除々に対応し、LAN が 2 バイト文字の宛先を解釈し、更に、インターネットが 2 バイト文字による宛先処理へと波及していくため、これには、どうみても 10 年単位の時間が必要である。それは、通信系の処

理プログラムだけでなく、回線速度が飛躍的に向上し、図形、動画などに加えて2バイト文字を転送しても、費用、時間のいずれも問題視されない程度にならなければ、普及しないからである。

4. システム構築

4.1 データベースとの連動

電子文書とデータベースとの連動は、比較的素直にできる事項ではあるが、電子文書(例:SGML文書)たる機能を強調しようとするとは容易ではない。例えば、SGML文書が文書断片を直接参照できるからといって、それら文書断片を直接変更し、変更管理したとすると、その特別なSGMLデータベース単体系で運用するにはよいが、作成方法の違う他のデータベース系との連動を考えるととたちまち破綻する。1998年現在、複数種類のSGMLデータベース管理系間で変更同期、変更情報交換、文書断片の検索などに対応した処理系は販売されていない。いわんや、関係データベースにしる木構造データベースにしる、ファイル単位又は同一長さの記録単位による処理と、SGML文書のような、任意長さの断片を特徴とする処理とでは、折り合いが悪い。

筆者らは、文書断片の変更を同期させるための情報を文書の前部に保持する手法を提案してきた^[7]。

4.2 STEP との連動

製品モードに関する国際規格群STEPの応用規格(設計における形態管理)からSGML文書などの違う形式を参照したい要望がある。STEP系の記述言語EXPRESSとSGML系とでは、目的、機能、育ちも違うため、相互乗り入れしたところで、処理系は二手に分れて開発できるものの、利用者側は、二つの規格の特徴をわきまえた運用をする必要がある。したがって、この連動は、木に竹を継ぐような行為であるという冷静な見方をしている。互いの得意分野を明確にして如何に補うかは、今後の課題である。

NCALSでは、当面の経過処置として、SGML文書中に利用者定義属性を設けてSTEPから参照されている文書断片に印を付けられるようにした。しかし、その印がついたことによって、STEP側とSGML側が自動的に変更同期されるわけではなく、変更指示などを体系的に作成する支援をするだけである。

STEP可視系からSGML文書を見る場合、指定された部分の仮割付けを行なって表示する簡易表示系は、割付け体裁が単純でよければ、作成可能であろう。しかし、STEPに限らないがSGML文書自体をそれ以外の処理系によって、直接、再生しようとする、変更箇所の傍線修飾、紙上等しい割付けなどを要求した途端に全文書の最初からの割付け処理に突入するなどの問題がある。大きな文書の割付けは、計算機の前で待てる時間では終了しないこともある。このあたりは、どの程度の情報粒度に対して、どういう処理を作用させると効果的かという要求分析が最も重要である。

4.3 EDI との連動

電子商取引の仕様指定のために、EDIの付属文書として電子文書が添付でき、再生処理できることが必要である。その添付物は、図面、写真、仕様書などが対象で

ある。この電子商取引の伝票の交換手順が EDI (Electronic Data Interchange に由来) がある。その EDI の添付物として、SGML 文書が添付できれば、文面 (text)、図面、仕様書などを詳細に指定できることになる。その一つの方法として、EDI 側から文書識別子 (文書名、文書断片) によって、SGML 文書中の断片 (章・節・項・号、図表など) を指定できる実験が NCALS で行なわれた。その試みでは、SGML 文書の論理構造に EDI 属性をもたせられるようにし、EDI で引用されている論理構造部分に印を付ける方法が行なわれた。この属性は、SGML 文書中では、どの EDI 伝票で引用されているかの指標となる (両方向指標式)。はん (汎) 用的な仕様書・図面などで、多くの伝票から引用される場合、この属性を使うと、引用を追加する度に原文書の属性値も追加することになるので、特定の伝票とだけ強く結び付く応用 (特別製作品など) に限って利用できる。逆に多くの伝票から引用される文書は、EDI 側からの一方向だけ指標付けされるのが普通である。応用によっては、このような密結合方式ではなく、参照対比表を別途用意し、それを介して結合先を獲得する粗結合方式による連係を薦める。

なお、EDI 伝票の作成時に仕様書中の一部の記述を複製したいことがあるが、それは入力時点での局所応用であって、記述内容の一貫性の保持のためには重要な機能ではあるが、対象とする仕様書を EDI 伝票添付物として転送するための規約と送受信者がその添付文書の一部へ自動指標付けすることとは別と考えられる。

5. 在来文化との比較

5.1 紙文化の利点の助長

人類は、石器文化時代では、情報の記録・保管・伝達が大変困難であったが、紙文化に進化してから、記録の容易さ、保管場所の寸法の縮小、軽量化、複製の容易さ、可視性、一覧性、きれいさなどの恩恵に浴してきた。

5.2 紙文化の欠点を電子で補完

紙文化の利点の他方では、紙の素材が木のパルプであることによる保存期間の制約 (風化、虫食い、破損など)、即時転送性のなさ、文書断片の再利用性の低さなどの欠点がある。これに対して、文書を電子化 (デジタル化) することによって、通信回線による即時閲覧性、再利用性の向上、保存期間の延長、文書の検索性の向上などが実現できる。しかし、デジタル化文書は、一般的に可視性が悪くなる。表示面の経済的な制約などから新聞紙に匹敵する一覧性は当面期待できない。更に、文書内の高速ページめくりに替わる検索手段が求められる。人が文書中の記述箇所を記憶にたよって探す場合、こんな割付のところであったとか、あんな絵の付近であったという淡い記憶に一致する箇所を探すことがある。この仕掛けが電子文書で実現できると、便利になることがある。このあたりは、現在、研究途上にある。

5.3 人間世界の習慣

CALS 時代には、人の習慣も変わらざるをえない。

5.3.1 面談なしでの事業の推進

CALS 時代が、通信を仲立ちとして、遠隔地、面識なしの人など通常組織の外と事業を連係できる頃と仮定すると、電子文書を真に活用するためには、人間側が次のよ

うに進化しなければならない。面談なしの事業推進は、如何に自分の考えを通信線の先の人へ伝えるか、何を考慮すれば伝わるかなどが基盤になる。

a) 閉鎖社会の共通認識を除外した文を書く。

閉鎖社会とは、ここでは、同じ組織、同じ業界、同じ地方、同じ国、同じ企業集団、同じ官庁などのような閉じた社会で基礎知識、常識、分野知識、用語、文の作り方などが独特な社会を表す。CALS時代には、これらの境界を越えても通じる文書を作成する必要がある。従来の身の文書は、そのほとんどが、ある特定の社会の常識にたよって解釈させる文面になっている。それ以上の配慮は、従来、必要性が低かったからである。逆にその社会の言葉でなければ、その特定の社会では通じないからでもある。しかし、文書を受信するのがその社会の外の人になる可能性が高まるのがCALS社会である。そう考えるとき、たとえ日本語で表記するとしても、万人に通じる日本語とは、どのあたりかという想定が問題である。その基準として、公の文書については、“公用文の書き方”が日本国内閣から発令されており、義務教育を終了していれば、文面を読む水準に設定されている。国は、自国民の誰もが平等の可読性、正しい理解及び行動ということが基本になっているからである。

この水準にそって、JISなどは書かれている。一見読み難そうであるが、正確性を追及している場合のうっとうしさであって、構文、漢字集合、仮名遣い、接続詞、副詞、用語などが制約されており、多義には解釈できない文面で書かれていて良文とは言いがたいが、最低限の正確さは保持している文である。現在、読者が目にしている、この論文は、その公用文の書き方に従って記述しようと努力しており、読む側では別段の奇異さを感じられないであろう。それは、記述対象が規格ではないから、文末が“とする”などの規格書の常用句を使っていないからである。

この論文の読者は計算機技術者だけではないと想定しているので、この論文は次の注意を払っている。

- 1) 極く少数の英字略号だけしか使用しない。
- 2) 稀にしか片仮名による外来語表記をしない。
- 3) 主語は必ず明示し、それが代名詞ならばその元を明示する。
- 4) 漢字は常用漢字に限定し、送り仮名は、名詞の複合語の場合削除している。
品詞によっては漢字が使えない部分を仮名書きとし、かつ、漢字が使える場合は、必ず、漢字にしている。
- 5) 読点は、横書き読点(“,”)とし、接続詞は、“あるいは”を用いない。
- 6) かぎ括弧は、用いない。

これらの公用文の制約に従ったところで、指摘されるまでは、その制約によって不自然とは感じにくいであろう。

b) 正確に書く

CALS時代の電子文書は、面談を前提にしなくても言いたい事項が文面に正確に表現されている必要性が高くなる。

この事項は、普遍的で、すでに実施していると自負されているかも知れな

いが、次のような観点を加えるとどうであろうか？

- 1) 用語の意味は、国語辞書に記述されているとおりにしか使用しない。
- 2) 片仮名の外来語は、中学生でも知っている範囲までをその意味だけに使用する。
- 3) 新語は、定義を付記する。
- 4) 特定の社会だけでしか通用しない方言は、使用しない。
- 5) 主語を省略しない。挿入句は、読点で挟んで明確にする。
- 6) 例だけ示して、すべてを説明したと思わない。
- 7) 図・表の説明があり、それらへの指標もある。

c) 正確に読む

文面を正確に読むのも、読者が考える以上に難しい。それは、内容の難易度ではなく、その文面に書いてある事象をその文面のおりにだけしか解釈しない習慣の形成についてである。筆者は、JISの一文ですら正確に読めないのが一般人の水準ではないかと思っている。まさかと思われるから、次に例を示す。

例文 公用文：Aは、B及びC並びにD又はEから構成する。

模式：A=(B and C) and (D or E)

どうであろうか？ 極端な例であるが、模式のおりに一意にだけ読めたであろうか？ 制約日本語は、その約束を知らなければ読めないのである。

6. ま と め

CALS以前(BC)からCALIS以後(AC)への電子文書の進化の技術的側面、日本語による技術文書であればすべて共通に処理できる世界の構築、CALIS以降の人に必要とされる電子文書の書き方の変化などを概観した。それらのうち、技術側面の共通化は、産業界の見識のある少数の専門家のもとで、次世代への遺産としては是非とも実現したい事項である。後者の制約日本語の正確な読み書きは、全国民が一挙に努力を要する事項であるから世紀単位の努力がいるであろう。すなわち、産業日本語教育をしっかりと行なう方策を実行しなければ解決しない課題である。

これは、高校、大学などでの産業日本語講座の新設などをするか、大学入試問題の回答が制約日本語で書けるか否かを問う形にすると比較的早く広がるかも知れない。

いずれの方法をとるにせよ、日本語をデジタル時代にも真に使えるようにするためには、産業日本語教育の再考が必要である。

謝辞 この論文は、通産省の依託研究を受けたCALIS技術研究組合での研究活動でえた知見をもとにしている。文末を借りて、通産省及びCALIS技術研究組合の関係者に感謝する。

参考文献 備考 末尾のくっこのJapaneseは、制約日本語記述の文献を示す。

- [1] ISO; ISO 8613 1~14, Open Document Architecture (ODA) and interchange format, 1994 [JIS X 4101~4114 (開放型文書体系) が相当]
- [2] ISO; ISO 8879, Standard Generalized Markup Language (SGML) 1986.
- [3] JISC; 文書記述言語 SGML, 日本規格協会, 1992 (Japanese)

- [4] 若鳥他 ; “ NCALS はん用文書型定義の開発方針と課題 ”, 情報処理学会第 53 回全国大会論文集 4 343, 情報処理学会, 1996 (Japanese)
- [5] WAKATORI, Rick ; “ Development of Generic document definition in NCALS ”, Proceeding of CALS EXPO '96, Long Beach California, USA., 1996.
- [6] WAKATORI, Rick ; “ Development of Document Type Definition for Interchange of Technical Documents between Companies ”, Proceeding of CALS Japan '96, Tokyo, 1996.
- [7] 若鳥他 ; “ SGML 文書と差分情報とから指定の版を合成する実現方式 ”, 情報処理学会第 54 回全国大会論文集 3 329, 情報処理学会, 1997 (Japanese)
- [8] 若鳥他 ; “ 自動投錨を意識した NCALS 平文文書から ISO-HTML との変換 ”, 第 55 回全国大会, 情報処理学会, 1997 (Japanese)
- [9] WAKATORI, Rick, et al. ; “ Derived DTD in NCALS ”, Proceeding of CALS EXPO USA. '97, Orlando, Florida, USA, 1997.
- [10] WAKATORI, Rikuo ; “ Creating an infrastructure Using Derived DTD Scheme ”, Proceeding of CALS EXPO International '97, Tokyo, 1997.

執筆者紹介 若鳥陸夫 (Rikuo Wakatori)

1961 年 3 月芝浦工業大学工学部電気工学科卒業。同年 4 月日本ユニシス(株)入社。研究・開発部門で文書処理関連技術の研究に従事。この間 ISO/JIS 委員として従事。1997 年 7 月情報技術標準調査会標準化貢献賞受賞。1995 年 7 月から 1998 年 3 月まで CALS 技術研究組合へ研究員として出向。現在、日本ユニシス(株)新事業企画開発部所属。横浜市立大学経営研究科非常勤講師。技術士(電子・電気)。
Email : Rick.Wakatori@unisys.co.jp