

## (仮称)x-file概念に基づく電子原稿用紙仕様の概念モデル

水島賢太郎

神戸女子短期大学・初等教育学科

Tel 078-303-4700 Fax 078-303-4703

e-mail mizusima@kwjc.kobe-wu.ac.jp

### 概要

ワープロやエディタといった文書作成支援ソフトは既に長い歴史があり、成熟しているとも言える。ところで、ここ数年、XMLに関する書籍が普通に書店に並ぶようになったように、ネットワークと安価なパソコンの普及は、従来の紙メディア出力を中心とした文書教育の再考を迫っている。この面から既存のワープロやエディタを見ると、正しい電子化文書教育には過剰なものや逆に足りないものが見られる。ここではその背景を振り返りながら、電子化文書教育にふさわしい電子原稿用紙ソフトの仕様について考察したものを報告する。なお、このソフトは主たる目的を教育においているが、考察内容はより一般的な文書作成支援ソフト開発も考慮している。

### 1 はじめに

#### 1.1 本研究テーマ

情報教育の目的の一つに電子化文書を「正しく」かつ「効率的」に作成すること能力の育成がある。本研究の目的は、この教育目的をサポートする文書作成支援ソフトの開発にある。このソフトを「電子原稿用紙」と呼び、現在その仕様を教育目標との観点から考察している。なお、このソフトは、実務でも利用でもより一般的な文章作成ソフトの開発も視野においている。

#### ●「正しく」かつ「効率的」とは

本研究での「正しく」かつ「効率的」という意味は次のことを意味する。

まず「正しく」とは、電子化された文書ファイルがコンピュータによって適切かつ効果的に処理・加工・再利用できる形式として作られていることを意味し、個々の文章や文書全体の意味内容的正しさではない。このことを文書教育の立場からいえば、作成された電子化文書が、論理構造と物理

表現とに明確に分離されており、かつ XML 文書における整形式(Well-formed)としての文書レベルとして作成できていることを意味する。

次に「効率的」であるが、ここで効率には2つの意味を持たせている。

その一つは、先の示した「正しさ」に対応したもので、文書インスタンスの2側面、すなわち論理構造と物理表現を明確に意識して作業できることを意味する。文書作成時にこの2つを明確に意識して作業すると、作業の純度が上がり、結果として文書作成の効果が上がる。また、文書作業を共同作業として行う場合にも、原稿作成作業と物理表現のデザイン作業とに分担できることにより、作業時間の軽減につながるからである。以上は、文書作成者という人間の問題である。

いま一つは、別種の効率問題である。通常、コンピュータを使って作業する場合、常にコンピュータ環境の問題が付きまとう。たとえばシステムクラッシュに対するバックアップや文書のバージョン管理、さらに異なった作業環境での同期問題などである。これらは本来の文書作成業務ではなく、むしろシステムがどれだけカバーできるかという問

題である。これをシステム的効率化と呼ぶ。

とはいって、システム的効率の問題をユーザーが理解しておくことは重要である。実際、一般家庭へのパソコン普及に伴い、同期等の問題が頻繁に見られるようになって来たが、その原因の理解は文書作業(より広くは電子化ファイル一般)の効率に関わってくる。

以上をまとめる電子原稿用紙の仕様策定は次の図1のような異質なフェイズを考えることになる。

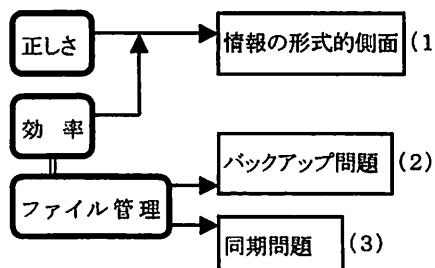


図1 仕様の3フェイズ

図1の内、(1)はユーザーが直接操作するインターフェイス設計に関わり、(2)(3)の本質部分はコンピュータが裏で補助する仕組みに関係する。

これら3フェイズはどれもが重要な問題であるが、紙幅の関係もあり、本稿では主に(1)を議論し、(2)(3)については簡単に触れるに留める。これらについては参考文献<sup>[1][2]</sup>を参照していただきたい。特に[2]は本報告の別章という性格を持っている。

## 1.2 電子原稿用紙ソフトは必要か

さて電子化された文書の作成を支援ソフトにはエディタ、ワープロ、Web ページ作成支援ソフト、また最近とみにその重要さが指摘されているXML 文書作成ソフトなど多数のものが開発・研究されている。とりわけ、エディタやワープロはもはや成熟ソフトという段階にあろう。しかし電子化文書教育の面から既存のエディタやワープロ、あるいはアウトラインプロセッサーを見ると次のような問題があるといえる。

### ● エディタ

エディタはいわばタイプライターの直系の

子孫で、親の持つ素顔を磨き上げるという形で進化をしたものといえる。つまり、エディタは本来、無構造の文字列作成機械の電子化版であり、したがって作成される電子化文書の構造化といったことを支援しない。つまりエディタは電子化文書教育のツールとしては自由度が高すぎるのである。

また、これから電子化文書で重要なとなるリンク機能を持たないことも、リンクがもたらす問題、つまり、同期問題を取り扱えないということもある。

### ● ワープロ

ワープロも本来はエディタ同様、タイプライターの子孫であったが、DTP 機能の追及のあまり、文書作成が目的なのかデザインが目的か分からなくなり、文書教育を混乱させている。しかし、ワードの象徴されるように、ワープという商品には文書作成現場の雑多な要求を貪欲に取り入れているという意味で、文書に対して人々が求めている機能を知る格好のものになっている。

### ● アウトラインプロセッサー

論理的な文書作成に焦点を絞れば、アウトラインプロセッサーが最もふさわしいといえる。問題は、ほとんどのアウトラインプロセッサーがエディタ同様、物理表現が貧弱なため、単体として論理構造と物理表現の差異を十分教育できないことである。

## 2 ソフト仕様のための考察

昨今盛んになっているWeb ページ作成教育は情報発信能力(デザイン力)の文脈で語られることが多いが、その言語のルーツである(情報関係者にとっては常識であろうが)SGML までさかのぼると、その開発目的が科学技術文明の進展に伴う膨大な文書情報処理問題であったと分かる。その議論は古く、SGML のもととなったIBM 社のGML は1963年だといわれている(ISO 規格は1986年)。

その間、コンピュータ環境の普及は著しく、それを受けてかネットワーク時代にふさわしい XML の議論が最近とみに激しくなってきた。時代は紙メディアの清書的ワープロ時代から確実に変わろうとしている。しかし、未だ教養教育のレベルでは十分扱われていない。というより、その事実すら知らないようなものも多い。

認知科学の研究が示すように、アーティファクト(道具)は人間の思考のあり方と深く関係している。学びもまた思考の一形態であるとすれば、適切な操作インターフェイスを持った電子化文書作成支援ソフトを開発し、初中高校段階から教育に用いれば、たとえば XML 普及を進める上で好ましい社会環境を作れるといえる。

以下、電子原稿用紙の仕様を固めるため必要となる事柄を簡単にまとめておく。

## 2.1 論理表現と物理表現

「作・文書教育」の実践を行ってきた経験からいって、物理表現の指導はそう難しいものではなく、「文書の見栄え、化粧だ」といった感じで説明すればほとんどの学生は理解してくれる。しかし、論理構造は厄介である。それは「論理」という言葉と「構造」という言葉が持つ日常的イメージと電子化文書で言われる意味とにある種の乖離があるからである。そこで、筆者は次のような定義を行って説明している。

### (1) 論理表現:

文書作成者(情報発信者)が発信した情報に込めた意図

### (2) 物理表現:

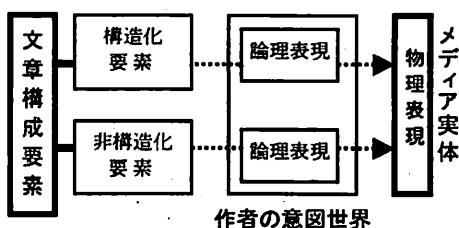
作者の意図にふさわしい見栄えでメディアに具体的な物理実体を与えたもの

### (3) 文書構成要素:

文書は幾つかの構成要素からなるが、その要素は2つに分かれる。一つは、文書の階層構造という骨組みに関わる要素、つまり目次として抽出されるような要素で、これを文書の構造化要素する。構造化要

素でないものは非構造化要素となる。

2つの異なる要素はそれぞれ論理表現をもち、その論理表現に応じた物理表現で形ある文書インスタンスが出来上がるというわけである。



このまとめ方は建物に関する私たちの常識的な言葉とも符合するため、学生には分かりやすいものとなる。すなわち「家は柱や梁といった構造体で骨組みが出来ている。それに扉や窓といったものをくっ付けて出来上がる。文書も同じである」というわけである。

電子原稿用紙のインターフェイスは、これら文書要素と論理表現と物理表現を可視化すること目的としている。

## 2.2 文書モデル再考

電子原稿用紙に限らず、息長く生き残り、かつ分かり易い操作インターフェイスを持ったソフトを作るには、そのソフトが目的とする内容を適切に表現した抽象的なモデルが必要となる。電子原稿用紙は文書作成支援ソフトであるから、まず文書の抽象化モデルを作ておく必要がある。筆者は先に HTML とワープロの同時教育のための文書論理構造形式モデル(積み木ブロックモデル)を提案した<sup>[3]</sup>。

このモデルは文書の形式上の最小単位を段落とし(レベル1)、その段落が(入れ子構造を含めて)多数集まって文書が出来上がっているというモデルである。このモデルでは文書構成要素として見出しや引用のように單一段落(レベル2-イ)と、箇条書きや HTML の「表」のように複数構成段落(レベル2)とに分類した。今回の電子原稿用紙のもととなるモデルを考えるにあたって、レベル2の

「表」の取り扱いを、図2から図3に変えた。これにより、電子原稿用紙の構造化文書作成モードでの「表」の記述は HTML のように構造まで書かず単に「表8 x-file のブロック構造」といった記述に留めることなる。

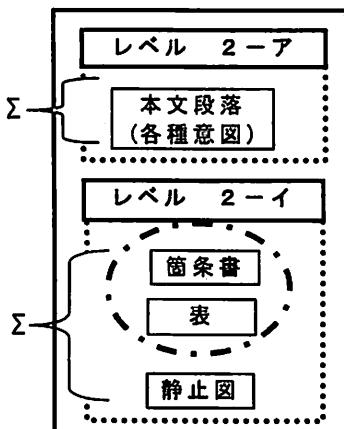


図2 積み木ブロックモデル

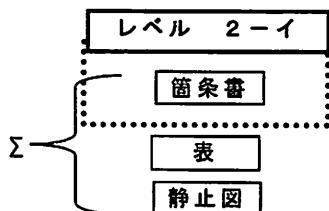


図3 レベル2-イの変更

### 2.3 電子原稿用紙の利用環境

「はじめに」で述べたように、本稿での報告は、1章の図1の(1)で示した文書情報の形式的側面を中心とした操作インターフェイスの仕様(デザイン)に重点がある。その際、情報の形式的側面と操作インターフェイスとの対応関係が単純なほど本研究の狙いである論理表現と物理表現の学習効果が高いと思われる。

そこで先ず電子原稿用紙が扱う文書ファイルのブロック構造について簡単に示しておく。

なお、電子原稿用紙で扱う文書ファイルのブロック構造には他のすべてのファイル同様、ファイル管理等のデータ領域がある。このデータ領域には電子原稿用紙の利用に関わるアイディア、つまり

図1の(2)(3)で示したバックアップと同期問題の解決に対するアイディアが関わっているため、電子原稿用紙利用環境について先ず簡単に紹介しておく。

以下、本稿の記述を簡潔にするため、幾つかの用語とその機能について紹介する。

#### 2.4 x-file、x-folder、v-router とは

電子原稿用紙が扱う文書ファイルを「拡張された文書ファイル(extended-file)」、それを管理・運用するための仕組みを「拡張されたフォルダ(extended-folder)」、「仮想的なルータ(virtual-router)」と仮称し、それぞれを、x-file、x-folder、v-router と略記する。

##### 1. x-file

x-file はそれ自体が独自のメソッドを持った一種のオブジェクトである。つまり、それがソフトで扱われたり、また複写や移動といった操作を受けた場合、自分の状態を他の x-file とやり取りしながら利用環境を整えるという「ある種の自立的なロボット」のようなファイルである。

x-file の実体は通常のテキストファイルのデータブロックに独立したヘッダーブロックとデータブロックもつた構造をテキストデータとして埋め込んだものである。この仕組みは HTML や XML、あるいは電子メールファイルの仕組みと同じである。

x-file をこのような仕様にした理由は次の通りである。

##### (1) データファイル側からソフトの仕様を規定

文書ファイルとしての x-file は今回提案する電子原稿用紙のための文書ファイルとして想定していない。つまり、マイクロソフト Word の文書ファイル(.doc)があくまで Word という固有のソフト向けであるのに対し、x-file は仕様をオープンすることで、幾つもの文書作成支援ソフトの開発されることを期待している。ちょうど HTML の仕様と各種ブラウザの開発関係と同じである。

## (2) ファイル管理データやソフトカスタマイズ情報等の柔軟な受け入れ

ほんの 10 年前のパソコン環境と今日を比べると、CPU の性能向上やメモリーやハードディスク容量の巨大化、そしてネットワークの高速化には目を見張るものがある。このような状況を考えると、たとえば Windows といった OS や応用ソフトのあり方を考え直してもいい時期に来ているのではないか。すなわち、メモリーやハードディスクが高価であった時代に考えられた工夫、アクロバット的テクニックをやめて平易なものにやりかえるというアイディアである。

たとえば Windows ではアプリケーションの設定をレジストリで行うか INI ファイルで行うかという議論があり、どちらにもそれなりのメリットデメリットがあるが、もし INI ファイルを個々のデータファイルが持てばどうなるのだろう。この場合、システムやアプリケーションソフトのあり方は単純になるはずである。つまり、データファイルが OS やアプリケーションソフトの機能の一部を肩代わりするとい

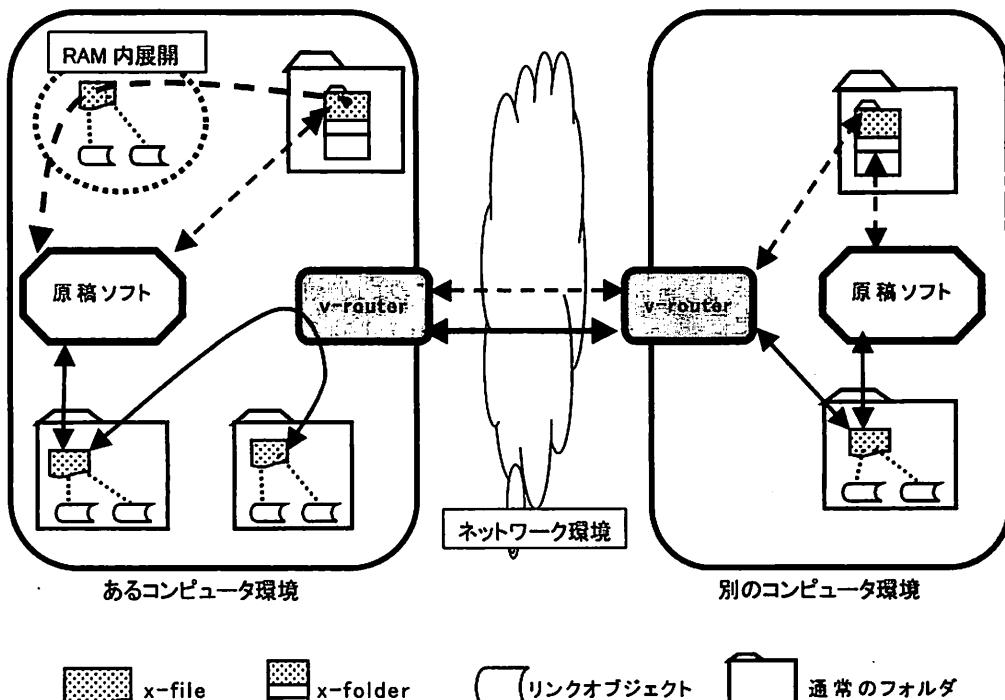
う考え方である。この方法は明らかに無駄が多いが、INI ファイルのサイズなど、今（未来）の環境からみればたいしたものではないだろう。

### 2. x-folder

x-folder は、電子メールの Mbox の考え方を使って、複数の x-file や外部オブジェクトを一つのアーカイブとして扱おうというものである。すなわち、電子メールソフトではあたかも個々のメールを一つずつ扱っているように見えるが、その実体は一つのファイルに過ぎず、適当な管理テーブルファイルを用いて、その一部を切り取ったりすることが簡単にできるということにある。

x-folder も x-file 同様、後述の v-router とのやり取りを通して自立的に自己管理できる構造を持たせる。

図4 電子原稿用紙活用全体イメージ図



### 3. v-router

インターネットにおけるルータと似た機能を果たすソフトで x-file や x-folder は v-router を介して他の x-file や x-folder とやり取りを行う。また、v-router は URL のようなアドレス管理機能をもつており、そのアドレスと x-file 名を指定することにより、ネットワークを介して x-file の処理が行えるようになっている。v-router は、いわば x-file を参照しようという場合の総合窓口といったもので、バックアップやファイルのバージョン管理、同期問題の解決を目指している。

x-file、x-folder、v-router、電子原稿用紙を含めた文書作成の環境イメージは図4のようなものである。

#### 2.5 x-file のブロック構造と電子原稿用紙のインターフェイス

右の図5は、x-file の詳細なブロック構造を示している。図の「原稿」部分は、XML の整形式の文書ファイルに対応している。「文書インスタンス」は、「原稿」にスタイルシートを加えたものを意味する。

各ブロックの機能の概略は次の通りである。

##### (1) ヘッダーブロック

バックアップやバージョン変更、あるいは同期のためのデータと奥付のような文書メタデータが置かれる。

##### (2) データブロック

文書内容のテキストデータ。図表等のオブジェクトは、本文中の「図3 ブロック構造図」といった文字列とリンクを通して結ばれている。

##### (3) プロパティブロック

データブロックの各段落の種類や文字列強調といったものスタイルを置く対応テーブルと物理表現を与えるスタイルデータブロック。XML の整形式な文書のタグセットもここにおかれる。

##### (4) オブジェクトブロック

図表等のデータを置くブロックで、置き方には直接ファイルに添付する方法と外部

データへのリンク情報だけのものがある。添付するものがバイナリファイルの場合は電子メールの添付ファイル同様、MIME でテキスト化してある。

##### (5) メソッドブロック

v-router とのやり取りといった x-file の自立性に関するプログラムというユーザーが扱えないプログラムとユーザカスタマイズ用スクリプトが置かれる。

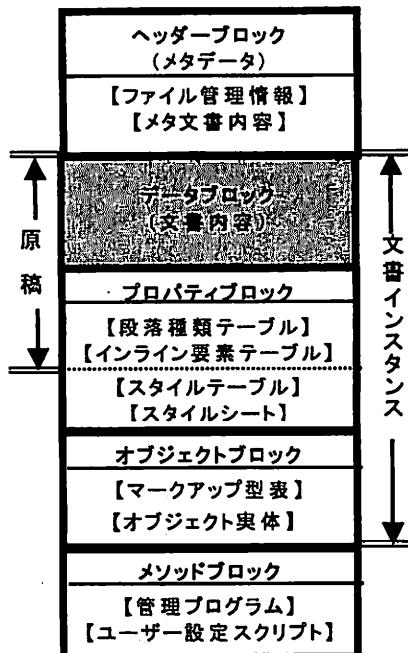


図5 x-file のブロック構造

各ブロックは、ある意味で予約領域的扱いが出来、それを扱うソフトによって具体的な運用が決められる。このような大雑把な仕様にしているのは、2章で示したように「データファイル側からソフトの仕様を規定」するという考え方から来ている。

たとえば、データブロック以降がヌルなら文書ファイルは単なるテキストファイルで、それを扱うソフトはエディタの一種となる。逆にプロパティブロックに XML の整形式だけを置き、出力ファイルとして「文書インスタンス」だけとすると、ソフトは XML 文書作成支援ソフトというまさに物理表現力を持たない原稿用紙となる。また、XML ファイルに対応し

た DTD とスタイルシートを置き、電子原稿用紙に高級な XLL-FO をアドオンすれば、電子原稿用紙は DTP ソフトにもなる。次章で説明する電子原稿用紙もその例といえる。

### 3 電子原稿用紙のインターフェイス

#### 3.1 オープニング画面

昔のワープロ専用機や一世を風靡したワープロソフト「松」では、オープニング画面が何を行つかのメニューになっていた。電子原稿用紙を教育的に考えれば、この仕様が良いように思われる。すなわち、メニューから新規作成と既存のファイル編集の選択から始まるというわけである。

新規作成の場合、まず文書テンプレート選択ウインドウが開いてくる。そこから適当なテンプレートを開くと、著者名、キーワードといった HTML のヘッド部分に meta タグで想定される各種情報入力ウインドウが出てくる。当然、何を文書メタ情報として必須にするかは教育段階に応じたものであり、ユーザーが組み込むテンプレートで指定できる。

#### 3.2 文書作成画面とメニューバー

文書内容入力モードでは、物理表現に関連したメニューは現れない。このモードではテキスト入力と論理表現だけを入力するからである。具体的にはコピー、貼り付け、検索・置換(正規表現サポート)といったものと、ワードにおけるスタイル指定機能などが現れる。

##### 1. 段落種類の表示

画面の左側には、ワードの下書きやアウトラインモードの「スタイル名表示領域」に当たるものがある。この部分は、XML の整形式な文書をサポートするため、ワードよりは詳しいものとなっている。

以下、実際の入力手順を示しながら説明する。

###### (1) テキスト入力

ブロック構造	段落種類	段落内容
	本文	はじめに

###### (2) 段落種類設定

今の場合、「はじめに」の段落が章のようなプロ

ックの場合は、段落にカーソルを置き、段落の種類をマークする。すると次のように変化する。

ブロック構造	段落種類	段落内容
<章>	見出し1	はじめに

##### (3) 改段時の継承

ブロック構造要素のマークで始まった場合、ユーザーが「ブロック終了」を選ばない限り、改段操作(enterキー)によってブロック構造に表示される終タグは位置を変えながら残る。

ブロック構造	段落種類	段落内容
<章>	見出し1 本文	はじめに 文書教育に… である。 そこで重要…
</章>	本文	

ただし、新たなブロック構造を入れ子にした場合は、それが入れ子として表示されるのは言うまでも無い。

ブロック構造	段落種類	段落内容
<章>	見出し1	文書とは
<節>	見出し2 本文	定義 文書が何か…
</節>		
</章>		

##### (4) ブロック構造の解除

「ブロック終了」を押すと、ブロック構造に新たなマークが表示される。次の例は「章」が改まった場合のイメージである。

ブロック構造	段落種類	段落内容
</章>		
<章>	見出し1	まとめ
</章>		

なお、以上の1行目にある表題部は本稿説明のためのもので、実際には表示されない。用語も同様である。

##### 2. 段落種類の変更

ウインドウの構造や段落の種類の変更は、該当マークにマウスをあわせることで行えるようになる。

### 3. インライン要素の表示

文書中の特定文字列にインライン的意図を埋め込む場合、ワープロなら太字や下線といった物理表現で識別できるが、文書作成モードは物理表現をサポートしていないのでこれは出来ない。かといってHTMLやXMLのように要素タグを表示すると原稿としての可読性が落ちる。そこで、このモードでは「ユーザーの意図としての物理表現」ではなく、「仕様としての約束」を示す物理表現が表示される(次の例では網掛け)。実際の雰囲気をHTMLのソースと対比させて示してみる。

#### HTML:

重要なことは<em>物理表現</em>ではなく

#### 原稿用紙:

重要なことは**物理表現**ではなく

ユーザーがどのような論理表現を埋め込んだかは該当箇所にマウスカーソルを持っていくとそれが表示され、かつ、変更できる。



重要なことは**物理表現**ではなく

### 4. アウトラインモード

これに関しては既存のアウトラインプロセッサーの機能と同等のものを組み込めばよい。

#### 3.3 物理表現モード

物理表現モードを文書教育に限れば、HTMLファイル作成支援ソフトのスタイルシート編集機能だけを組み込み、実際の見栄えは編集できないブラウザ機能でよい。しかし、このソフトを実務で使えるようにするには、既存のワープロソフトレベルの物理表現力が必要となろう。しかし、先に述べた「データファイル側からソフトの仕様を規定」の立場から言えば、その仕様は別の議論となるため、ここでは次の1つの問題だけを指摘して終わることにする。

#### ●番号自動作成問題

ワードの段落番号や Tex に見られるように、電

子化文書支援ソフトでは、連番サポート機能がある。この場合、「第1節」とするか「1. 2」とするかを物理表現とみるか入力された文書内容の実体と見るかは大きな問題である。HTMLで見てみよう。

```
<ol>
  <li>ヘッダーブロック</li>
  <li>データブロック</li>
</ol>
<p>さて、この■■■■■で示したデータブロックとは</p>
```

この場合■■■■■何となんと書いたらよいのだろう。というのは、li 要素の表示はブラウザに支配されているのだから。私の考えでは、連番の書式は作者側の文書内容の一部と考える。したがって電子原稿用紙は、連番のサポートした結果を DTP 機能や外部 XLL-FO に送る場合は、テキストに変換してしまう方が良いと考えている。これを HTML の面から見れば ol 要素は不要で、接頭記号なしの ul 要素だけを考えることである。

### 4 まとめ

文書教育現場からの経験をもとに新しいタイプの文書作成支援ソフトの仕様について提言した。

文書の電子化は SGML 以来の長い歴史を持った課題であり、私自身の理解も不十分である。このため、具体的な仕様については問題があることは重々承知しているので、いろいろ指摘していただければありがたい。本考察が文書電子化や文書教育現場の参考になれば幸いである。

#### 参考文献

- [1]水島賢太郎、「作・文書」教育のための電子原稿用紙の仕様について、情報処理学会シンポジウム SSS2002 論文集 p167-174、2002.8
- [2]水島賢太郎、人間をモデルとして拡張したオブジェクト指向ファイル概念とその思想-x-file、x-folder、v-router とは -、情報処理学会、情処研報 Vol.2003, No.70, p35-42、2003.6
- [3]大学における文書電子化教育カリキュラムの提案と指導の要点－新学習指導要領実施を踏まえて－、情処研報(情報処理学会)、Vol.2002, No.39の p22