

3 . 情報発信プラットフォームの検討

3 . 1 発信方式の概要

ネットワークへの情報提供を考えた時、最適なフォーマットで発信することが、一人でも多く閲覧することにつながる。表 3.1.1 に代表的なフォーマットの一覧を示す。

ドキュメントのフォーマット		画像圧縮フォーマット	
P D F	米国・アドビシステムズ社のドキュメント表示用のファイル形式	G I F	イラスト向き、256色限定、圧縮率は低い
H T M L	ホームページのドキュメント記述言語	T I F F	米国でよく使用されている画像表示用の形式、フルカラー、モノクロ2形式、圧縮率は中、可逆圧縮
S G M L	構造を持ったドキュメントの記述言語	J P E G	写真向き、フルカラー、圧縮率は任意だが高いほど画質が劣化、不可逆圧縮
X M L	構造を持ったドキュメントの記述言語	(M P E G)	(動画像用)

表 3.1.1 フォーマットの一覧

表 3.1.1 に示す各ドキュメントのフォーマットには以下の特徴がある。

- ・ PDF (Portable Document Format) は、レイアウト情報もほぼ完全に維持され、クオリティーを余り落とすことなく、ファイルサイズもかなり小さくできる。
- ・ HTML (HyperText Markup Language) はインターネット上での標準的フォーマットとしてお馴染みである。レイアウトは大まかにしか指定できない上、複数ページを一つのファイルにまとめられないので、携帯性・完成度の点で難がある。
- ・ SGML (Standard Generalized Markup Language) は、構造化された文書を複数のスタッフで改訂する場合等に向いている。システム規模も大きく値段も高い。機種や OS を問わずに文書の交換ができるが、レイアウト情報は失われる。
- ・ XML (eXtensible Markup Language) は、SGML を軽量化した拡張版 HTML の様なものである。ブラウザの違いによりレイアウトがくずれたり、字体が指定できない等の制約がある。値段は SGML よりかなり安い。

画像圧縮フォーマットについては、ドキュメント中の画像処理に適宜選択して使用される。GIF (Graphics Interchange Format) や TIFF (Tagged-Image Files Format) が良く使われ、JPEG (Joint Photographic Experts Group) は写真に良く使われる。動画像に最適な MPEG (Motion Picture Experts Group) は本目的では対象外である。

表 3.1.2 には、フォーマットを選択する上での大切な要素を一覧にした。

種別	閲覧 ソフト 対価	クロスプラット フォーム 対応	セキュ リティ	アナログ データ からの変換	デジタルデータ からの変換		圧縮
					画像のみ	文書含む	
P D F	無償						
H T M L	無償		×	×			
S G M L	有償			×			×
G I F	無償		×			×	
T I F F	有償		×			×	
J P E G	無償		×			×	

表 3.1.2 フォーマットの機能別の適否⁴⁾

すなわち、

- ・ 閲覧ソフトが無償でクロスプラットフォーム対応: 誰が何処でも同じ内容で見られる環境が容易に得られること。
- ・ セキュリティ: デジタルコンテンツは複製や改ざんが容易であり、それらを判別し難いので防止できること。
- ・ データの変換性: 過去の紙媒体(アナログ)のものと今後のデジタル化したものとを異なるフォーマットで管理するのでは手間や費用が無駄となる上、利用者側にも不便をかける。一貫して同じフォーマットで管理できること。
- ・ 圧縮: 圧縮してデータ量が少なくなればダウンロードの時間が短くなる等ネットワークへの負担も低減し、記憶媒体の容量も少なくて済む。

この表からも配信に最適なドキュメントフォーマットはPDFと考えられる。従って、以下ではPDFを対象に述べることにする。

3.2 PDFの概要

PDFは、あらゆるOS間で電子文書配信が可能となるように米国・アドビシステムズ社が開発した電子出版ソフトウェア“Adobe Acrobat”で使用されているドキュメント表示用のファイル形式である。1997年に登場してその後注目されるようになり、米国ではデファクトスタンダードになりつつある。日本でも多数のインターネットサイトから新聞、パンフレット、製品カタログ、スペックシート、雑誌等がPDFで発信され、従来印刷物として郵送するしかなかった分野で盛んに採用されている。

日本語の場合は、4.6(35ページ)で触れる様にフォントの形や字詰め量が変ることがあるが、レイアウト情報もほぼ完全に維持され、ファイルサイズもかなり小さくなり、配信に向いている。以下にPDFの大切な要素を記す。

1. 閲覧ソフト“Adobe Acrobat Reader”の自由配布・無償入手が可能なこと。普及率が高くメジャーであること。
(<http://www.adobe.co.jp/product/acrobat>よりダウンロードが可能)

- 2 . Windows、Mac、UNIX など様々なマシン用に閲覧ソフトがあり、印刷イメージとほぼ同一の画面表示・印刷出力が得られること。
- 3 . 冊子体が不足したときは、その代替りとなる程度の印刷品質であること。
- 4 . ファイル内容が改ざんされずに配布できることを保証するセキュリティ機能があること。
- 5 . 過去の文書と今後の文書とを同一の PDF フォーマットで一元管理できること。
- 6 . 効率の良いファイルサイズの圧縮が可能であり、データ配信に向いていること。
- 7 . 文書の検索が可能であること。
- 8 . 文字や画像を一つのファイルの中で違和感無く扱うことができること。
- 9 . 研究者の作成したワープロや表計算等のファイルからレイアウトもそのままに直接 PDF ファイルを作成できること。

この中で、6 の圧縮率の高さは伝送時間や保存容量に影響し、使用上かなり気になる要素である。論文一編を HTML と PDF で作成した場合のファイルサイズを比較したところ、PDF は HTML の 1/7 に圧縮できたという報告もある。(画像ファイル数は6点)⁴⁾

5 の一元管理は表 2.1.1 (5 ページ) のように紙媒体の構成に幾通りもの状態が存在する場合に問題となろう。例えば方式 A は良いが方式 B は取扱えないようなフォーマットでは、管理の手間・費用など無駄が多くなる。

以上の様に、複雑な条件の全てを満たすのが PDF であり、当財団の情報発信プラットフォームに採用することとした。