

デジタルドキュメント(7)

高久雅生

2013年5月30日(木)3・4時限

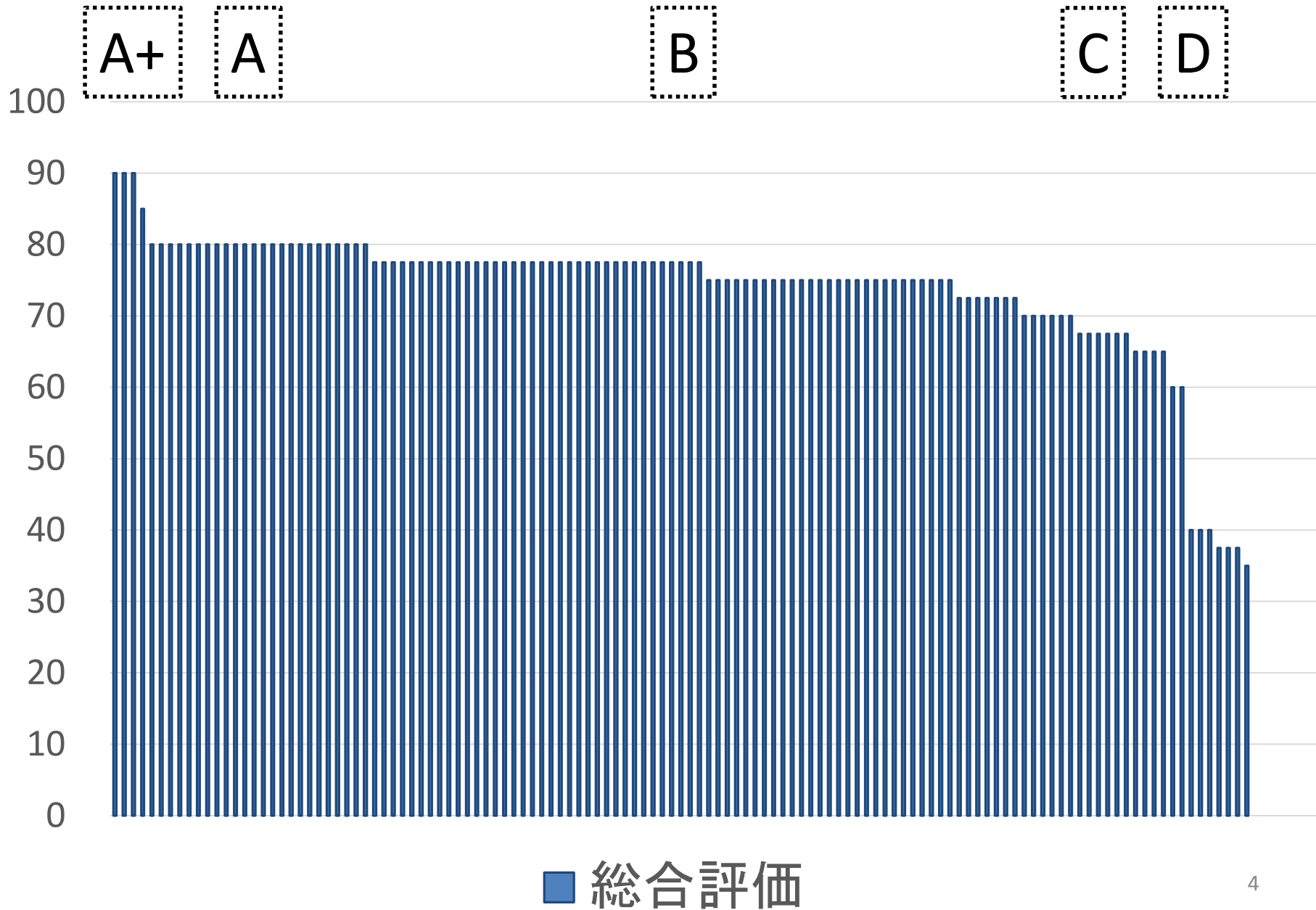
（前回の復習 = ふりかえり）

- 電子書籍の事例を確認しながら、それぞれの特徴を考えてみた
 - 電子書籍端末、書店、コンテンツ、図書館サービス
- 電子書籍の特徴
 - 紙の書籍との比較では...
- 電子書籍の課題
- 最近の動向
- （第3回レポート課題を出題）

前回の出席カードから(質疑: レポート関連1)

- レポートのA+, A-, B-といった +, - の付いた評価は最終的な評価にどう影響するのか？
 - 各レポート課題に関し、以下のとおりの傾斜配点を行っています。最終的には全ての課題分を積算して総合的に考慮したうえで最終評価とします。
 - A+ : 90点以上, A: 80点以上, B: 70点以上, C: 60点以上, D: 60点未満
 - ※未提出課題がある場合には不合格となりますし、遅れ提出の場合には減点等としている他、最終評価は相対評価としますので、評価上の閾値は授業期間中には定まらないものと思われます。ご容赦ください。
- ドキュメントのフォーマットはたとえば、PDFとかJPEGなどを記載するのか？ Kindleストアのものの場合はどうなるのでしょうか？
 - 第3回レポート課題のこととして答えます。基本的に、PDFやJPEG、HTMLというのが「文書形式(フォーマット)」にあたります。
 - Kindleストアにあるモノは大半がKindle(AZW)形式と呼ばれるフォーマットですが、PDF形式等、ほかのフォーマットにも対応しており、それぞれのフォーマットの意味、違い等は適宜調べた上でレポートにて説明してください。

(参考:5/30現在の成績分布)



前回の出席カードから(質疑: レポート関連2)

- 要件は満たしたにも関わらず、分量が少ないとの指摘があり、評価が低かった。
 - 「文章で説明すること」もできるだけ怠らないように。卒研に向けての練習と思って取り組むように。
 - 冗長にならないよう、レポート課題で課されたトピックの説明として、自身が調べ、まとめる内容に関し、端的な事実に加え、前提条件を紛れなく説明することを意識してください。
 - 例: ある論文がオンライン上で提供されているとして...

書誌事項: 高久雅生, 江草由佳. 「東日本大震災発災期のツイッターにおける地域発信情報の分析の試み」. 第5回知識共有コミュニティワークショップ, 福山, 2012, p.81-84. <http://masao.jpn.org/pub/2012/infosocio-ks.pdf>

- 文書フォーマット: PDF
- 提供形態: オンラインで提供されていた。

書誌事項：高久雅生, 江草由佳. 「東日本大震災発災期のツイッターにおける地域発信情報の分析の試み」. 第5回知識共有コミュニティワークショップ, 福山, 2012, p.81-84.

<http://masao.jpn.org/pub/2012/infosocio-ks.pdf>

この論文は著者が自身のウェブサイトで公開していたものであり、文書フォーマットはPDF形式である。この論文は上記書誌事項で示すように、2012年に広島県福山市で開催された研究会において発表された4ページから成る論文である。著者が自身で投稿したファイルをそのまま自身のウェブサイトで提供していると思われる。このため、提供ファイルもスキャン画像によるPDF画像ではなく、文字を拡大縮小してもピクセル単位にギザギザの線(ジャギー)が出ることも無く、なめらかに描画される(右図参照)。

論文のテーマは東日本大震災の直後におけるツイッター上での情報提供の様子を調査したものである。論文では、著者名の英ローマ字表記と英語による論題が付されているほかは、全体を通じて日本語で書かれており、論文の読者は日本語話者が想定されていると思われる。

東日本
After the Great

論文の構成要素として、論題・著者・Abstract・キーワードの書誌的な要素が書かれたのち、本文は「1. はじめに」から「6. まとめ」まで6章から構成され、最後に参考文献が付いている。参考文献では全部で12の文献が番号付きで参照され、そのうち8つの文献にはURLが付けられている。URL付きの文献はURL部分がリンクとなっており、PDFファイルを閲覧しながら、その場でリンクをクリックすることにより、参照文献を読むことができる。参照文献の番号[1]～[12]は、それぞれ本文中の関連する記述の箇所に付されており、どの記述においてどの参考文献を読めばよいか明らかなになっている。

また論文中、表は1つ、図は2つが付けられており、参照文献と同様、表1、図1等として番号を付した上で、本文中で参照している。なお、この論文の中では「図1」が2つあり、これは著者の校正ミスと思われ、後者は「図2」となるのが正しいと思われる。また、図の一つは…

前回の出席カードから(質疑)

- Kindleではどの本がどこまで読まれているか、という事がAmazon側にわかるようになっていて、という話が少し怖いと感じた。
 - 参考まで、以下に利用規約、参考記事を示します:
 - <http://www.amazon.com/gp/help/customer/display.html?nodeId=200506200>
 - <https://www.eff.org/pages/reader-privacy-chart-2012>
 - http://www.dotbook.jp/magazine-k/an_e-book_buyers_guide_to_privacy/ (初版; 日本語訳)
- 国立国会図書館デジタル化資料が、今後全国の図書館へのオンライン配信を予定、とあるが、いつ頃の実施を目指したものか？
 - 平成26年1月からの実施を予定
 - 国立国会図書館.「デジタル化資料の図書館送信に関する改正著作権法の施行について」.
2012年12月17日(プレスリリース)
http://www.ndl.go.jp/jp/news/fy2012/_icsFiles/afieldfile/2012/12/17/pr121217.pdf
 - 国立国会図書館. デジタル化資料の図書館送信に伴う手続.
http://www.ndl.go.jp/jp/aboutus/digi_distribution.html (2013年5月27日アクセス)
- アマゾンや楽天など、今は様々な電子書籍のハードがありますが、先生
の目から見て、何が一番お勧めですか？
 - (正直、よく分かりません...)
 - 軽いのが良い、とか。書籍の種類: 小説が読みたい、マンガが読みたい、洋書が読みたい、、、とか条件に応じて、いろいろあります。
 - 西田宗千佳監修:「これ1冊で完全理解電子書籍: 端末・サービスの選び方から使い方、楽しみ方まで」.日経BP社, 2013, 97p. (ISBN: 978-4-8222-6955-5)

前回の出席カードから(感想)

1. レポートのことにに関してだが、参考文献の書き方を次回こそはきちんと書けるようにしたい。
2. 楽天によるkobo向けの電子書籍販売ページにて、マンガが多くピックアップされていることが意外だった。
3. これまで電子ジャーナルが多く紹介されていたこともあってか、電子書籍といえば小説などの文書だという印象が強かった
4. 個人的には、文書中にリンクを張れるのがデジタルドキュメントの大きな利点だと思っているので、やはりデジタルで読むなら文書がいいと思う。
5. 電子書籍は実はほとんど利用したことが無かったので、読んでみたいと思う
6. 電子書籍の時代がこれから到来すると思いました
7. 楽天のkoboが、無料配布されている青空文庫作品の多くを収録しているのを見て思い出したが、日本においては無料の電子書籍の利用が多いという話を聞いたことがある。背景には、紙の本にはお金を払うが、電子的なコンテンツを(無料で見られるウェブサイトの影響より)「購入する」という感覚があまりないということがあるらしい。ただ、最近は電子書籍を読める媒体もだいぶ広まり、ウェブ上で音楽などのデジタルコンテンツを購入する行為も普及していると思うので、状況はまた変わっているのかもしれない。
8. 電子書籍と紙の書籍の長所・短所が良く分かった。私はどちらが良いかと問われれば「紙の本」と答えるが、最近の高校生が皆電子辞書を持ち歩いているように、検索を目的とするならば、電子書籍に軍配が上がると思う
9. 電子書籍が紙の本に比べて、何が利点で何が欠点なのかが分かった。情報社会の現代ではやはり電子書籍の利点の方が多いと思った
10. iPadやiPhoneが汎用端末だから、電子書籍を読むkoboやkindleと同じ扱いができるというのが面白いと思った。アナログ書籍とデジタル書籍の良いところがよく分かった
11. 電子書籍を閲覧できる端末の紹介がわかりやすかった。
12. もう少しデジタルドキュメントの電子書籍の値段を下げしてほしい。100円ぐらいでないと気軽に買えない
13. 電子書籍ではデータで取り扱われていることで、紙の本のように貸したり借りたりが出来ない。現段階では誰かにすすめたい本があっても、「試しに読んでみて」という部分が簡単にいかない。図書館サービスのあり方も含めて、その部分が今後の課題なんだなと思った。
14. スマートフォンでない携帯電話も意外と売られているんですね。その割には店で見かけないけど・・・
15. いわゆるガラケーも案外売れていてびっくり
16. 電子書籍と電子ペーパーの関連を知りたくなりました

本日のお品書き

- 前回の復習
- ドキュメントフォーマット
 - ドキュメントフォーマットの切り口
 - テキスト / 文字コード
 - ドキュメントフォーマットの例
 - メールフォーマット
 - HTML
- 提出物
 - 出席票

ドキュメントフォーマット (1)

ドキュメントフォーマットとは？

- デジタル文書の内容を機械的に解釈するための「決まりごと」「様式」「形式」
 - データをどのように並べるか（配置）
 - 内容解釈のためのデータ・情報は、デジタル情報として、符号化（エンコード）される
- 例：
 - プレインテキスト？
 - 文書内容を文字コードにより符号化し、テキスト（文字の連なり）として解釈できるようにしたもの。
 - 画像形式？
 - 表現すべき画像要素を、ピクセル・描画要素単位の情報として符号化し、2次元画像として解釈できるようにしたもの。

ドキュメントフォーマットの切り口 (1)

- テキスト (text) vs バイナリー (binary)
 - ビットデータ
 - 文字コードによる解釈
 - 外字
- フォーマットの指定・識別・判別
- シンプルコンテンツ vs 複合メディア
 - 埋め込みコンテンツ
 - ハイパーリンク
- メタデータ
 - 埋め込みメタデータ
 - 外部メタデータ記述
- 文書レイアウト
 - ページ概念
- 文書内の書式要素
 - 見栄え / スタイル
 - フォント
- ファイル vs ストリーム
 - データの保存・蓄積と配信

ドキュメントフォーマットの切り口 (2)

- オープンフォーマット
 - 移植可能性 / ソフトウェア独立性
 - Free / proprietary
- 標準化
 - デファクト標準とデジュール標準 (“de facto” vs “de jure”)
- 文書フォーマットのバージョン
- フォーマット変換
 - テキスト → HTML
 - LaTeX → PDF
- 圧縮
 - 可逆 vs 非可逆
- セキュリティ
 - パスワード
 - 電子署名
- 長期保存
- デジタルフォレンジック

プレインテキストとは？

- 文書内容を文字コードにより符号化し、テキスト（文字の連なり）として解釈できるようにしたもの。
- もっとも基本的なドキュメントフォーマットのひとつ
 - ドキュメントフォーマットの基礎
 - （別の定義：バイナリフォーマットではない）
- 特徴：
 - ほぼあらゆる環境で特別なソフトウェア無しに用いることが出来る
 - テキストフォーマットだけでは、書式要素を保持しない
 - 複合的なオブジェクトとの関連付けはそれ単体ではできない

プレインテキスト読解例 (1)

```
01001000011001010110110001101  
10001101111001000000101011101  
10111101110010011011000110010  
00010000100001010
```

プレインテキスト読解例 (2)

- 先頭からバイト(byte)単位で解読してみよう。
 - 1バイト = 8ビット (bit)

(48)₁₆

(65)₁₆

(6C)₁₆

(6C)₁₆

(72)₁₀

(101)₁₀

(108)₁₀

(108)₁₀

01001000 01100101 01101100 01101100

01101111 00100000 01010111 01101111

01110010 01101100 01100100 00100001

00001010

プレーンテキスト読解例 (3)

- 文字コードと照合
- ASCIIコード表を基にしてみる

48 65 6C 6F 20 57 6F 72 6C 64 21 0A

01001000 01100101 01101100 01101100
01101111 00100000 01010111 01101111
01110010 01101100 01100100 00100001
00001010

※文字コードとは？

- 文字を表現する集合を配列し、それを符号化
する方式を定めたもの
 - 文字集合 (Character Set)
 - 文字符号化方式 (Character Encoding)

文字集合 (Set)

○ = F Z
+ A K

(配列規則)

A B C
+ = ,

(符号化規則)

文字コード

A → 41
B → 42
C → 43
...

ASCIIコード表

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0	NUL	SOH	STX	ETX	EOT	ENQ	ACK	BEL	BS	HT	LF	VT	FF	CR	SO	SI
1	DLE	DC1	DC2	DC3	DC4	NAK	SYN	ETB	CAN	EM	SUB	ESC	FS	GS	RS	US
2		!	“	#	\$	%	&	‘	()	*	+	,	-	.	/
3	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
4	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
5	P	Q	R	S	T	U	V	W	X	Y	Z	[¥]	^	_
6	`	a	b	c	d	e	f	g	h	i	j	k	l	M	n	o
7	P	q	r	s	t	u	v	w	x	y	z	{		}	~	DEL

文字コードの実際

- 欧米語の場合
 - ASCII
 - ISO-8859-1
- 日本語の場合
 - JIS X 0208
 - JIS X 0201
 - ...
- その他
 - Unicode
- よく使われる文字コード
 - ASCII
 - UTF-8
 - Shift_JIS
 - EUC-JP
 - ISO-2022-JP
 - ISO-8859-1

文字コードにおける制御文字

- 計算機上で表現される「文字」には、情報交換の役割があり、たとえば、下記のような特殊な役割をもつ制御文字が含まれている：
 - 空白文字
 - タブ
 - 改行文字
 - 削除記号 (BS, DEL)
- これらの文字以外の表現は、歴史的な文脈で必要とされたという経緯に加え、テキストのみで文書構造を表現するためにも用いられる。

00101101001011010010000000001010101110011110001010110101110101110
01000001011001011101101110000001011100000100000001010000100110101
10000101110011011000010110111100100000010101000110000101101011011
00001011010110111010100101001001000000010000000100000001000001010
00101010100000110010001100000011000100110011110001111010111100110
10010110111111011101010010011001011110000111101111011000111110010
00110000101110011110110011110110001010010011001011101100001101110
01010010011101010101001001101111010100100101101111010010010111111
10100001101000110000101000101111001011110010000000100000110000111
10111101100011111001000110000101110011110110011110110001011111111
01111010111101111100011011010011011011101111101111000011001010111
10011101001011110000110100101110001111010010110100011101001011010
00101011100010100110101101011110011010110111110011110000101000101
11100101111001000000010000001000101011011010110000101101001011011
00001110100010000001101101011000010111001101100001011011110100000
00111001101101100011010010111001100101110011101000111001101110101
01101011011101010110001001100001001011100110000101100011001011100
11010100111000000001010001011110010111100100000001000000101010001
10010101101100001110100010000000110000001100100011100100101101001
11000001101010011100100101101001100010011001100111001001101000000
1010

プレーンテキストに基づくドキュメント フォーマット

- プレーンテキストに基づくドキュメントフォーマットには多くの事例がある
 - Eメール
 - HTMLファイル(ウェブ)
 - SGML / XML
 - LaTeX

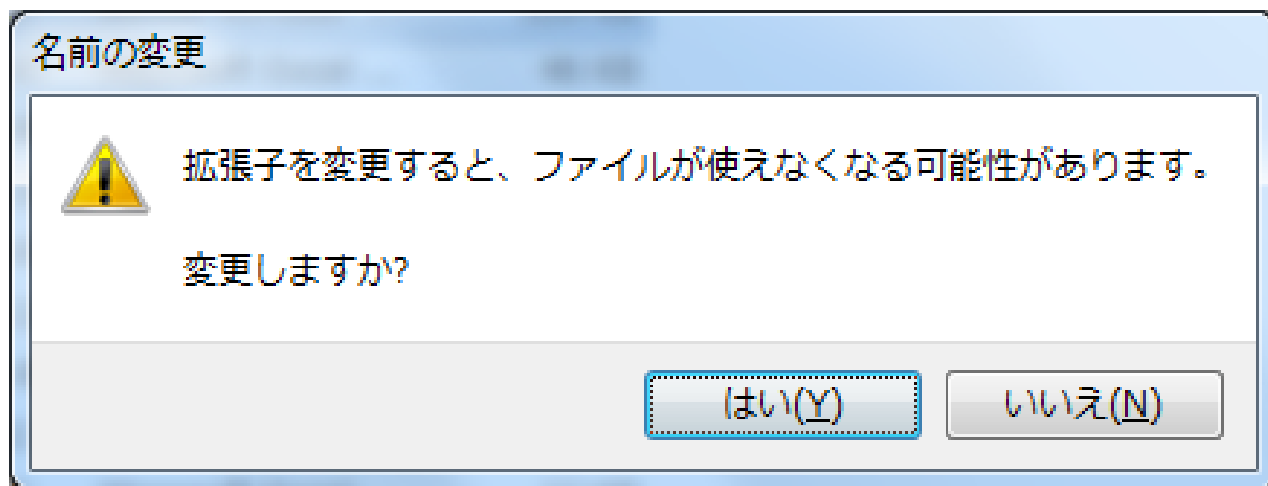
ドキュメントフォーマットの識別, 判別

- メタデータとして付与
 - 拡張子
- 内容で判断
 - マジックナンバー (File magic)
- 通信・転送プロトコル上で指定
 - MIME Type (IANA)
 - その他

フォーマットの識別: ファイル拡張子

ファイル名の末尾を「.」(ピリオド)で区切り、ファイル形式を示す2～3文字からなる文字列を付与

- テキストファイル: text.txt
- HTMLファイル: index.html, welcome.htm
- PDFファイル: example.pdf
- などなど。



フォーマットの識別： ファイルマジックナンバー

- 古典的なヒューリスティック手法
- ファイルの冒頭に分かりやすいフォーマットを示すバイト列ヘッダを挿入しておく方式
- 例：
 - XML形式：<?xml version=" ...
 - GIF形式(画像)：GIF87a または GIF89a

フォーマットの識別:

MIMEタイプ (MIME type)

- Multipurpose Internet Mail Extensions (MIME)
- コンテンツ用のフォーマット指定タイプ
 - Internet Media Type, Content Type
- ウェブやEメールでのドキュメント形式の指定に使われる
- IANAによる公式登録リスト:
<http://www.iana.org/assignments/media-types>
- [type] / [subtype]
 - text/plain, text/html, application/pdf, image/gif, vnd.openxmlformats-officedocument.presentationml.presentation
- タイプ
 - application, audio, image, message, model, multipart, text, video

メールフォーマット

- Eメール: 最も古典的なインターネットアプリケーションの一例
- テキスト形式によるフォーマットの例
 - 基本的に、テキストデータのみでやり取りする
 - (一種、テキスト形式への符号化がなされる)
- オープンフォーマット
 - 様々なソフトウェア(メーラー)がメール内容を扱う
 - 表示、作成、送信
- 送信者と受信者は異なる環境にいることを前提
 - 全世界で幅広く分散して使われており、新しい仕様に改善しようとしても一度に一斉に切り替えることは不可能
- RFC 5322 “Internet Message Format” (2008):
<http://tools.ietf.org/html/rfc5322>
 - RFC 2822 (2001) ← RFC 822 (1982)

メールフォーマットの例

Date: Thu, 20 Sep 2001 07:00:00 JST

From: 首相官邸 <koizumi@mmz.kantei.go.jp>

Reply-To: koizumi@mmz.kantei.go.jp

Subject: 【小泉内閣メールマガジン 2001/09/20】難局にひるまず立ち向かおう！！

To: koizumi-ml-user@mmz.kantei.go.jp

小泉内閣メールマガジン ===== 2001/09/20

☆☆ 今週のキーワード「補正予算」 ☆☆

小泉総理が平成13年度補正予算の編成について準備を進めるよう指示をしました。(解説は最後に)

□ 目次

[らいおんはーと ～ 小泉総理のメッセージ]

● 難局にひるまず立ち向かおう！！

メールフォーマットの例

メールヘッダ

Date: Thu, 20 Sep 2001 07:00:00 JST ↵

From: 首相官邸 <koizumi@mmz.kantei.go.jp> ↵

Reply-To: koizumi@mmz.kantei.go.jp ↵

Subject: 【小泉内閣メールマガジン 2001/09/20】難局にひるまず立ち向かおう！！ ↵

To: koizumi-ml-user@mmz.kantei.go.jp ↵

小泉内閣メールマガジン ===== 2001/09/20

☆☆ 今週のキーワード「補正予算」 ☆☆

小泉総理が平成13年度補正予算の編成について準備を進めるよう指示をしました。(解説は最後に)

メール本文

□ 目次

[らいおんはーと ～ 小泉総理のメッセージ]

● 難局にひるまず立ち向かおう！！

メールフォーマットの例(ヘッダ部)

Received: from goat.ipc.ulis.ac.jp (goat.ipc.ulis.ac.jp [133.51.184.12])

by nile.ulis.ac.jp (8.9.3+Sun/8.9.1) with ESMTP id HAA06473

メールサーバ転送の記録

for <masao@nile.ulis.ac.jp>; Thu, 20 Sep 2001 07:06:10 +0900 (JST)

Received: from mails1.mmz.kantei.go.jp (mails1.mmz.kantei.go.jp [202.212.115.161])

by goat.ipc.ulis.ac.jp (8.8.6 (PHNE_14041)/3.7W) with SMTP id HAA02715

for <masao@ulis.ac.jp>; Thu, 20 Sep 2001 07:08:52 +0900 (JST)

Received: (gmail 6408 invoked by uid 113); 20 Sep 2001 07:03:39 +0900

Received: from localhost (HELO mails1.mmz.kantei.go.jp) (sendmail-bs@127.0.0.1)

by localhost with SMTP; 20 Sep 2001 07:03:39 +0900

Received: (gmail 6325 invoked by uid 200); 20 Sep 2001 07:00:00 +0900

Date: 20 Sep 2001 07:00:00 +0900

From: =?ISO-2022-JP?B?GyRCPHNBJQxRSEbKEI=?= <koizumi@mmz.kantei.go.jp>

Reply-To: koizumi@mmz.kantei.go.jp

Subject: =?iso-2022-jp?B?GyRCIVo+LkB0RmlzVSVhITwlayVeJSwIocVzGyhKIDlwMDEv=?=

=?iso-2022-jp?B?MDkvMjAbJEIhW0ZxNkkkSyRSJGskXiQ6TikkQTh+JCskKiQmISobKEo=?=

=?iso-2022-jp?B?GyRCISobKEo=?=

To: koizumi-ml-user@mmz.kantei.go.jp

Message-Id: <20010919220000.6323.qmail@msgs1.mmz.kantei.go.jp>

Mime-Version: 1.0

Content-Type: text/plain;

charset="iso-2022-jp"

Precedence: bulk

Content-Length: 12860

時刻

送信者

返信先の指定

件名

宛先

メッセージID

MIMEバージョン

フォーマット指定

メールマガジン配信指定

本文長

ここまで、メールヘッダ

メールフォーマットの例（ヘッダ部）(2)

- Date
- From
- To
- Subject
- Message-Id
- Received
- Reply-To
- Content-Type
- MIME-Version
- Precedence
- Content-Length

複合型ドキュメントとしての メールフォーマット

- 複数のドキュメントをテキストフォーマットの中に埋め込む
 - マルチパート: 添付ファイルの実現
- MIME: Multipurpose Internet Mail Extensions
 - メール用の交換フォーマットにおいて、添付ファイルやプレーンテキスト以外の形式でのやりとりを行うための拡張仕様
- メールヘッダの拡張
 - MIME-Version
 - Content-Type
 - Content-Disposition
 - Content-Transfer-Encoding

添付ファイル付きのメールフォーマット例

Received: from SINPRD0310MB391.apcprd03.prod.outlook.com ([169.254.5.109]) by
SINPRD0310HT002.apcprd03.prod.outlook.com ([10.255.17.37]) with mapi id
14.16.0311.000; Thu, 30 May 2013 02:38:55 +0000
Content-Type: multipart/mixed; boundary="_002_9BB0C5FDD26EB142B3B76E53820B0FB9020BC402SINPRD0310MB391_"
Content-Transfer-Encoding: binary
From: Masao Takaku <masao@slis.tsukuba.ac.jp>
To: Masao Takaku <masao@slis.tsukuba.ac.jp>
Subject: =?iso-2022-jp?B?GyRCJUYIOSVIGyhC?=
Date: Thu, 30 May 2013 02:38:54 +0000
Message-ID: <9BB0C5FDD26EB142B3B76E53820B0FB9020BC402@SINPRD0310MB391.apcprd03.prod.outlook.com>

--_002_9BB0C5FDD26EB142B3B76E53820B0FB9020BC402SINPRD0310MB391_
Content-Type: text/plain; charset="iso-2022-jp"

これはテストです。

高久

--_002_9BB0C5FDD26EB142B3B76E53820B0FB9020BC402SINPRD0310MB391_
Content-Type: text/plain; name="20130529-klis-tokuron.txt"
Content-Description: 20130529-klis-tokuron.txt
Content-Disposition: attachment; filename="20130529-klis-tokuron.txt"
Content-Transfer-Encoding: base64

jcCSa4nvgUaR5Yp3iUANCG0KgUWOv5biDQoNCi0glI6ObYzjivqJ25L2gsyQaZhlgS2Cx4KkgsiC
wYLEgtyCt4KpgUgNCi0gk/yOJoLMjduCzINlgVuDfYLGjKSLhoNlgVuDfYLNlc+CpoLEguCR5Y/k
lXaCxYK3gqmBSA0KLSCM9paxiPWODZZdgsaCqYLBgsSDQYOKgsWCt4KpgUgNCg0K

--_002_9BB0C5FDD26EB142B3B76E53820B0FB9020BC402SINPRD0310MB391_--

- Content-Type: multipart/mixed;
boundary="_002_9BB0C5FDD26EB142B3B76E5
3820B0FB9020BC402SINPRD0310MB391_"

複数パートの区切り文字を
全体で指定

- 各パートでの内容種別
 1. Content-Type: text/plain; charset="iso-2022-jp"
 2. Content-Type: text/plain; name="20130529-klis-tokuron.txt"

Content-Disposition: attachment;
filename="20130529-klis-tokuron.txt";

HTML (Hypertext Markup Language)

- ウェブ上でのコンテンツ記述用言語
- ウェブブラウザによる解釈と描画
- タグ;メタデータ;構造
- ハイパーメディア;ハイパーテキスト
– リンク
– ページ埋め込み
- テキストデータのやり取り
- マルチメディア(画像、音声、動画)
- ダイナミックな表現;インタラクション

```
<html>
<head>
<title>タイトル</title>
</head>
<body>
<h1>見出し</h1>
<p>段落</p>
</body>
</html>
```

WebとHTML

- HTTP, URI, HTMLの3点セットによるWebの実現！
- HTTPレイヤでのデータ転送とドキュメントフォーマット指定
- URI指定によるリンクを通じたハイパーメディアの実現

XML (Extensible Markup Language)

- 順序付き木構造による表現
- スキーマ(文書内容の定義)に則った、個別のインスタンス言語
 - HTML
 - その他の文書形式
- 参照関係
- 文字コード: Unicode
- 元情報としてのXML文書
 - 変換することにより、他のデータ形式へ
 - One source multiple use...

まとめ

- ドキュメントフォーマットとは？
- ドキュメントフォーマットの切り口
- フォーマットの識別、判別
- プレインテキストのフォーマット
 - テキスト / 文字コード
 - ドキュメントフォーマットの例
 - Eメールのフォーマット
 - HTML

出席票の提出

提出年月日、学籍番号、所属、氏名、感想コメント等(あれば)を記入のうえ、提出してください。

提出位置:

3編生

2013xxxxx

3編生

2012xxxxx

3年次
(知識)

2011xxxxx

3年次
(創成)

2011xxxxx

4年次

2010xxxxx,
etc.

※前回に欠席し、レポート課題の返却を受けていない者は申し出て、返却レポートを受けとること。