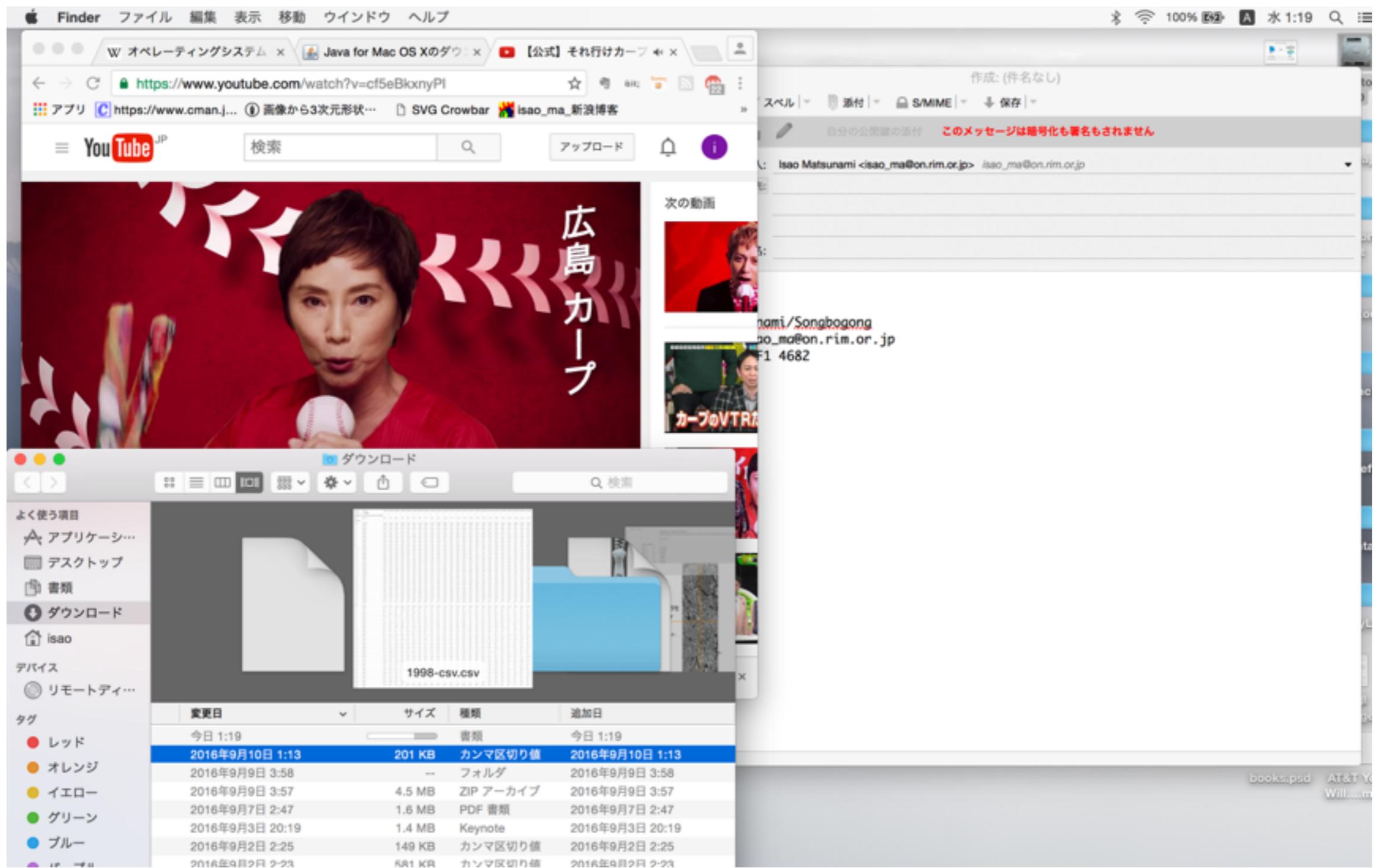


# ダウンロード+YouTube+メール



## OSの目的

### リソースの管理

マルチタスクを可能にする。

あるソフトウェアは、同時に別のソフトウェアが実行されていることを心配する必要がない。

### 基本サービスの提供

文字を表示する、漢字変換する、ファイルに保存する、印刷する機能を、全ソフトが各自用意する必要がない。

### ハードウェアの抽象化

どの会社のHD、ディスプレイでも同じように動く。

ハード会社は「ドライバ」だけを提供する

## OSの目的

### 神様の役割

ハングアップしたソフトは、自分自身を終了できないから、OSが殺してやらなければならない。

コピー＆ペーストは、OSがデータを仲介している。

### デザインの統一

メニューに印刷、設定、ビューなどがあることなど、共通デザインを半ば強制する。Ctrl+Cなども共通

OSが一つの世界を作り上げてしまう  
→データを共有するときに大問題

# データのフォーマット

データはどのように保存されているか

## 1.プロプライエタリ

psdはフォトショップ、xlsはエクセルでしか開けない  
多くの場合、企業秘密  
[ユーザーが「対応しなければならない」](#)

## 2.オープンな業界標準フォーマット

画像のjpg,png、音声のmp3、映像のmp4（ただし乱立）  
地図のshpなど。業界団体によって規約が公開されている  
[ソフト側が「対応しなければならない」](#)

## 3.素朴なフォーマット

シンプルテキスト、CSV、固定長、XML

# バイナリエディタでxlsxファイルを見る

	参考資料シート.xlsx															
0	50 4B 03 04 14 00 08 08 08 00 66 2E 29 PK f.)															
13	49 00 00 00 00 00 00 00 00 00 00 00 00 I															
26	0B 00 00 00 5F 72 65 6C 73 2F 2E 72 65 _rels/.re															
39	6C 73 AD 92 CF 4A 03 31 10 87 EF 7D 8A lsi.?.J 1 ..}.															
52	90 7B 77 B6 15 44 64 B3 BD 88 D0 9B 48 .{wカ Ddウス.ミ.H															
65	7D 80 98 CC FE 61 37 99 30 19 75 7D 7B }\.フ™a7.0 u} {															
78	83 08 5A A9 A5 07 8F 49 7E F3 CD 37 43 . Zウ・ .I~.^7C															
91	9A DD 12 66 F5 8A 9C 47 8A 46 6F AA 5A .ソ f...G.FoIZ															
104	2B 8C 8E FC 18 7B A3 9F 0E F7 EB 1B BD +... {」. .ス															
117	6B 57 CD 23 CE 56 4A 24 0F 63 CA AA D4 kW^#赤VJ\$ cハエヤ															
130	C4 6C F4 20 92 6E 01 B2 1B 30 D8 5C 51 トL. .n イ 0リ¥Q															
143	C2 58 5E 3A E2 60 A5 1C B9 87 64 DD 64 ツX^:. `・ ケ.dソd															
156	7B 84 6D 5D 5F 03 FF 64 E8 F6 88 A9 F6 {.m]_ .d...ウ.															
169	DE 68 DE FB 8D 56 87 F7 84 97 B0 A9 EB ^ h` . .V....-ウ.															
182	46 87 77 E4 5E 02 46 39 D1 E2 57 A2 90 F.w.^ F9A.W『.															
195	2D E7 20 46 2E 22 BC 11 AF CE 44 F3 FF 6E /23 07DSU															

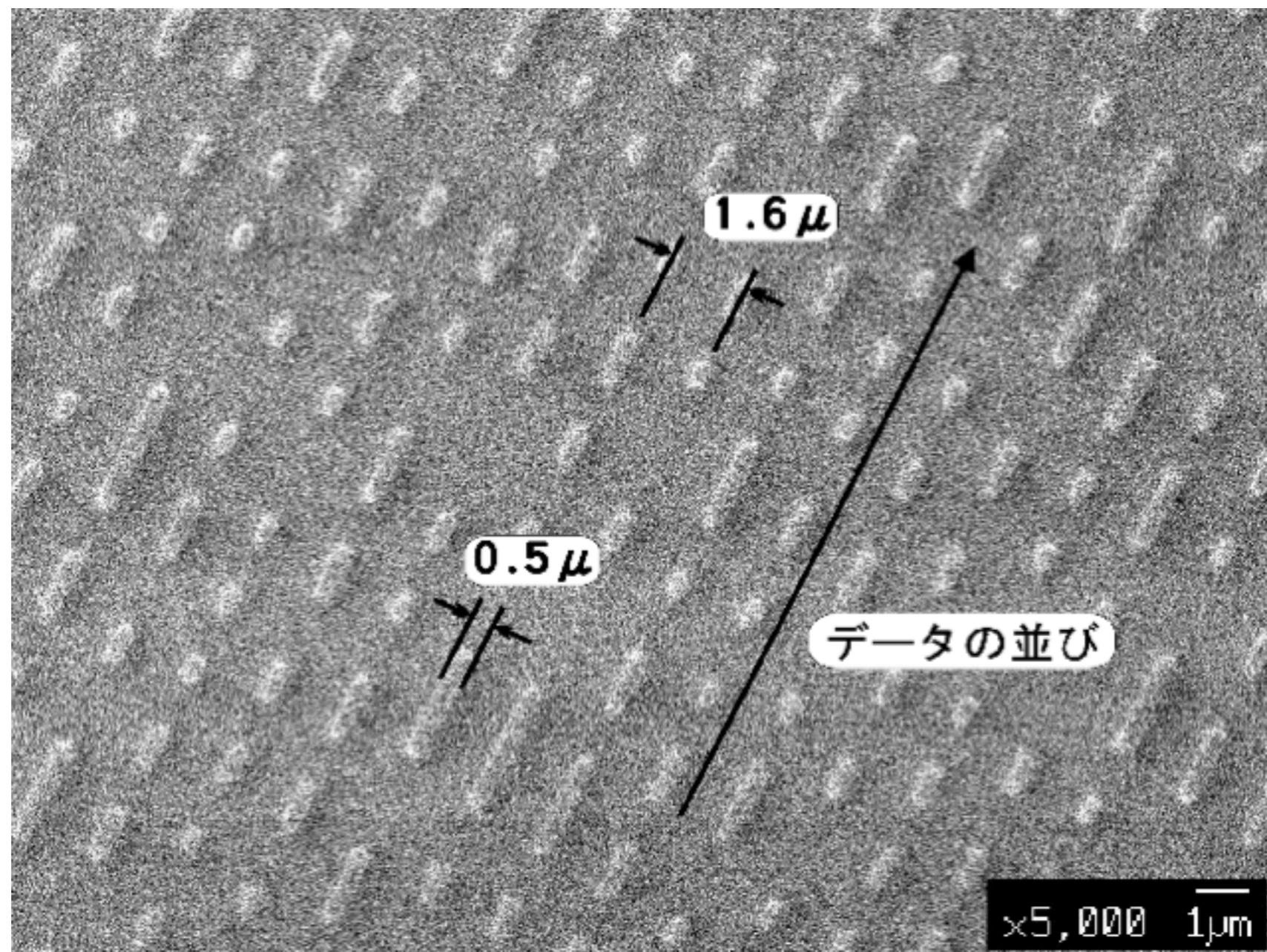
Signed Int big (select some data)

0 out of 30052 bytes

# バイナリエディタでjpgファイルを見る

P1100206 (2).jpg											
0	FF	D8	FF	E0	00	10	4A	46	49	46	00
11	01	01	01	00	B4	00	B4	00	00	FF	E1
22	61	42	45	78	69	66	00	00	4D	4D	00
33	2A	00	00	00	08	00	0E	01	0F	00	02
44	00	00	00	0A	00	00	08	C2	01	10	00
55	02	00	00	00	09	00	00	08	CC	01	12
66	00	03	00	00	00	01	00	01	00	00	01
77	1A	00	05	00	00	00	01	00	00	08	D6
88	01	1B	00	05	00	00	00	01	00	00	08
99	DE	01	28	00	03	00	00	00	01	00	02
110	00	00	01	32	00	02	00	00	00	14	00
121	00	08	E6	01	3B	00	02	00	00	00	01
132	00	00	00	00	02	13	00	03	00	00	00
143	01	00	02	00	00	87	69	00	04	00	00
154	00	01	00	00	00	5A	61	A5	00	07	00

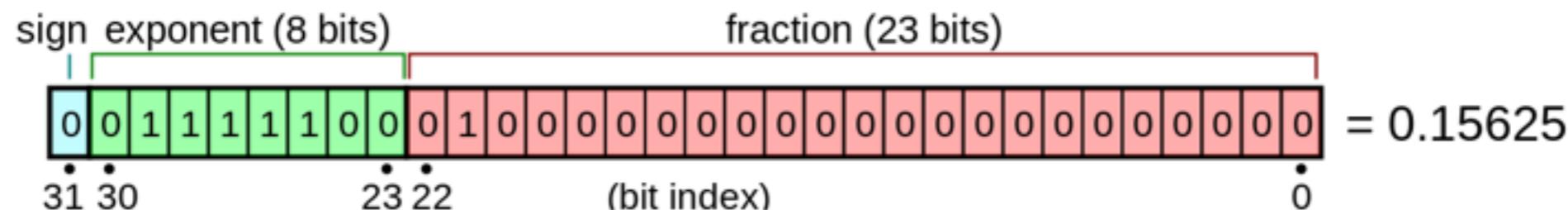
# 「デジタルなんて所詮0と1」



# IEEE 754

## 32ビット単精度 [ソースを編集]

单精度二進化浮動小数点数は、32ビットワードに格納される。



sign は符号、exponent は指数部、fraction は仮数部である。

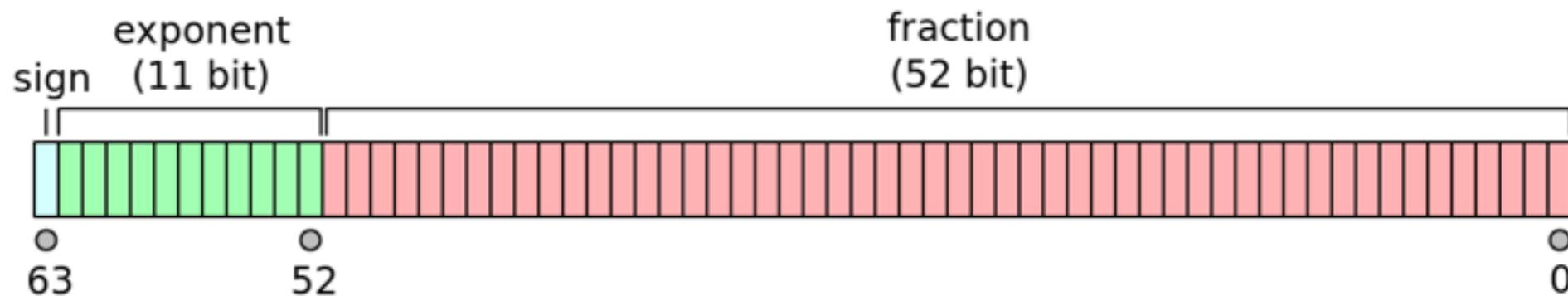
指数部は下駄履き表現（バイアスまたはエクセスとも。[符号付数値表現を参照](#)）と呼ばれる形式であり、実際の値に、ある固定値（ここでは  $emax = 127$ ）を加算したものである。このような表現にしているのは浮動小数点数同士の比較を単純にするためである。指数部は大きな値も小さな値も表せるように負の値にもなるが、これを単に2の補数で表すと、全体の符号 sign とは別に exponent も符号を持つことになり、単純な大小比較ができなくなってしまう。そのため、指数部はバイアスされて常に正の値となるような形式で格納される。単精度では $-126 \sim +127$ に127を加えて、 $1 \sim 254$ としている（0と255は特殊な意味を持つ。後述）。この表現により「指数が正の数」「指数が1の数」「指数が負の数」「0」を、この順に自然に並べることができる。浮動小数点数を解釈するときは、バイアスを減算して実際の指数を求める。

表現可能なデータは指数部の値によって区別され、仮数部の値にも影響される。指数部も仮数部も符号無しの二進整数であることに注意されたい（指数部は  $0 \sim 255$ ）。

# 現在は64bitが標準

## 64ビット倍精度 [編集]

倍精度も、各フィールドの幅が広くなっているだけで、考え方は同じである。

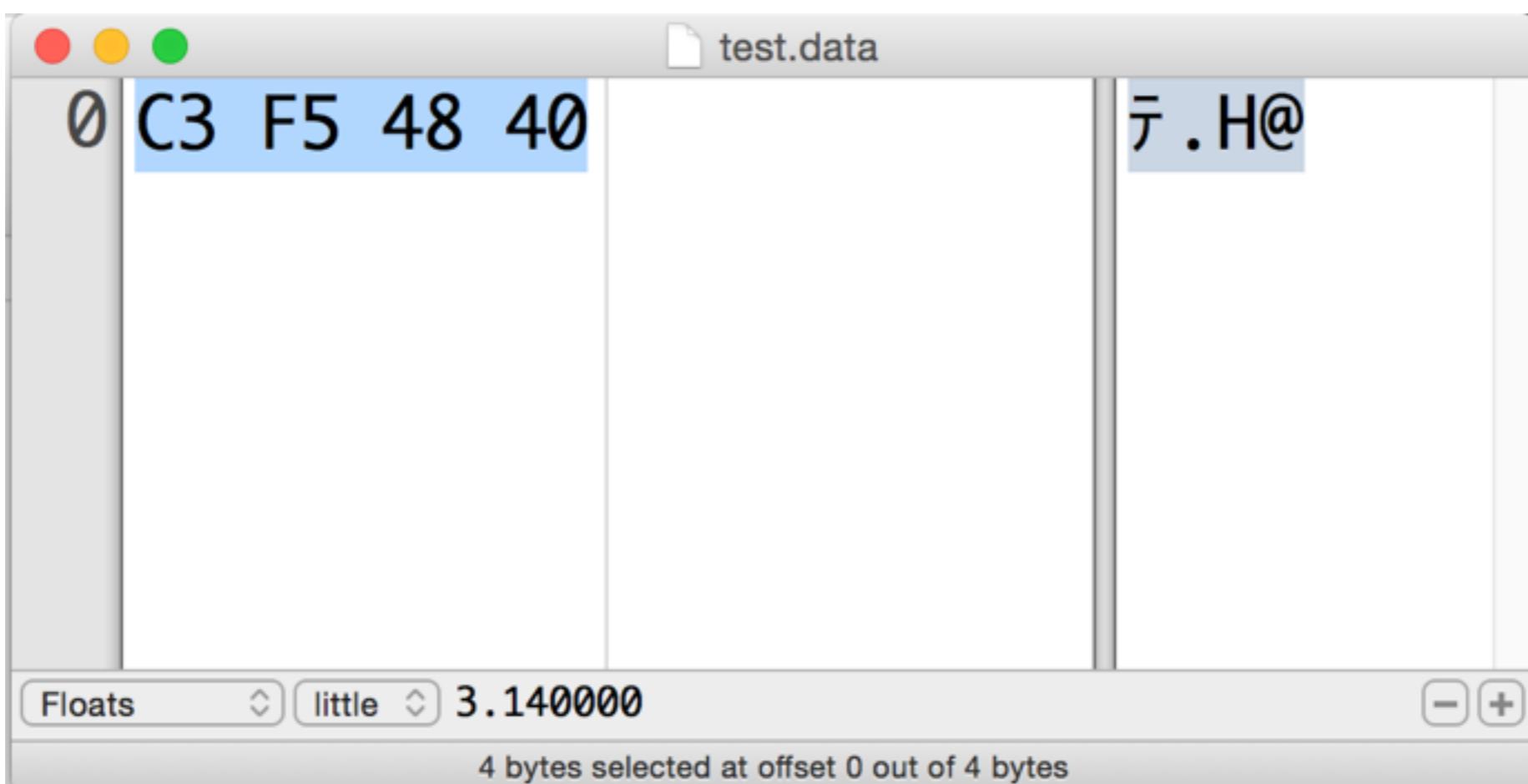


正規化数では、指数は  $emax = +1023$  でバイアスされる（したがって  $e$  は  $\text{exponent} - 1023$ ）。正規化数の指数は  $e: +1023 \sim -1022$  ( $\text{exponent}: 2046 \sim 1$ ) である ( $\text{exponent} = 2047$  の時無限大またはNaN、 $\text{exponent} = 0$  の時非正規化数で  $e = -1022$ )。正規化数の時仮数部はけち表現である。

# バイナリ・ファイル

数字を直接書き込む形式

```
>>> import struct  
>>> fout = open("test.data", "wb")  
>>> a = 3.14  
>>> fout.write(struct.pack("f", a))  
>>> fout.close()
```



文字を表すにはどうすればよいか？



2の7乗 = 128まで表す

1か0でプラスとマイナスを表す

数字としては-128から128までしか表せない

# アスキー文字コード

American Standard Code for Information Interchange, 1963

		LSB	0000	0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111
		MSB	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0000	0	NUL	SOH	STX	ETX	EOT	ENQ	ACK	BEL	BS	HT	LF	VT	FF	CR	SO	SI	
0001	1	DLE	DC1	DC2	DC3	DC4	NAK	SYN	ETB	CAN	EM	SUB	ESC	FS	GS	RS	US	
0010	2	SP	!	*	#	5	%	&	'	(	)	*	+	,	-	.	/	
0011	3	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?	
0100	4	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	
0101	5	P	Q	R	S	T	U	V	W	X	Y	Z	I	\	I	^	-	
0110	6	\	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	
0111	7	p	q	r	s	t	u	v	w	x	y	z	l	l	l	-	DEL	

BINARY — HEX — ASCII

# Latin1文字コード

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0																
1																
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	0
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	{	}	l	~	
8																
9																
A	í	¢	ƒ	¤	¥	ƒ	§	“	”	ø	¤	«	¬	¬	®	¬
B	°	±	²	³	‘	µ	¶	·	,	¹	º	»	¼	½	¾	¿
C	À	Á	Ã	Ä	Å	Æ	Ç	È	É	Ê	Ë	Ì	Í	Î	Ï	Ї
D	Ð	Ñ	Ò	Ó	Ô	Õ	Ö	×	Ø	Ù	Ú	Û	Ü	Ý	Þ	Þ
E	à	á	ã	ä	å	æ	ç	è	é	ê	ë	ì	í	î	ï	ї
F	ð	ñ	ò	ó	ô	õ	ö	÷	ø	ù	ú	û	ü	ý	þ	ÿ

# JIS X 0201(半角カナ)

	00	10	20	30	40	50	60	70	80	90	A0	B0	C0	D0	E0	F0
00	NL	DE	SP	0	@	P	'	p			ー	タ	ミ			
01	SH	D1	!	1	A	Q	a	q		.	ア	チ	ム			
02	SX	D2	"	2	B	R	b	r		「	イ	ツ	メ			
03	EX	D3	#	3	C	S	c	s		」	ウ	テ	モ			
04	ET	D4	\$	4	D	T	d	t		、	エ	ト	ヤ			
05	EQ	NK	%	5	E	U	e	u		・	オ	ナ	ユ			
06	AK	SN	&	6	F	V	f	v		ヲ	カ	ニ	ヨ			
07	BL	EB	,	7	G	W	g	w		ア	キ	ヌ	ラ			
08	BS	CN	(	8	H	X	h	x		イ	ク	ネ	リ			
09	HT	EM	)	9	I	Y	i	y		ウ	ケ	ノ	ル			
0A	LF	SB	*	:	J	Z	j	z		エ	コ	ハ	レ			
0B	HM	EC	+	;	K	[	k	{		オ	サ	ヒ	ロ			
0C	CL	→	,	<	L	¥	l			ヤ	シ	フ	ワ			
0D	CR	←	-	=	M	]	m	}		ュ	ス	ヘ	ン			
0E	SO	↑	.	>	N	^	n	—		ヨ	セ	ホ	。			
0F	SI	↓	/	?	O	—	o	DL		ツ	ソ	マ	。			

# 漢字の登場

256種類では無理→2byteなら256\*256まで可能

## ■ 第1水準漢字(16区～47区)

区	点	JIS	SJIS	EUC	+0	+1	+2	+3	+4	+5	+6	+7	+8	+9	+A	+B	+C	+D	+E	+F
16	0	3020	889E	B0A0		亜	啞	娃	阿	哀	愛	挨	始	逢	葵	茜	穢	惡	握	渥
	16	3030	88AE	B0B0	旭	葦	芦	鯵	梓	庄	幹	扱	宛	姐	虻	飴	絢	綾	鮎	或
	32	3040	88BE	B0C0	粟	拾	安	庵	按	暗	案	闇	鞍	杏	以	伊	位	依	偉	囬
	48	3050	88CE	B0D0	夷	委	威	尉	惟	意	慰	易	椅	為	畏	異	移	維	緯	胃
	64	3060	88DE	B0E0	萎	衣	謂	違	遺	医	井	亥	域	育	郁	磯	一	壻	溢	逸
	80	3070	88EE	B0F0	稻	茨	芋	鰯	允	印	咽	員	因	姻	引	飲	淫	胤	蔭	

区	点	JIS	SJIS	EUC	+0	+1	+2	+3	+4	+5	+6	+7	+8	+9	+A	+B	+C	+D	+E	+F
17	0	3120	893F	B1A0		院	陰	隱	韻	吺	右	宇	烏	羽	迂	雨	卯	鶴	窺	丑
	16	3130	894F	B1B0	碓	臼	渦	噓	唄	鬻	蔚	鰐	姥	廸	浦	瓜	閨	噂	云	運
	32	3140	895F	B1C0	雲	茌	餌	叡	營	嬰	影	映	曳	栄	永	泳	洩	瑛	盈	穎
	48	3150	896F	B1D0	穎	英	衛	詠	銳	液	疫	益	駅	悅	謁	越	閱	榎	厭	円

## 文字コードの種類

- |        |   |
|--------|---|
| JIS    | JIS X 0208。1978年。第一水準2965字、<br>第二水準3384字を決める。<br>メールではISO-2022-JPと呼ばれる                                      |
| S_JIS  | JIS X 0208の亜種。アスキーと半角カナ<br>を共存させる工夫。Microsoftはcp932、<br>MacはMacJapanese。MS-kanji,x-sjisと<br>も呼ぶ場合も。各社微妙に違う！ |
| EUC-JP | UNIXの世界で、S_JISと同じ目的で作られ<br>たコード<br><br><span style="color: blue;">文字化けは大抵このレベル</span>                       |

# 機種依存文字 = 外字

各社が勝手に決めた=別のPCでは表示されない

# フォントがない

	00	01	02	03	04	05	06	07	08	09	0A	0B	0C	0D	0E	0F	10	11	12
0800	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□
0820	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□
0840	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□
0860	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□
0880	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□
08A0	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□
08C0	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□
08E0	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□	□
0900	ঁ	ঁ	ঁ	ঁ	ঁ	ঁ	ঁ	ঁ	ঁ	ঁ	ঁ	ঁ	ঁ	ঁ	ঁ	ঁ	ঁ	ঁ	ঁ
0920	ঠ	ড	ঢ	ণ	ত	থ	দ	ঘ	ন	জ	প	ফ	ব	ভ	ম	য	ৰ	ৱ	ল
0940	ঠী	ডু	ঢু	ণু	তু	থু	দু	ঘু	নু	জু	পু	ফু	বু	ভু	মু	যু	ৰু	লু	-
0960	ঠু	ডু	ঢু	ণু	তু	থু	দু	ঘু	নু	জু	পু	ফু	বু	ভু	মু	যু	ৰু	ঁ	ঁ
0980	□	ঁ	ু	ু	ু	ু	ু	ু	ু	ু	অ	আ	ই	ঈ	উ	উ	ঞ্চ	ু	ু
09A0	ঠ	ড	ঢ	ণ	ত	থ	দ	ঘ	ন	জ	প	ফ	ব	ভ	ম	য	ৰ	ল	ু

## Unicodeの登場

日本語はS\_JIS、簡体字はGB2312(EUC\_CN)、繁体字はBIG5/EUC\_TW、韓国語はJohab/UHCで共存できない。フォントも違う。

→ 対訳テキストが作れない！

全ての言語の文字を表す文字コード=Unicode

4byteで文字を表現する=4byte使うとは限らない

UTF-8: アスキーは1byte、よく使う字は2byte、稀に使う時は3,4byte

UTF-16: 2byteか4byte

UTF-32: 常に4byte

公的文書も最近はUTF8になりつつあります

## 改行コード問題

LF(10): Line Feed=行送り

CR(13): Carriage Return=印刷位置の復帰

Windows

LF+CR

旧Mac

CR

Unix, MacOS

LF

Macで作ったtext/csvファイルを読み込んだ  
Windowsは、いつまでたってもCRが来ない！

ただし、賢いソフトは補ってくれる

# テキスト・ファイル

abc<改行>あ<改行>

encode_utf8.txt	0   61 62 63 0A   E3 81 82 0A   abc ...
encode_utf16le.txt	0   61 00 62 00   63 00 0A 00   42 30 0A 00   a b c B0
encode_shiftjis.txt	0   61 62 63 0D 0A   82 A0 0D 0A   abc .

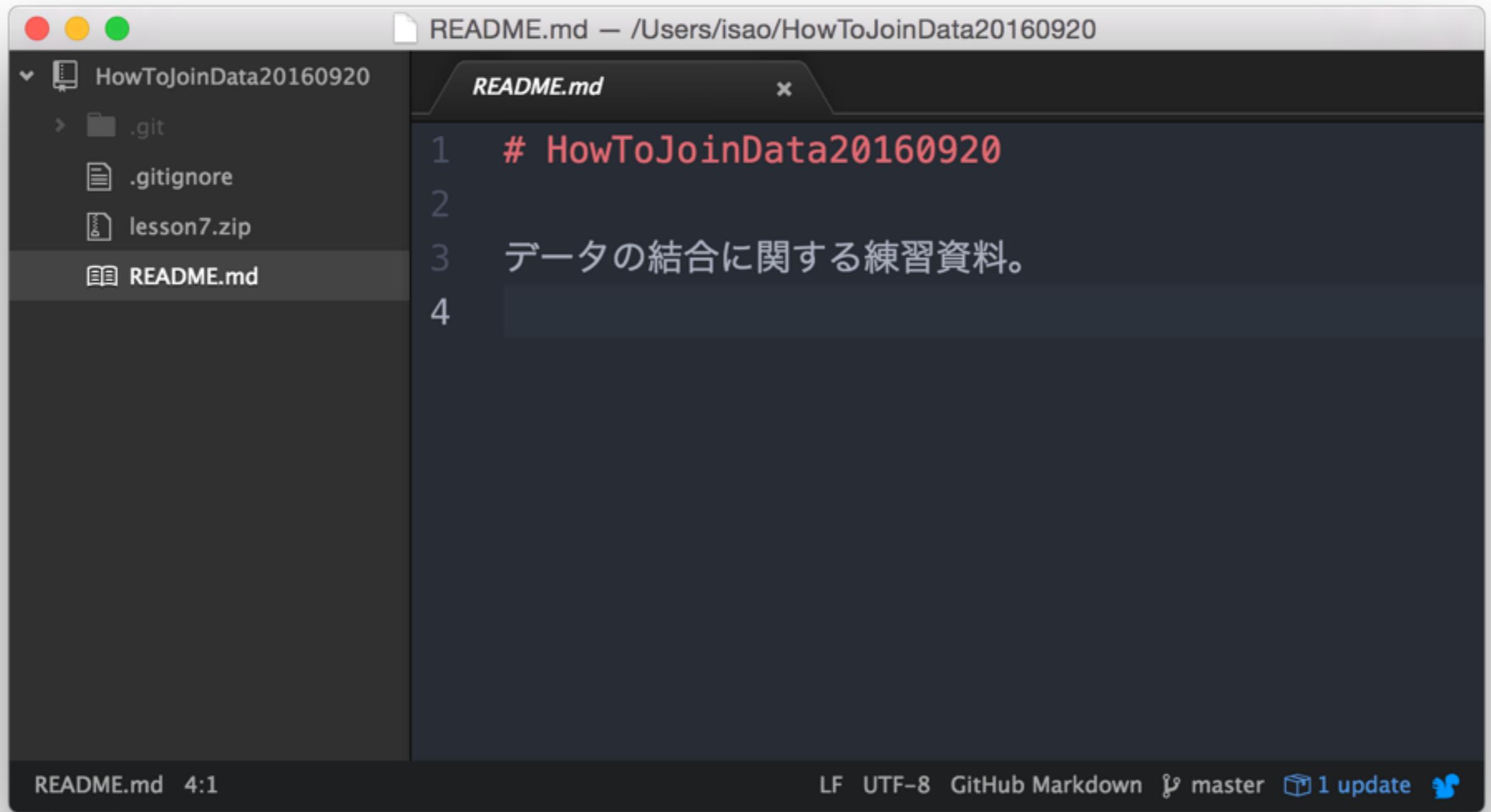
abc: utf8とshiftJISは同じ、utf16は00がついている

改行: utf8は0A、shiftJISは0D0A（Macは0D）、utf16は0A00

あ: utf8はE38182、shiftJISは82A0、utf16は3042(LEは順序反転)

# 実践：文字化け対策

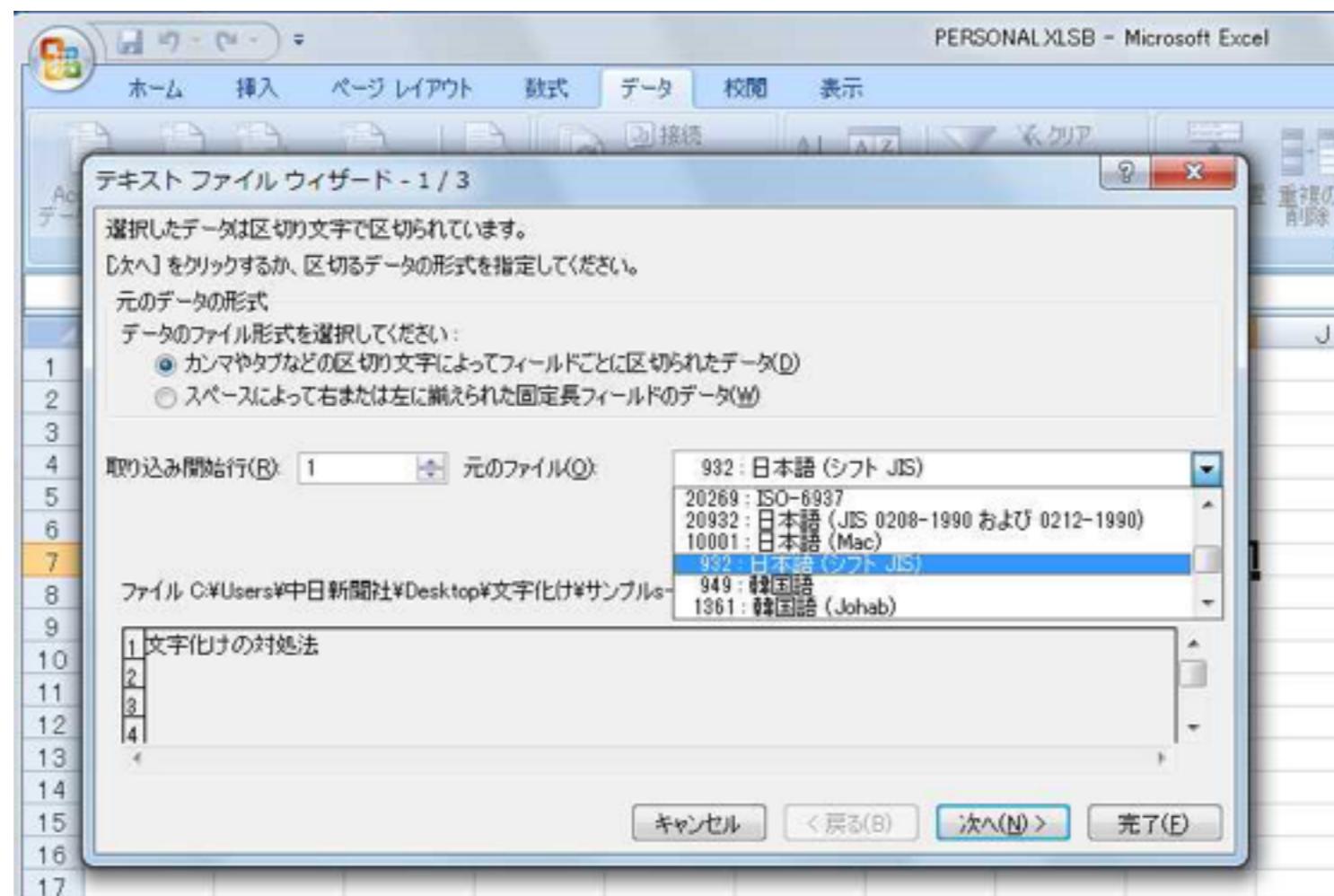
## 1.ほとんどのソフトは文字コード選択可能



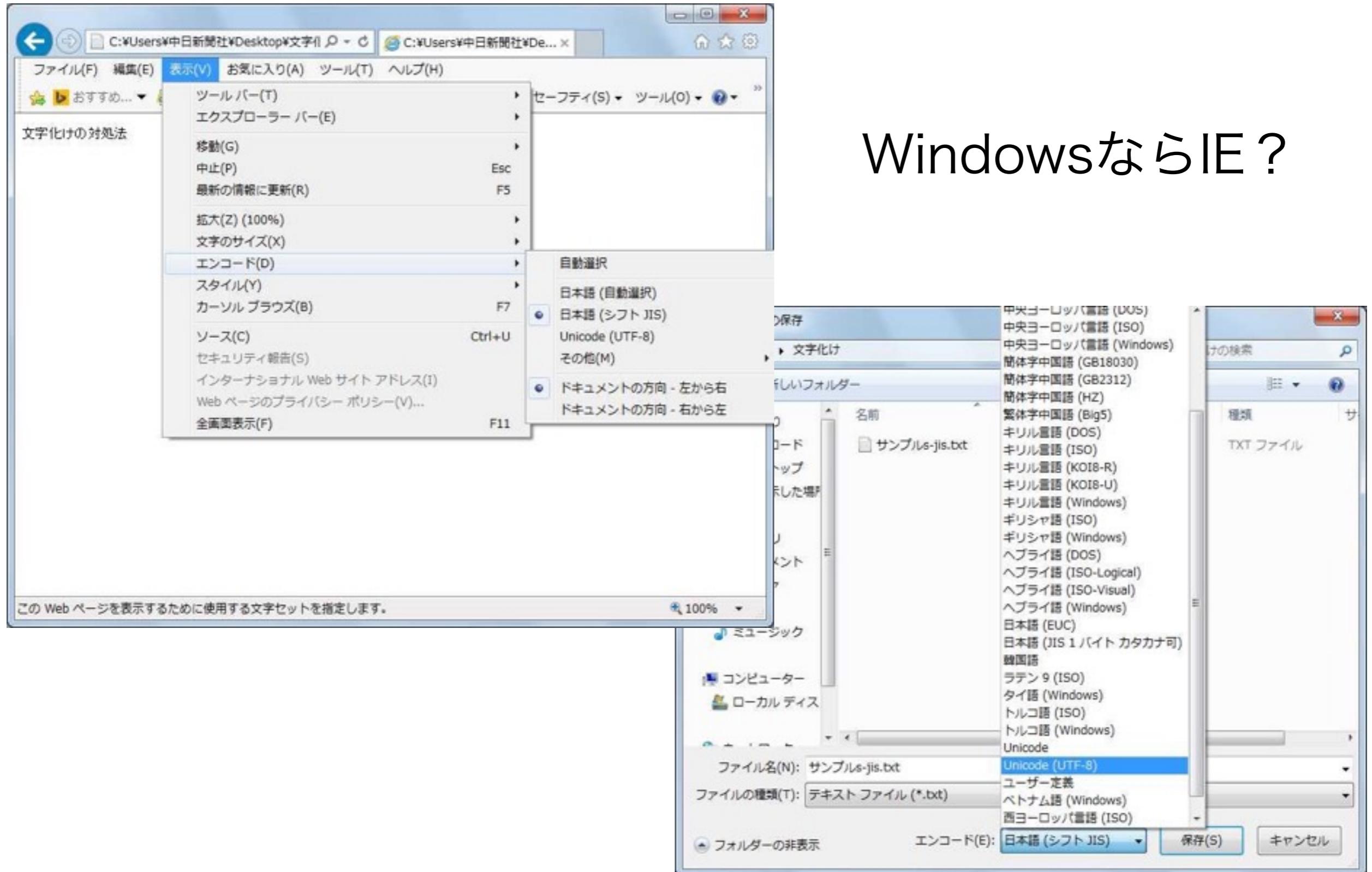
The screenshot shows a macOS desktop environment. On the left, a Finder window displays a folder named "HowToJoinData20160920" containing files: ".git", ".gitignore", "lesson7.zip", and "README.md". The "README.md" file is selected. On the right, a text editor window titled "README.md" shows the following content:

```
1 # HowToJoinData20160920
2
3 データの結合に関する練習資料。
4
```

The text "データの結合に関する練習資料。" is displayed in Japanese. At the bottom of the screen, a dark bar shows the file name "README.md" and the line number "4:1". To the right of the file name, there are status indicators: "LF", "UTF-8", "GitHub Markdown", "master", "1 update", and a small squirrel icon.



## 2. ファイルを完全に変換したい



## 2. ファイルを完全に変換したい

Mac/Ubuntuならコマンドライン

nkfは富士通製変換コマンド（超多機能！）

```
> nkf -guess sample.txt  
UTF-8  
> nkf -s sample.txt > s-jis.txt // 文字コード  
> nkf -Lw sample.txt > CRLF.txt // 改行形式
```

iconvは世界共通（改行形式はそのまま）

```
> iconv -f SHIFT-JIS -t UTF8 sample.txt > utf8.txt
```

腹をくくってコマンドラインを習得しよう

```
> nkf -w -Lu --overwrite *.txt
```

あるフォルダのtxtファイルを一括変換

文字コードが違う文書で  
コピペはなぜ可能？

# CSVファイル

```
◀ ▶ earthquake.csv *
```

1 "time","day","lat","lon","mag","basho"  
2 "16/04/20 13:47",20,33,131.11,1.4,"阿蘇山周辺"  
3 "16/04/20 13:45",20,32.79,130.95,1.6,"阿蘇山周辺"  
4 "16/04/20 13:42",20,33.04,131.14,1.4,"熊本県北東部・大分県境付近"  
5 "16/04/20 13:31",20,32.66,130.67,2.6,"熊本県中部"  
6 "16/04/20 13:30",20,32.62,130.68,0.9,"熊本県中部"  
7 "16/04/20 13:27",20,32.55,130.55,1,"八代海・熊本県南西部"  
8 "16/04/20 13:26",20,32.16,130.51,1.8,"八代海・熊本県南西部"  
9 "16/04/20 13:26",20,32.19,130.48,1.6,"八代海・熊本県南西部"  
10 "16/04/20 13:22",20,32.92,130.99,2.2,"阿蘇山周辺"  
11 "16/04/20 13:14",20,32.57,131.55,2.6,"宮崎県北部"  
12 "16/04/20 13:07",20,32.59,130.65,1.1,"熊本県中部"  
13 "16/04/20 13:04",20,33.28,131.41,1.9,"別府付近"  
14 "16/04/20 13:03",20,32.87,130.84,1.7,"熊本県北部"  
15 "16/04/20 12:47",20,32.96,131.1,2.4,"阿蘇山周辺"  
16 "16/04/20 12:38",20,32.89,130.91,1.4,"阿蘇山周辺"  
17 "16/04/20 12:38",20,32.83,130.85,1.5,"熊本県北部"  
18 "16/04/20 12:34",20,32.97,131.12,1.3,"阿蘇山周辺"  
19 "16/04/20 12:33",20,32.71,130.68,3.2,"熊本県中部"  
20 "16/04/20 12:31",20,32.92,131.03,1.7,"阿蘇山周辺"  
21 "16/04/20 12:26",20,32.51,130.62,0.9,"熊本県中部"  
22 "16/04/20 12:21",20,33.05,131.14,1.2,"熊本県北東部・大分県境付近"

# Escape問題

単純CSV: 「コンマ付き文字列」を誤解する

引用符付きCSV:

「引用符付き文字列」を誤解する

エスケープ文字( 文字列内の引用符は「\」)

TSV: コンマではなくタブで区切る

文字列内のタブは?

# 固定長データ

measure\_20150624\_7.txt \*

1	J2015062400013548	008 362699	017 1374127	017 21012102V	511 4143HIDA MOUNTAINS	REGION	20K
2	_N.OMCH2740	%24P	00013665S	013743 1762 00	1 1139 01	1 1181 00	01 15 6AM 1
3	_E.MTU	807 t24P	00013795S	013970 607 02	3 627 01	3 655 01	2J 15 6MM 1
4	_N.HOTH2701	h24P	00013813S	013993 1116 02	2 1033 01	2 832 00	0J 15 6AM 1
5	_N.OM2H2700	h24P	00013824S	014018 769 00	2 1279 00	2 484 00	2J 15 6AM 1
6	_N.AHMH2796	N24P	00013903	2830 01	2 3014 01	3 581 01	3J 15 6M 1
7	_N.AAKH2809	s24P	00014011	221 00	4 245 00	4 221 02	0J 15 6M 1
8	_DP.TYJ1367	%24P	00013993				15 6M 1
9	_N.IKSH2699	h24P	00014035	110 01	4 95 01	4 88 01	4J 15 6M 1
10	_N.SNBH2795	N24P	00014053	594 02	4 745 02	4 356 03	4J 15 6M 1
11	_N.HKKH2812	h24EP	00014043S	014388 453 00	4 301 01	4 235 02	4J 15 6MM 1
12	_N.MSTH2515	h24P	00014045S	014380 119 01	4 118 01	4 100 01	4J 15 6AM 1
13	_N.KTRH2667	h24P	00014055S	014378 106 01	4 88 01	4 86 01	4J 15 6MM 1
14	_N.KOKH2577	h24P	00014096S	014453 726 00	4 486 01	4 491 01	4J 15 6MM 1
15	_N.HBAH2709	h24S	00014442	125 01	1 130 00	1 151 00	0J 15 6A 1
16	_TTATEY	572 N24P	00014242	1043 01	6	473 01	6J 15 6A 1
17	_NSAKAI	569 N24P	00014231S	014699 89 02	5 143 01	5 142 01	0J 15 6AM 1
18	_DP.KTJ1318	t24S	00014670	183 01	0 215 00	0 140 01	0J 15 6M 1
19	_E.ASI	408 t24S	00014734	120 01	0 101 01	0	J 15 6M 1
20	_NIUKAW	578 N24P	00014303S	014805 216 00	5 299 01	5 172 00	5J 15 6MM 1
21	_MATSUS	67 G24P	00014386S	014961 100 01	6 56 01	7 84 02	7J 15 6AM 1
22	E						
23	J2015062400034143	013 352805	041 1411067	078 349210620V	511 3110NEAR CHOSHI CITY		28K
24	_TENNOD	699 N24P	00034980ES	035554 4444 00	1 3579 00	1 5668 01	1J 15 6MM 1
25	_N.CHSH2725	s24P	00034982S	035579 1818 03	7 1809 04	9 2669 02	1J 15 6AM 1
26	_N.CH2H2631	h24P	00035016ES	035689 2782 04	9 2473 04	8 1692 03	9J 15 6AM 1
27	_N.HA2H2764	h24P	00035201	2012 05	15 1466 01	13 1681 02	13J 15 6A 1
28	_N.HSNH2799	h24P	00035246ES	040045 1321 05	16 1423 02	14 1353 02	8J 15 6MM 1
29	_N.SSAH2824	h24EP	00035331	2522 05	13 3610 04	15 1631 03	2K 15 6M 1

メモリーが貴重な時代：最も無駄のないデータ

# データの位置、コードは事前に定義

レコードレイアウト																		
レコード名		震源レコード												レコード長				
項目型	ヘッダーダー	西暦					日時			分			秒			標準誤差 (秒)		
		A1	I4	I2	I2	I2	I2	I2	I2	I2	I2	I2	F4.2	F4.2				
項目型	緯度 (度)	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	標準誤差 (分)	
		I3		F4.2		F4.2							I4		F4.2		F4.2	
項目型	深さ (km)	43	44	45	46	47	48	49	50	51	52	53	54	M1種別	M2種別	使用走時表	震源評価	
																A1	A1	
項目型	マグニチュード 1種別	F5.2 or I3.2														震源情報	最大震度	被害規模
		43	44	45	46	47	48	49	50	51	52	F2.1	A1	F2.1	A1	A1	A1	A1

# XMLデータ

HTML (XMLの一種) のように、データをタグで記述する方式

要素数が事前に決められない場合

データの定義が固定できない場合

データの検証が必要な場合

データ交換の業界標準方式

ただし、データサイズは巨大

```
<bookstore>
  <book isbn="978-4-7741-5715-3">  <- 属性(attribute)として
    <title>本のタイトル</title>      <- 必ず必要な要素として
    <author>著者の名前 1 </author>   <- 著者は 1 人以上
    <author>著者の名前 2 </author>
  </book>
<bookstore>
```

# WNPAデータ（国別団体）

```
< FSX4000000.xml >

1  <?xml version="1.0" encoding="UTF-8"?><OfferingXML Version="5.0">
2    <Header Category="0" CreateDate="2014/02/10" CreateTime="02:58:39" DeliveryXmlVersion="1"
3      ReportNumber="00003371" ResultSystemCode="FSX400000" XmlVersion="0"/>
4    <Meta CategoryTitle="オフィシャルリザルト" SubTitle="団体順位" Title="フィギュア" TransmitDate="
5      2014/02/10 02:58:39"/>
6    <Body>
7      <ResultList>
8        <ParticipantUnit CountryName="ロシア" IOCCode="RUS" MemberCount="15" Rank="1">
9          <Result Type="Point" Value="75"/>
10         <Member>
11           <Athlete FamilyName="フルシェンコ" FamilyNameEntry="Plyushchenko" FormatType="4"
12             GivenName="エフゲニー" GivenNameEntry="Evgeny" Id="2001933" TVName="Evgeny
13               PLYUSHCHENKO"/>
14           <ParticipantInfo Type="MemberNumber" Value="1"/>
15         </Member>
16         <Member>
17           <Athlete FamilyName="ソトニコワ" FamilyNameEntry="Sotnikova" FormatType="4" GivenName=
18             "アデリナ" GivenNameEntry="Adelina" Id="2001886" TVName="Adelina SOTNIKOVA"/>
19           <ParticipantInfo Type="MemberNumber" Value="2"/>
20         </Member>
21         <Member>
22           <Athlete FamilyName="リプニツカヤ" FamilyNameEntry="Lipnitskaya" FormatType="4"
23             GivenName="ユリア" GivenNameEntry="Yulia" Id="2001879" TVName="Yulia LIPNITSKAYA"/>
24           <ParticipantInfo Type="MemberNumber" Value="3"/>
25         </Member>
26         <Member>
27           <Athlete FamilyName="ボロソジヤル" FamilyNameEntry="Volosozhar" FormatType="4"
28             GivenName="タチアナ" GivenNameEntry="Tatiana" Id="2001872" TVName="Tatiana VOLOSOZHAR
29             "/>
30           <ParticipantInfo Type="MemberNumber" Value="4"/>
31         </Member>
32         <Member>
33           <Athlete FamilyName="トランコフ" FamilyNameEntry="TRANKOV" FormatType="4" GivenName=
34             "マキシム" GivenNameEntry="Maxim" Id="2001943" TVName="Maxim TRANKOV"/>
```

# WNPA個別データ

```
FSX400101S.xml < XML Editor >

1  <?xml version="1.0" encoding="UTF-8"?><OfferingXML Version="5.0">
2    <Header Category="S" CreateDate="2014/02/09" CreateTime="15:11:00" DeliveryXmlVersion="1"
3      ReportNumber="00000278" ResultSystemCode="FSX400101" XmlVersion="0"/>
4    <Meta CategoryTitle="スタートリスト" SubTitle="団体男子フリー" Title="フィギュア" TransmitDate="2014
/02/09 15:11:00"/>
5    <Body>
6      <StartList EventDay="2014/02/09" EventTime="19:00">
7        <ParticipantUnit CountryName="イタリア" IOCCode="ITA" MemberCount="1" StartingNumber="1">
8          <Member>
9            <Athlete FamilyNameEntry="Parkinson" GivenNameEntry="Paul Bonifacio" Id="2015464"
10              TVName="Paul Bonifacio PARKINSON"/>
11            </Member>
12            <ExerciseInfo Code="4S" Description="Quad Salchow" Order="1"/>
13            <ExerciseInfo Code="3A+2T" Description="Triple Axel+Double Toeloop" Order="2"/>
14            <ExerciseInfo Code="3A" Description="Triple Axel" Order="3"/>
15            <ExerciseInfo Code="CCoSp" Description="Change Foot Combination Spin" Order="4"/>
16            <ExerciseInfo Code="FCSp" Description="Flying Camel Spin" Order="5"/>
17            <ExerciseInfo Code="StSq" Description="Step Sequence" Order="6"/>
18            <ExerciseInfo Code="3Lz+3T" Description="Triple Lutz+Triple Toeloop" Order="7"/>
19            <ExerciseInfo Code="2A+1Lo+3S" Description="Double Axel+Single Loop+Triple Salchow"
20              Order="8"/>
21            <ExerciseInfo Code="3F" Description="Triple Flip" Order="9"/>
22            <ExerciseInfo Code="3Lo" Description="Triple Loop" Order="10"/>
23            <ExerciseInfo Code="ChSq" Description="Choreo Sequence" Order="11"/>
24            <ExerciseInfo Code="3Lz" Description="Triple Lutz" Order="12"/>
25            <ExerciseInfo Code="CSSp" Description="Change Foot Sit Spin" Order="13"/>
26          </Member>
27        <ParticipantUnit CountryName="米国" IOCCode="USA" MemberCount="1" StartingNumber="2">
28          <Member>
29            <Athlete FamilyName="ブラウン" FamilyNameEntry="Brown" FormatType="4" GivenName="
ジェーソン" GivenNameEntry="Jason" Id="2036672" TVName="Jason BROWN"/>
</Member>
<ExerciseInfo Code="2A" Description="Double Axel" Order="1"/>
<ExerciseInfo Code="3A+3T" Description="Triple Axel+Triple Toeloop" Order="2"/>
```

# ちなみに国際スケート連盟はPDF([www.isu.org](http://www.isu.org))

Rank	Name	Nation	Starting Number	Total Segment Score	Total Element Score	Program Component Score (factored)	Total	Total Deductions
2	Mao ASADA	JPN	18	121.29	56.08	65.21	0.00	
#	Executed Elements	Base Value	GOE	J1 J2 J3 J4 J5 J6 J7 J8 J9		Ref	Scores of Panel	
1	2A	3.30	0.75	1 2 2 1 2 1 1 2			4.05	
2	3F+2Lo	7.10	0.70	1 1 1 1 1 0 1 1			7.80	
3	3Lz	!	-0.23	-1 1 0 -1 0 0 0 -1			5.77	
4	FCSp3	2.80	0.50	1 1 1 1 2 0 1 1			3.30	
5	ChSq1	2.00	1.75	2 3 3 3 3 0 2 2			3.75	
6	2A+2T	5.06	x 0.50	1 1 1 1 1 1 0 1			5.56	
7	3S	4.84	x 0.58	1 1 0 1 1 0 2 1			5.42	
8	2F	2.09	x 0.00	0 0 0 -1 0 0 0 0			2.09	
9	3Lo	5.61	x 0.70	1 1 1 1 1 1 1 1			6.31	
10	CCoSp4	3.50	0.92	2 2 2 2 2 0 2 1			4.42	
11	LSp3	2.40	0.58	1 1 1 1 2 1 1 2			2.98	
12	StSq3	3.30	1.33	2 3 3 3 3 1 2 3			4.63	
		48.00					56.08	
	Program Components		Factor					
	Skating Skills		1.60	8.25 7.75 8.25 8.25 8.50 6.50 8.25 7.75			8.08	
	Transitions		1.60	7.25 8.00 8.00 7.25 8.00 5.75 7.75 7.50			7.63	
	Performance		1.60	8.50 8.75 8.75 7.75 8.50 6.00 8.50 8.00			8.33	
	Composition		1.60	8.00 8.50 8.00 8.25 8.75 6.25 8.50 8.00			8.21	
	Interpretation of the Music		1.60	8.25 8.50 8.75 8.75 8.75 6.00 8.75 8.00			8.50	
	Judges Total Program Component Score (factored)						65.21	
	Deductions						0.00	

x Credit for highlight distribution, base value multiplied by 1.1 ! Not clear edge

フィギュアは要素数も可変。減点要素も任意

## 参考資料

大学の1年生向け情報処理の講義

[http://www.coins.tsukuba.ac.jp/~yas/coins/  
literacy-2016/2016-05-06/](http://www.coins.tsukuba.ac.jp/~yas/coins/literacy-2016/2016-05-06/)