Cracking the Code

Tanya Goetz

Assistant Professor

New York City College of Technology

http://www.citytech.cuny.edu

```
{\rtf1\mac\ansicpg10000\cocoartf102
{\fonttbl\f0\fswiss\fcharset77 Helvetica;\f1\fswiss\fcharset77 Helvetica-
BoldOblique;\f2\fswiss\fcharset77 Helvetica-Bold;
}
{\colortbl;\red255\green255\blue255;}
\margl1440\margr1440\vieww9000\viewh7500\viewkind0
\pard\tx720\tx1440\tx2160\tx2880\tx3600\tx4320\tx5040\tx5760\tx6480\tx7200\tx7920\tx8640\ql\qnatura
l

\f0\fs24 \cf0 \
\
Greetings P3 Members!!\
\
How are you
\f1\i\b RIGHT
\f2\i0
\f0\b0 now?}
```
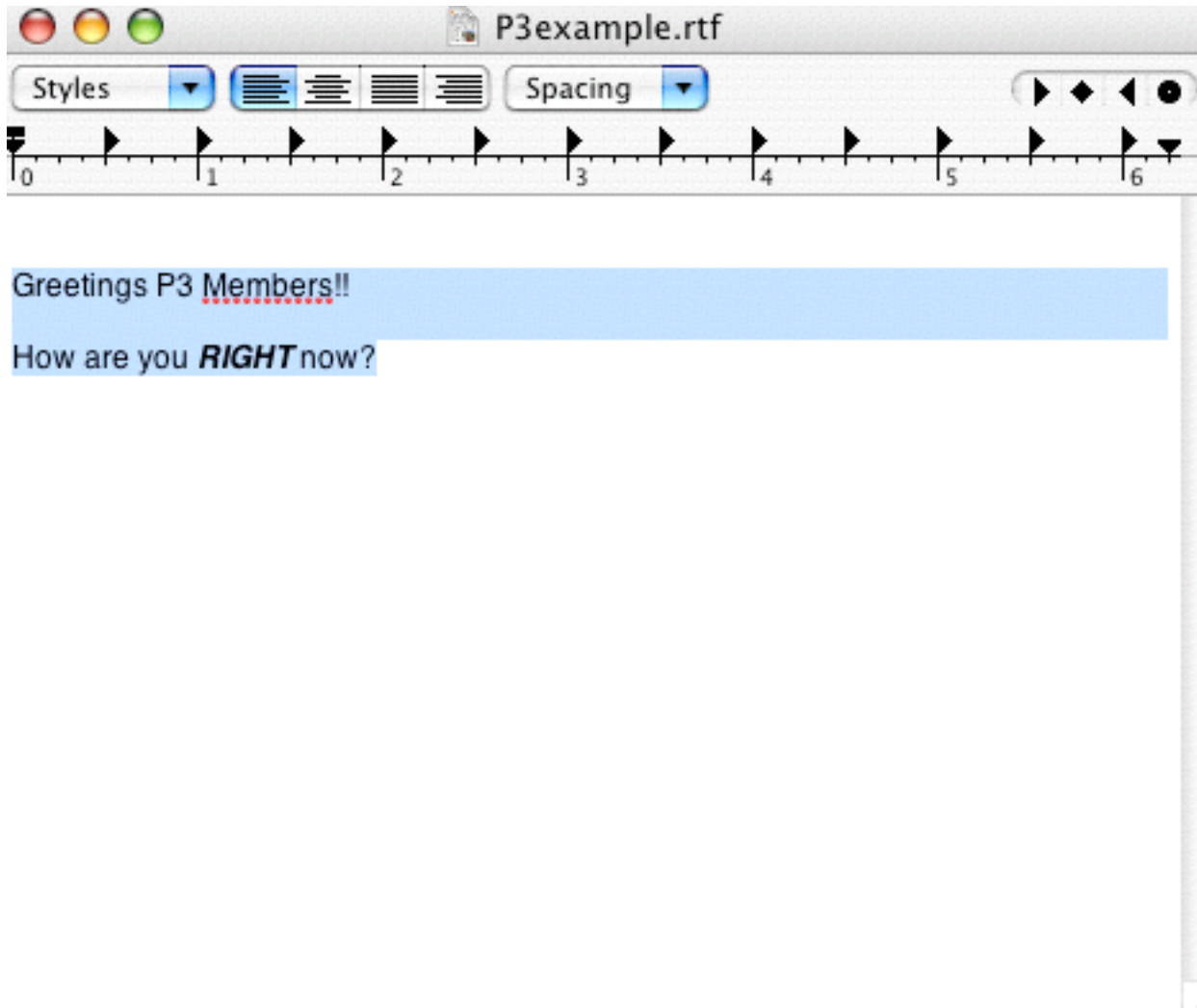
P3example.rtf

Styles | Spacing

0   1   2   3   4   5   6

Greetings P3 Members!!

How are you *RIGHT* now?

General information

Basic general information about the markup languages: creator, version, etc.

	Creator	First public release date	Latest stable version	Editor	Viewer
DocBook	The Davenport Group	1992	4.4	Text editor	HTML and PDF converters (For printing)
HTML	Tim Berners-Lee	1993	4.01	Text editor, HTML editor	Web browser
MathML	W3C	July 1999	2.0	Text editor, TeX converter	Web browser, Word processor
MIF	Adobe Systems	before or during 1993	7.0	Text editor, FrameMaker	FrameMaker
RTF	Microsoft	1987	1.8	Text editor, Word processor	Word processor
TeX	Donald Knuth	1978	3.141592	Text editor	DVI or PDF converter
XHTML	W3C	January 26, 2000	1.1	Text editor, HTML editor	Web browser
WML	WAP Forum	?	2.0	Text editor	Microbrowser
	Creator	First public release date	Latest stable version	Editor	Viewer

Characteristics

Some characteristics of the markup languages.

	Major purpose	Based on	Markup type	Structural markup	Presentational markup [†]
DocBook	Technical document	SGML / XML	Tag	Yes	No
HTML	Hypertext document	SGML	Tag	Yes	Yes
MathML	Mathematical document	XML	Tag	Yes	No
MIF	Technical document		Tag	Yes	Yes
RTF	Rich text document		Control code	Yes	Yes
TeX	Academic document		Control code	Yes	Yes
XHTML	Hypertext document	XML	Tag	Yes	No
WML	Hypertext document	XML	Tag	Yes	Yes
	Major purpose	**Based on**	**Markup type**	**Structural markup**	**Presentational markup**

- **^** Many markup languages have purposely avoided presentational markups. For markup languages based on SGML and XML, CSS is used as a presentation layer.

Three types of documents

- Text
- Image
- Page

Documents fall into one of these categories.

Text Files

- Can be saved as "open format," making it possible to move the file to other software and platforms or in a specific software format such as the ubiquitous .doc format from Microsoft Word.

- Program specific formats become problematic when transferred to different software, even earlier versions of the same software.

Open Text Format: ASCII

- ASCII (American National Standard Code for Information Interchange) is a standard format for digital information, both text and graphic, but usually used for text. It represents the text as numerical data since numbers are what computers understand.

- ASCII is "raw" text. It contains no information about the text's design or formatting (no bolds, italics, tabs etc.)

- The Good News: It will import anywhere

- The Bad News: You have no control over its appearance

For further information, see www.asciitable.com

And http://www.en.wikipedia.org/wiki/ASCII

ASCII Table and Description

ASCII stands for American Standard Code for Information Interchange. Computers can only understand numbers, so an ASCII code is the numerical representation of a character such as 'a' or '@' or an action of some sort. ASCII was developed a long time ago and now the non-printing characters are rarely used for their original purpose. Below is the ASCII character table and this includes descriptions of the first 32 non-printing characters. ASCII was actually designed for use with teletypes and so the descriptions are somewhat obscure. If someone says they want your CV however in ASCII format, all this means is they want 'plain' text with no formatting such as tabs, bold or underscoring - the raw format that any computer can understand. This is usually so they can easily import the file into their own applications without issues. Notepad.exe creates ASCII text, or in MS Word you can save a file as 'text only'

Dec	Hx	Oct	Char		Dec	Hx	Oct	Html	Chr	Dec	Hx	Oct	Html	Chr	Dec	Hx	Oct	Html	Chr
0	0	000	NUL	(null)	32	20	040	 	Space	64	40	100	@	@	96	60	140	`	`
1	1	001	SOH	(start of heading)	33	21	041	!	!	65	41	101	A	A	97	61	141	a	a
2	2	002	STX	(start of text)	34	22	042	"	"	66	42	102	B	B	98	62	142	b	b
3	3	003	ETX	(end of text)	35	23	043	#	#	67	43	103	C	C	99	63	143	c	c
4	4	004	EOT	(end of transmission)	36	24	044	$	$	68	44	104	D	D	100	64	144	d	d
5	5	005	ENQ	(enquiry)	37	25	045	%	%	69	45	105	E	E	101	65	145	e	e
6	6	006	ACK	(acknowledge)	38	26	046	&	&	70	46	106	F	F	102	66	146	f	f
7	7	007	BEL	(bell)	39	27	047	'	'	71	47	107	G	G	103	67	147	g	g
8	8	010	BS	(backspace)	40	28	050	(	(	72	48	110	H	H	104	68	150	h	h
9	9	011	TAB	(horizontal tab)	41	29	051	)	)	73	49	111	I	I	105	69	151	i	i
10	A	012	LF	(NL line feed, new line)	42	2A	052	*	*	74	4A	112	J	J	106	6A	152	j	j
11	B	013	VT	(vertical tab)	43	2B	053	+	+	75	4B	113	K	K	107	6B	153	k	k
12	C	014	FF	(NP form feed, new page)	44	2C	054	,	,	76	4C	114	L	L	108	6C	154	l	l
13	D	015	CR	(carriage return)	45	2D	055	-	-	77	4D	115	M	M	109	6D	155	m	m
14	E	016	SO	(shift out)	46	2E	056	.	.	78	4E	116	N	N	110	6E	156	n	n
15	F	017	SI	(shift in)	47	2F	057	/	/	79	4F	117	O	O	111	6F	157	o	o
16	10	020	DLE	(data link escape)	48	30	060	0	0	80	50	120	P	P	112	70	160	p	p
17	11	021	DC1	(device control 1)	49	31	061	1	1	81	51	121	Q	Q	113	71	161	q	q
18	12	022	DC2	(device control 2)	50	32	062	2	2	82	52	122	R	R	114	72	162	r	r
19	13	023	DC3	(device control 3)	51	33	063	3	3	83	53	123	S	S	115	73	163	s	s
20	14	024	DC4	(device control 4)	52	34	064	4	4	84	54	124	T	T	116	74	164	t	t
21	15	025	NAK	(negative acknowledge)	53	35	065	5	5	85	55	125	U	U	117	75	165	u	u
22	16	026	SYN	(synchronous idle)	54	36	066	6	6	86	56	126	V	V	118	76	166	v	v
23	17	027	ETB	(end of trans. block)	55	37	067	7	7	87	57	127	W	W	119	77	167	w	w
24	18	030	CAN	(cancel)	56	38	070	8	8	88	58	130	X	X	120	78	170	x	x
25	19	031	EM	(end of medium)	57	39	071	9	9	89	59	131	Y	Y	121	79	171	y	y
26	1A	032	SUB	(substitute)	58	3A	072	:	:	90	5A	132	Z	Z	122	7A	172	z	z
27	1B	033	ESC	(escape)	59	3B	073	;	;	91	5B	133	[	[	123	7B	173	{	{
28	1C	034	FS	(file separator)	60	3C	074	<	<	92	5C	134	\	\	124	7C	174	|		
29	1D	035	GS	(group separator)	61	3D	075	=	=	93	5D	135	]	]	125	7D	175	}	}
30	1E	036	RS	(record separator)	62	3E	076	>	>	94	5E	136	^	^	126	7E	176	~	~
31	1F	037	US	(unit separator)	63	3F	077	?	?	95	5F	137	_	_	127	7F	177	&#127;	DEL

Open Text Format: RTF

- RTF (Rich Text Format) is an open text format that also contains codes for formatting and specific typefaces.
- Developed by Microsoft in 1987
- Works well with Mac platform--textedit on Mac has as its default save option Rich Text format.
- This is a good choice to use when sending text documents to others

The Meta- Language: SGML

- Standardized Generalized Markup Language. Became an ISO (International Standards Organization) standard in 1986. It is a standard for the description of a marked up electronic text.
- SGML is designed for managing large documents that need to be output in different print formats.
- It is a standardized approach to defining the elements of a document. It does NOT specify any formatting. It specifies the rules for tagging elements.
- In SGML, tags are meant to describe contents, not the presentation. A tag would denote a headline, a product description, a movie review, a caption, etc. NOT the font or the leading, etc.
- device-independent and system independent

Markup Languages: SGML

- So, one would not use <I> , an italics command when one had a book title. Instead, one would use <booktitle>.

- Very important to realize this distinction as XML, the subset of SGML that is all the buzz now in our world, acts more like SGML than HTML in this crucial aspect.

For further information, see:
http://www.w3,org/MarkUP/SGML

Markup Languages: XML

- XtensibleMarkup Language, a subset of SGML
- It avoids the pitfalls of html in that is focuses on content not presentation. More rigid in its structure.
- XHTML (Xtentsible Hypertext Markup language)-the W3C's recommendation for replacing HTML with an XML version of HTML.Simon St. Laurent notes in his book, XML: A Primer that "XML requires a different focus, demanding that designers examine the way their documents are built rather than the way they are formatted.
- As with SGML, you need to use a DTD when using XHTML. XHTML has now replaced HTML

http://www.w3.org/MarkUp/Guide/

Markup Languages: HTML 5

- Hypertext Markup Language, a derivative of SGML, was used for presentation of text on the World Wide Web. It identifies headings and paragraphs. It identified content in a document BUT also specified presentation of that document. This later changed when XHTML was introduced which used Cascading Style Sheets for Presentation

- HTML 5 does allow for separate presentaton and content. See links below:

  http://www.w3.org/MarkUP/html-spec/

  http://www.w3.org/MarkUp/Guide/

- Browsers interpret the HTML code. Every browser interprets the code slightly differently

Terms You should know

- HTML Document

  - Plain text file, that can be created using a text editor (Notepad in Windows, or TextEdit in Mac OS)
  - A Web page editor

- Web Page Editors

  - Example: Adobe Dreamweaver
  - Allows you to create and edit the page visually without having to manually add markup tags

Terms You should know

- URL

  - Stands for Uniform Resource Locator
  - This is the standard for specifying the addresses of web pages and resources on World Wide Web

```
<!DOCTYPE html PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN"
   "http://www.w3.org/TR/html4/loose.dtd">
<html>
<head>
 <meta http-equiv="Content-Type" content="text/html; charset=UTF-8">
 <title>Dave Raggett's Introduction to HTML</title>
 <style type="text/css">
 body {
   margin-left: 10%;
   margin-right: 10%;
   font-family: sans-serif;
 }
 h2,h3,h4,h5,h6 { margin-left: -3% }
 pre {
    color: green; font-weight: bold;
    white-space: pre; font-family: "Courier New", monospace;
```

```html
<!DOCTYPE html>
<!--[if IE 9]> <html lang="en" class="ie9"> <![endif]-->
<!--[if !IE]><!-->
<html lang="en-us">
    <!--<![endif]-->
    <!-- BEGIN HEAD -->
    <head>
        <title>City Tech - New York City College of Technology</title>
        <meta name="robots" content="all" />
        <meta name="robots" content="index, follow" />
        <meta name="GOOGLEBOT" content="INDEX, FOLLOW" />



<!-- <meta charset="utf-8" /> -->
<meta http-equiv="Content-type" content="text/html; charset=utf-8">
<meta http-equiv="X-UA-Compatible" content="IE=edge">
<meta name="viewport" content="width=device-width, minimum-scale=1.0, maximum-scale=1.0, user-scalable=no">
<meta content="New York City College of Technology (City Tech) is the designated college of technology of The City University of New York, currently offering both baccalaure
<meta name="keywords" content="City Tech, College of technology, higher education, NYC College of Technology, Entertainment Tech, Hospitality Management, Nursing, City Tech
<meta content="CIS Web Team" name="author" />
<!-- <meta name="robots" content="noindex,nofollow"> -->
<meta name="msvalidate.01" content="A75D65E5F5BF70F45CCD0754BD405044" />
<!-- BEGIN GLOBAL MANDATORY STYLES -->
<link href='//fonts.googleapis.com/css?family=Roboto+Condensed:300italic,400italic,700italic,400,300,700&amp;subset=all' rel='stylesheet' type='text/css'>
<link href="/assets/socicon/socicon.css" rel="stylesheet" type="text/css" />
<link href="/css/plugins/bootstrap-social.css" rel="stylesheet" type="text/css" />
<link rel="stylesheet" href="https://use.fontawesome.com/releases/v5.8.1/css/all.css" integrity="sha384-50oBUHEmvpQ+11W4y57PTFmhCaXp0ML5d60M1M7uH2+nqUivzIebhndOJK28anvf" cro

<link href="/assets/font-awesome/css/font-awesome.min.css" rel="stylesheet" type="text/css" />
<!-- <link href="/assets/simple-line-icons/simple-line-icons.min.css" rel="stylesheet" type="text/css" /> -->
<link href="https://cdnjs.cloudflare.com/ajax/libs/simple-line-icons/2.2.4/css/simple-line-icons.css" rel="stylesheet" />
<link href="/css/plugins/animate.min.css" rel="stylesheet" type="text/css" />
<link href="/assets/bootstrap/css/bootstrap.min.css" rel="stylesheet" type="text/css" />
<!-- END GLOBAL MANDATORY STYLES -->
<!-- BEGIN: BASE PLUGINS  -->
<link href="/assets/revo-slider/css/settings.css" rel="stylesheet" type="text/css" />
<link href="/assets/revo-slider/css/layers.css" rel="stylesheet" type="text/css" />
<link href="/assets/revo-slider/css/navigation.css" rel="stylesheet" type="text/css" />
<link href="/assets/cubeportfolio/css/cubeportfolio.min.css" rel="stylesheet" type="text/css" />
<link href="/assets/owl-carousel/owl.carousel.css" rel="stylesheet" type="text/css" />
<link href="/assets/owl-carousel/owl.theme.css" rel="stylesheet" type="text/css" />
<link href="/assets/owl-carousel/owl.transitions.css" rel="stylesheet" type="text/css" />
<link href="/assets/fancybox/jquery.fancybox.css" rel="stylesheet" type="text/css" />
<link href="/assets/slider-for-bootstrap/css/slider.css" rel="stylesheet" type="text/css" />
<!-- END: BASE PLUGINS -->
<!-- BEGIN THEME STYLES -->
<link href="/css/base/plugins.css" rel="stylesheet" type="text/css" />
<link href="/css/base/components.css" id="style_components" rel="stylesheet" type="text/css" />
<link href="/css/base/themes/blue10.css" rel="stylesheet" id="style_theme" type="text/css" />
<link href="/css/base/custom.css" rel="stylesheet" type="text/css" />
<!-- END THEME STYLES -->
<link rel="shortcut icon" href="/favicon.ico" />
<!-- analytics -->
```

Firefox's view of CityTech's website: Found Under Tools/
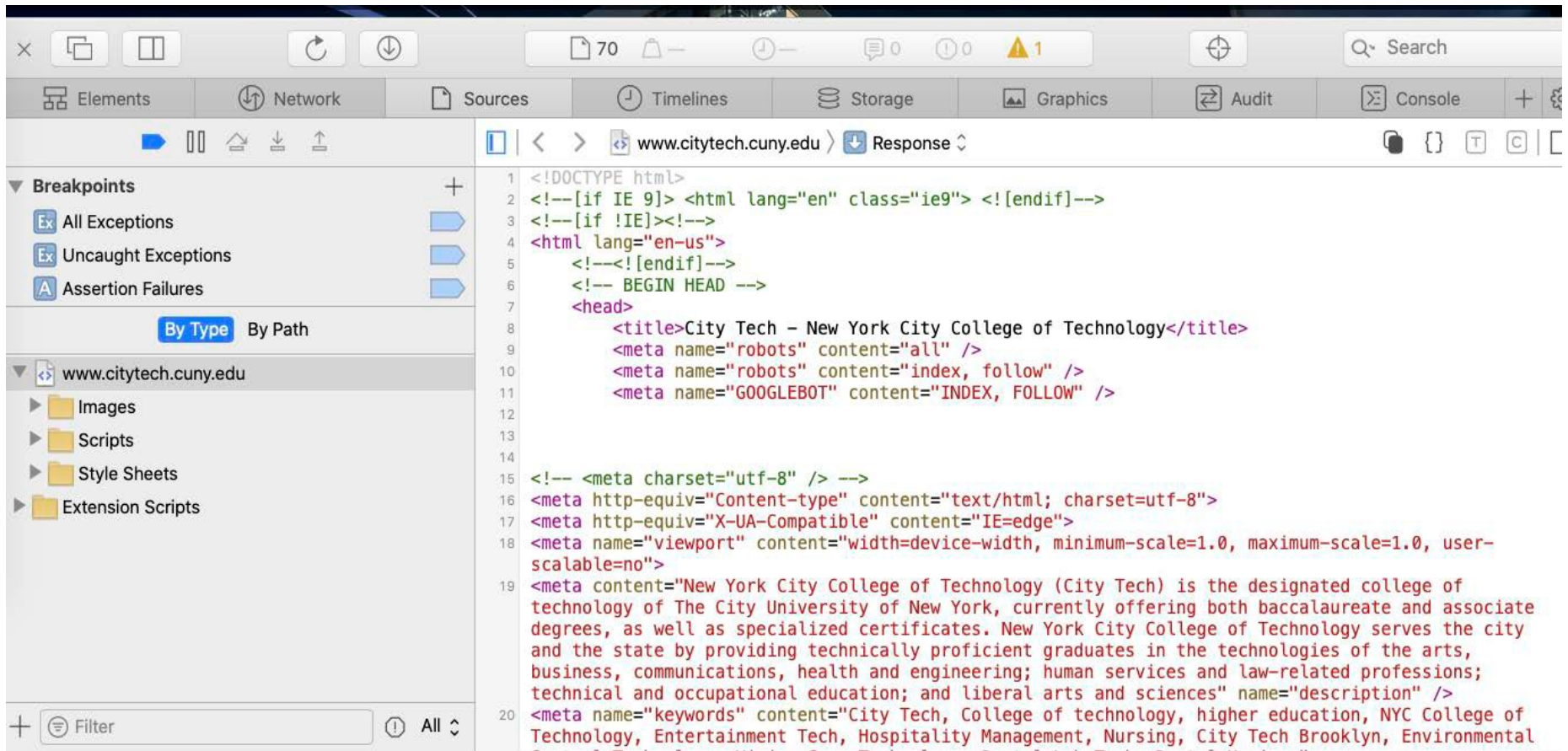Web Developer/Page Source/Command U

Here's how:

1. Open **Safari**.

2. Select 'Preferences' from the '**Safari**' menu.

3. In the 'Advanced' section and select 'Show Develop menu' in menu bar. '

4. Visit the **web page** you want to **view HTML** source for.

5. Select 'Show Page Source' from the 'Develop' menu that has been added to **Safari**.

Mar 4, 2013

www.socialmeteor.com › 2013/03/04 › how-to-view-ht... ▼

Quick Tip: How to 'View HTML Source' in Safari - Social Meteor

A look at the source Code from Safari's Developer Source Code

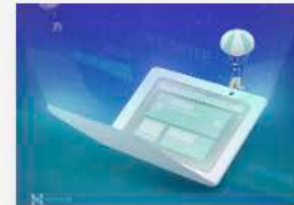https://www.hostinger.com/tutorials/best-html-editors

For the entire page . . .

The DeFacto PDL: Postscript

• Apage description language that RIPs (raster image processors) interpret for output to a proofer or to film (today it would most likely be output directly to the printing plate).

• A page description language is a programming language that describes the layout and appearance of a page. Quark Xpress, Adobe Indesign and even MS Word all write PostScript on the fly when you print to PostScript printer.

• Postscript is not the only Page Description language. HP printers, for example, use one called PCL (Printer Control Language).

• However, PostScript is used in almost all higher end graphic arts devices so it is by far the most important.

Profile: PostScript

- Vector File Format that can also contain bitmap data but key point is that text and line drawings are still a series of mathematical instructions rather than dots.

Identifying the PostScript File

- PostScript files are one file and usually have the .ps extension after it.

- One can open a Postscript file in a text editor or word processing program to view the code, but this will only tell you a limited amount of viable data and will not provide you with any visual clues.

```
%%%!PS-Adobe-2.0
%%Creator: dvips 5.521 Copyright 1986, 1993 Radical Eye Software
%%Title: html-spec.dvi
%%CreationDate: Thu Sep 21 18:28:17 1995
%%Pages: 76
%%PageOrder: Ascend
%%BoundingBox: 0 0 612 792
%%DocumentFonts: Times-Roman Times-Bold Courier Helvetica-Bold Helvetica
%%EndComments
%DVIPSCommandLine: dvips html-spec.dvi
%DVIPSSource: TeX output 1995.09.21:1828
%%BeginProcSet: tex.pro
/TeXDict 250 dict def TeXDict begin /N{def}def
```

Enter PDF:

Portable

Document

Format.

And a PDF is…

- **A Document Description format that describes a document for printing or display, It uses the imaging model of the PostScript language to render text & graphics.**
- **It is more efficient and predicable.**
- **It is a compact, sophisticated format for storing, transmitting, and displaying a wide variety of environments**
- **One can view it as the 21st century spawn of PostScript-- quicker, better and cheaper.**
- Essentially, **PDF is a shorthand version of PostScript. It eliminates idiosyncrasies within the PostScript instructions.**

For example, a cake recipe might be:

- Take the eggs and put them in a bowl. Take the milk and it in the same bowl. Take the sugar and put it in the bowl with the eggs and milk. Mix what is in the bowl.

Or I could say:

- Mix together eggs, sugar, and milk in a bowl.
- That's like what PDF does for PostScript files when you distill them. Same information, just standardized and compressed so that it is easier for the RIP (Raster Image Processor) to interpret the instructions. The RIP interprets the PDF file just as the browser interprets the html markup

But, PDF. . .

- Was created originally for the paperless office. It would solve the problems of exchanging documents across platforms that would preserve formatting of text and graphics. In other words, a page would look identical on my Mac and my boss's Windows NT machine.

- Therefore, PDF is designed for more applications that just high-end print and people can make PDF just for the screen or just for printing out on their laser printers.

- So, they can make a "bad" PDF for print unless you give very specific instructions. To solve this, a standard for the print industry called PDFX-1a was developed.

Which brings us to PDFX-1a… The X File

- American National Standards Institute specification for exchange of Digital Ads

- If PDF is a subset of PostScript, PDF-X1 is a subset of PDF that is designed only for printing on high-end devices.

- CMYK only and fonts must be embedded

- dpi is not specified as newspapers, magazines, etc. have different line screens

But wait, there's more: PDF-2

- this is a standard for the exchange of non-blind information between parties. (editorial pages, etc.)

- Therefore, PDFX-2 allows for an OPI -like (open prepress interface) workflow where the low-res is embedded in the file with the assumption that the high-res will be swapped out later before the page is ripped to a printing plate

- superset of PDFX-3 and therefore allows for device independent color spaces, LAB and ICC profiles.

Yet more: PDFX-3

- chief distinction is that this allows for tranferring data in CIELab or RGB with a profile attached
- Not used here in North America, but used in Europe

Adobe Acrobat is:

- The program that makes PDF, but it is NOT the only program that can do so.
- The part of Acrobat that makes PDFs from PostScript files is Acrobat Distiller
- Acrobat, the main body of the Application, allows one to add features such as links, bookmarks, notes, etc.

Acrobat Reader is:

- **The free viewer of PDF files that gives the view the ability to see the page and some navigational and search capabilities. One can also comment in Reader an view comments.**

Acrobat Professional is:

- **This program is for professional users who wish to create PDF files, preflight them, view separations and many other features. This is the program the a graphics arts professional should purchase.**