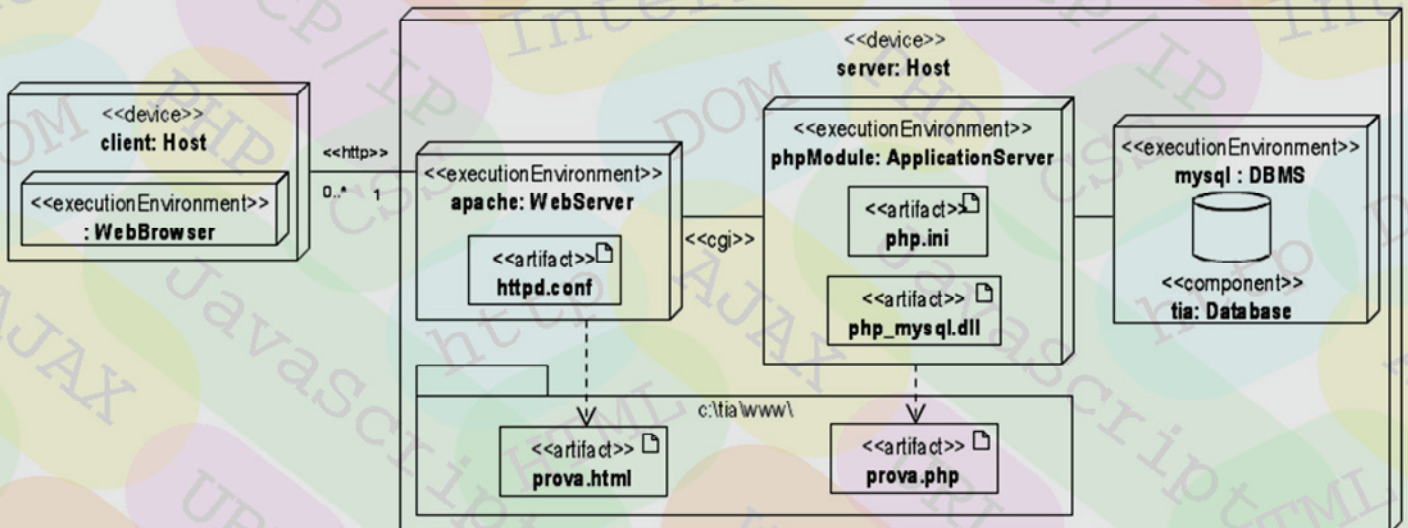


Marco Avvenuti Gabriele Cecchetti Mario G.C.A. Cimino

Lezioni e laboratori di tecnologie web lato server

con riferimento al linguaggio PHP



Prefazione

Dal punto di vista dell'utente il **Web** è una collezione di documenti collegati tra loro e accessibili tramite **Internet**, la rete globale costituita dalla interconnessione di reti di varia natura (private, pubbliche, aziendali, universitarie, commerciali, ecc.). Alla base dell'affermazione di Internet vi è un protocollo standard (**TCP/IP**) che permette alle diverse reti di scambiarsi dati in modo relativamente indipendente dagli apparati di connessione e dai sistemi operativi.

Ufficialmente, la nascita del Web risale al 6 agosto 1991, quando venne pubblicato il primo sito nella rete Internet. Vi sono oggi più di 100 milioni di siti attivi, 30 miliardi di pagine pubblicamente accessibili, e più di un miliardo di utenti (un sesto della popolazione mondiale). Nel Web non c'è limite al numero di pagine o di collegamenti (link) possibili; aggiungere una pagina non richiede coordinamento con alcuna autorità centrale e ha un costo addizionale estremamente basso; il protocollo che permette di seguire questi link (**HTTP**) è non discriminatorio, ossia consente di seguire ogni link indipendentemente dal contenuto o dalla proprietà della pagina. Il linguaggio di descrizione delle pagine web (**HTML**) è di pubblico dominio, le pagine possono essere scritte come testo in chiaro tramite un qualsiasi editor di testo, e sono quindi accessibili nella loro struttura, oltre che nel contenuto. Infine, grazie all'impiego di standard internazionali per le codifiche di carattere, una pagina HTML può essere scritta in qualsiasi lingua ed interpretata su qualsiasi piattaforma di sistema. Sin dall'inizio del Web, si poteva già "navigare" da un Macintosh, da un PC Windows o da una installazione Linux, e il formato HTML (assieme al formato PDF) è stato sin dagli anni novanta un formato privilegiato per lo scambio di documenti elettronici.

Un esempio attuale della estrema facilità di redazione, pubblicazione ed accesso al Web è il mondo dei Blog, divenuti rapidamente un popolare mezzo di espressione per qualsiasi settore, dalla politica all'arte, dalla scienza alla cronaca locale. L'approccio aperto e neutrale ha reso l'impatto sociale del Web più pervasivo e dinamico, e per certi versi antagonista, alle tradizionali reti di comunicazione, quali stampa e televisione. Il Web è diventato non solo una sede per il libero scambio di idee, ma anche una piattaforma per avere accesso ovunque a una serie di servizi innovativi: eBay, Google, Amazon, Wikipedia, Google Maps, iTunes, Youtube, BBC, Wheather, MySpace, SecondLife, Answers, Expedia, Monster, SourceForge, sono solo alcuni esempi della straordinaria capillarità resa possibile dalle tecnologie Web.

Il linguaggio HTML è oggi concepito per definire il contenuto logico e non l'aspetto del documento. Infatti, i terminali che possono accedere ad un file HTML sono molteplici, e per garantire la massima diffusione dei contenuti è opportuno evitare che i documenti siano destinati esclusivamente a classi ristrette di terminali, e quindi ad una élite di utenti. Non sempre i terminali sono dotati di adeguate capacità di presentazione grafica; possono anzi averne di limitate o non averne affatto (se ad esempio sono obsoleti, o sono terminali mobili, o sono adoperati da persone diversamente abili), o magari sono dispositivi che hanno necessità di estrarre informazioni dal documento senza visualizzarlo (ad esempio, per creare degli indici nei motori di ricerca, o per importare dei documenti HTML in un formato proprietario). Questo significa che uno stesso documento deve essere visualizzabile in modo diverso, e queste modalità di rappresentazione possono essere stabilite da un documento complementare, scritto in un linguaggio specifico (**CSS**), e diverso per ogni tipo di terminale.

Una volta prodotto, un documento web può essere pubblicato e reso accessibile attraverso un indirizzo (**URI**) formato da una sequenza di caratteri facili da memorizzare. Per trovare i documenti desiderati in questa gigantesca biblioteca sono indispensabili i motori di ricerca, i quali forniscono i link alle pagine attinenti ad alcuni termini di interesse, riordinandoli secondo un punteggio "neutralmente" stabilito in base al numero di collegamenti che conducono alle medesime pagine.

Inizialmente vi erano solo pagine HTML statiche, vale a dire file testuali arricchiti con informazioni di formattazione, e "navigabili" tramite opportune applicazioni (Web browser). Nonostante le limitazioni delle pagine statiche, la semplicità della tecnologia ne decretò un immediato successo, prima in campo accademico e di ricerca, quindi anche in quello commerciale, facendo nascere quella che oggi è anche nota come era del Web.

*Per superare i limiti delle pagine statiche, sin da subito furono definiti strumenti per generare dinamicamente pagine HTML. La prima soluzione di questo tipo fu l'interfaccia **CGI**, un meccanismo per consentire al sistema servente (o "server", che fornisce pagine Web) l'invocazione di applicazioni e la produzione "al volo" di contenuti, ad esempio adoperando dati estratti da un database interrogato in accordo alle specifiche inserite dall'utente tramite una precedente pagina. Per fornire maggiore interattività e dinamicità alle pagine, da un lato i browser furono potenziati tramite linguaggi di scripting (**JavaScript**) in grado di modificare le pagine direttamente sul terminale (o "client"), dall'altro nacquero dei linguaggi di programmazione specifici per la produzione dinamica di pagine sul lato server (tra i più diffusi abbiamo oggi **JSP**, **ASP.NET**, e **PHP**), trasformando i Web server da semplici archivi di pagine a motori di produzione delle medesime (application server).*

La complessa evoluzione degli standard del Web è stata da sempre guidata dal W3C, un organismo composto dalle principali aziende informatiche, società, associazioni, università ed istituti di ricerca al mondo. Ad esempio, sin dagli inizi il W3C ha cercato di combattere la progettazione di siti ad-hoc per terminali con specifiche risoluzioni (le etichette "best viewed with 800x600 screen resolution") o per specifici browser ("Best viewed by ? browser"), e qualsiasi schema di progetto che privilegiasse alcune tipologie di terminale. Ciò era particolarmente difficile quando i dispositivi palmari, i terminali industriali connessi ad internet, o i browser per utenti diversamente abili, erano ancora una rarità, e tutti in ufficio o a casa sembravano avere lo stesso tipo di personal computer. Un'altra questione che da sempre ha costituito un ostacolo alle potenzialità del Web è la cosiddetta "guerra dei browser", ossia il conflitto commerciale e di immagine tra diversi produttori che tentano di imporsi sul mercato, differenziandosi e implementando diverse concezioni degli standard sia nei browser che nelle piattaforme di sviluppo. Persino oggi, vi sono dei costrutti che occorre scrivere in due diverse versioni, perché non esiste una sola forma sintattica valida per tutti i principali browser.

In un contesto dalle esigenze sempre sempre crescenti, l'attività del W3C è stata caratterizzata da continue evoluzioni della struttura dei linguaggi. Per garantire compatibilità all'indietro, le successive versioni dei linguaggi mantengono molti dei costrutti precedenti etichettandoli come "disapprovati" ma al contempo validi all'interno di uno schema "di transizione". Di conseguenza, i browser divennero sempre più "tolleranti", al fine di poter visualizzare pagine scritte in versioni diverse, e contemporaneamente in grado di comprendere estensioni dello standard, specifiche per ogni browser e anticipatrici rispetto agli standard. Questo ha comportato anche problemi di sicurezza dei browser, che si sono protratti anche dopo la formulazione di un modello per le restrizioni dello scripting lato client. L'evoluzione dell'HTML culminò nel dicembre 1999 con una versione "stabile" ed ancora oggi di riferimento, HTML 4.01. L'evoluzione parallela del CSS portò nel 2004 ad emanare la specifica CSS 2.1, ancora oggi di riferimento perché supportata dai principali browser.

È estremamente importante scrivere pagine web ponendo attenzione agli standard di riferimento. Tutti i principali browser sanno interpretare oggi l'HTML 4.01 ed il CSS 2.1 in modo abbastanza uniforme. Ma, per poter visualizzare anche i milioni di documenti pubblicati in passato, essi sono in grado di visualizzare documenti non corretti secondo tali standard, ed addirittura documenti non corretti secondo alcuno standard. Tuttavia, poiché non tutti i browser riparano i documenti nello stesso modo, ciò introduce delle differenze, così che un documento non valido può mostrarsi e lavorare in maniera diversa su browser differenti. Queste differenze diventano ancora maggiori considerando le diverse versioni del JavaScript. Il W3C conduce da tempo una eccellente campagna di "bonifica" volta alla produzione di pagine formalmente corrette, attraverso la diffusione di servizi di "validazione", e di targhette di riconoscimento delle pagine valide.

*Particolarmente sul piano delle piattaforme di scripting è ancora in atto una competizione sui diversi linguaggi e sui relativi ambienti di sviluppo e di esecuzione, per cui oggi una qualsiasi soluzione web-based di una certa dimensione coinvolge l'accesso a piattaforme ibride. Questo si traduce nella necessità di dover sviluppare adoperando diverse sintassi, di sapersi orientare nella consultazione di diverse specifiche, e di saper programmare secondo uno stile facilmente portabile da un linguaggio all'altro. È quindi di importanza fondamentale per un progettista web avere una conoscenza dei paradigmi quali l'Object-Oriented (**OO**), il Document Object Model (**DOM**), l'architettura **three-tier**, il pattern model-view-controller (**MVC**) che consentono di vedere i sistemi ad un livello di astrazione superiore, ponendo in secondo piano la sintassi e mettendo in condizioni di*

identificare le analogie tra i diversi linguaggi, ad esempio tra JavaScript e VBScript, tra ASP.NET, JSP e PHP.

L'obiettivo di questo testo è quello di fornire una corretta metodologia e un'idonea collezione di strumenti e di standard aperti per poter progettare e sviluppare un servizio Web completo. Il materiale di questo testo costituisce parte integrante dell'insegnamento di Tecnologie Informatiche Applicate tenuto alla Facoltà di Ingegneria, corso di laurea in Ingegneria Informatica, e dell'insegnamento di Tecnologie Web nel Master di I livello in Tecnologie Internet dell'Università di Pisa.

Il codice sorgente presente nel testo è liberamente scaricabile all'indirizzo:

`http://tweb.ing.unipi.it`

con nome utente:

e password:

Indice

1	IL LINGUAGGIO PHP	9
1.1	INTRODUZIONE	9
1.2	SINTASSI DI BASE DI PHP	9
1.2.1	<i>Inclusione di PHP in una pagina HTML</i>	9
1.2.2	<i>Separatori delle istruzioni</i>	10
1.2.3	<i>Commenti</i>	10
1.3	TIPI E VARIABILI	10
1.3.1	<i>Stringhe</i>	12
1.3.2	<i>Array</i>	14
1.3.3	<i>Funzioni sulle stringhe</i>	15
1.3.4	<i>Funzioni sugli array</i>	16
1.3.5	<i>Conversioni di tipo</i>	17
1.3.6	<i>Variabili statiche e globali</i>	18
1.3.7	<i>Variabili predefinite</i>	19
1.4	OPERATORI	19
1.5	FUNZIONI DEFINITE DALL'UTENTE	20
1.6	STATEMENTS	22
1.7	DIRETTIVE DI INCLUSIONE	24
1.8	ACQUISIZIONE DEI DATI DALLE FORM HTML	25
1.9	USO DEI COOKIES	26
1.10	USO DELLE SESSIONI	27
1.11	PRINCIPALI FUNZIONI PREDEFINITE SUI DATI:	28
1.12	FUNZIONI SUI FILE	30
1.13	ESPRESSIONI REGOLARI	32
1.14	ALTRE FUNZIONI	34
1.15	OGGETTI CLASSE	34
1.16	PHP E MYSQL	37
1.16.1	<i>Connessione e Disconnessione</i>	37
1.16.2	<i>Selezionare un database</i>	37
1.16.3	<i>Inviare una query</i>	38
1.16.4	<i>Visualizzare il risultato di una query</i>	38
1.16.5	<i>Una semplice applicazione</i>	39
2	LABORATORI	43
2.1	ARCHITETTURE WEB A TRE STRATI	43
2.1.1	<i>Traccia</i>	43
2.1.2	<i>Soluzione proposta</i>	44
2.2	PHP CORE	51
2.2.1	<i>Traccia</i>	52
2.2.2	<i>Soluzione proposta</i>	54
2.3	DATABASE-DRIVEN WEB SYSTEMS	57
2.3.1	<i>Traccia</i>	57
2.3.2	<i>Soluzione proposta</i>	59
3	F.A.Q	67
3.1	PHP	67
3.1.1	<i>Sito oscurato</i>	67
3.1.2	<i>Iniezione html</i>	67
3.1.3	<i>Passaggio variabili</i>	68
3.1.4	<i>Gestione files</i>	70
3.1.5	<i>Includere frammenti esterni</i>	70
3.1.6	<i>http GET</i>	71
3.1.7	<i>http POST</i>	72
3.1.8	<i>File upload</i>	73
	INDICE ANALITICO	75
	INDICE DEI SITI WEB	77

Indice Analitico

\$

\$_GET, PHP; 26
\$_POST, PHP; 26
\$_REQUEST, PHP; 26
\$_SERVER, PHP; 19
\$_SESSION, PHP; 27
\$GLOBALS, PHP; 18
\$php_errormsg, PHP; 20
\$this, classi PHP; 35

&

&, riferimento, PHP; 11

.

., concatenazione stringhe PHP; 13

:

:
, risolutore PHP; 35

@

@, operatore di controllo errori,
PHP; 20

—

__CLASS__, PHP; 12
__construct, PHP; 54
__FILE__, PHP; 12
__FUNCTION__, PHP; 12
__LINE__, PHP; 12
__METHOD__, PHP; 12

<

<<, heredoc PHP; 13

=

===, identità operatore PHP; 37

A

abs, PHP; 28; 53
acos, PHP; 28
apache, installazione; 43
Apache-PHP-MySQL; 58
architetture a livelli; 43
Array, PHP; 14
array_combine, Array PHP; 17
array_fill, PHP; 14
array_flip, Array PHP; 17
array_keys, Array PHP; 17
array_merge, Array PHP; 17

array_pop, Array PHP; 17
array_push, Array PHP; 17
array_values, Array PHP; 16
asin, PHP; 28
asort, Array PHP; 17
atan, PHP; 28

C

ceil, PHP; 28; 53
chr, stringhe PHP; 15
class, PHP; 34
commenti, PHP; 10
compact, Array PHP; 17
conversioni di tipo, PHP; 17
cookies, PHP; 26
copy, file PHP; 31
cos, PHP; 28
count array, PHP; 53
crypt, stringhe PHP; 15

D

date, PHP; 29
debugging PHP; 54
define, costanti PHP; 12
do statement, PHP; 23

E

echo, PHP; 53
echo, stringhe PHP; 15
espressione regolare, PHP; 32
exp, PHP; 28
explode, stringhe PHP; 15

F

fclose, file PHP; 31
feof, file PHP; 31
fflush, file PHP; 31
fgets, file PHP; 31
fgetss, file PHP; 31
file, file PHP; 31
file_exists, file PHP; 31
filesize, file PHP; 31
flock, file PHP; 31
floor, PHP; 28
fopen, file PHP; 30
for statement, PHP; 23
foreach, PHP; 23
form, prelievo dati, PHP; 25
fputs, file PHP; 31
fread, file PHP; 31
fseek, file PHP; 31
funzione, PHP; 20
funzioni variabili, PHP; 22

G

getdate, PHP; 30
gettype, PHP; 11
global, PHP; 18

I

if statement, PHP; 22
implode, PHP; 37
implode, stringhe PHP; 15
import_request_variables, PHP; 26
include, PHP; 24
include_once, PHP; 25
inclusione, PHP; 9
iniezione di SQL, PHP; 59
intval, PHP; 53
is_array, PHP; 34
is_double, PHP; 34
is_float, PHP; 34
is_int, PHP; 11; 34
is_null, PHP; 11
is_object, PHP; 34
is_string, PHP; 34

K

ksort, Array PHP; 17

L

localtime, PHP; 30
log, PHP; 28

M

max, PHP; 28; 53
min, PHP; 28; 53
mkdir, file PHP; 32
mktime, PHP; 30
mysql e PHP; 37
mysql, installazione; 43
mysql_close, PHP; 37
mysql_connect, PHP; 37
mysql_error, PHP; 37
mysql_fetch_array, PHP; 38
mysql_free_result, PHP; 39
mysql_query, PHP; 38
mysql_select_db, PHP; 37

N

new, classi PHP; 35
NULL, PHP; 11
number_format, stringhe PHP; 15

O

operatori, PHP; 19

Indice Analitico

ord, stringhe PHP; 15
ordinamento array, PHP; 17
overloading funzioni , PHP; 21

P

parent, classe PHP; 35
parse_str, stringhe PHP; 15
passaggio parametri, PHP; 21
php, installazione; 43
PHPSESSID, PHP; 27
pi, PHP; 28
porting Java to PHP; 54
pow, PHP; 28
preg_match, espressione regolare
PHP; 33
preg_match_all, espressione
regolare PHP; 33
preg_replace, espressione regolare
PHP; 33
preg_split, espressione regolare
PHP; 33
print_r, stampa oggetti PHP; 34

Q

QUERY_STRING, PHP; 26

R

rand, PHP; 28

require, PHP; 24
require_once, PHP; 25
return, PHP; 21
rmdir, file PHP; 32
round, PHP; 28; 53

S

sequenze di escape, PHP; 13
serialize, classi PHP; 36
session_start, PHP; 27
sessione, PHP; 27
setcookie, PHP; 26
settype, PHP; 11
sin, PHP; 28
sizeof, Array PHP; 17
sleep, PHP; 37
sort, Array PHP; 17
split, PHP; 53
sqrt, PHP; 28
srand, PHP; 28
static, PHP; 54
str_split, stringhe PHP; 16
strcasecmp, stringhe PHP; 16
stringhe, PHP; 12
strip_tags, stringhe PHP; 16
strlen, PHP; 14
strlen, stringhe PHP; 16
strncmp, stringhe PHP; 16
strrev, stringhe PHP; 16
strstr, stringhe PHP; 16
strtotime, PHP; 29

strts, stringhe PHP; 16
strval, PHP; 13
superglobals, PHP; 19
switch statement, PHP; 23

T

tan, PHP; 28
time, PHP; 27; 30
tipi fondamentali, PHP; 10
tmpfile, file PHP; 32

U

unlink, file PHP; 32
unserialize, classi PHP; 36
unset, Array PHP; 16
unset, PHP; 12; 35

V

variabili predefinite, PHP; 19
variabili statiche e globali, PHP;
18

W

weakup, PHP; 37
while statement, PHP; 23

Indice dei siti web

[1]	http://tools.ietf.org/html/rfc1180	TCP/IP
[2]	http://www.domaintools.com/internet-statistics/	statistiche
[3]	http://news.netcraft.com/archives/web_server_survey.html	statistiche
[4]	http://www.boutell.com/newfaq/misc/sizeofweb.html	statistiche
[5]	http://www.internetworldstats.com/stats.htm	statistiche
[6]	http://www.geohive.com/default1.aspx	statistiche
[7]	http://www.w3.org/Protocols/	HTTP
[8]	http://tools.ietf.org/html/rfc3986	URI
[9]	http://www.ebay.com/	Ebay
[10]	http://www.altavista.com/	Altavista
[11]	http://www.live.com/	Live
[12]	http://www.google.com	Google
[13]	http://www.amazon.com/	Amazon
[14]	http://wikipedia.org/	Wikipedia
[15]	http://maps.google.com/	Google Maps
[16]	http://www.apple.com/itunes/	iTunes
[17]	http://www.youtube.com/	Youtube
[18]	http://www.bbc.co.uk	BBC
[19]	http://www.weather.com/	Weather
[20]	http://docs.sun.com/app/docs/coll/S1_Javascript_13	Javascript
[21]	http://www.myspace.com/	MySpace
[22]	http://secondlife.com/	Secondlife
[23]	http://www.answers.com/	Answers
[24]	http://www.expedia.com/	Expedia
[25]	http://www.monster.com	Monster
[26]	http://www.w3schools.com/browsers/default.asp	browsers
[27]	http://java.sun.com/products/jsp/	JSP
[28]	http://www.asp.net/	ASP
[29]	http://www.php.net/download-docs.php	PHP
[30]	http://www.adobe.com/devnet/pdf/pdf_reference.html	PHP
[31]	http://www.w3.org/CGI/	CGI
[32]	http://www.omg.org/	OO
[33]	http://www.w3.org/DOM/	DOM
[34]	http://msdn2.microsoft.com/en-us/library/d1wf56tt.aspx	VBScript
[35]	http://sourceforge.net/	SourceForge
[36]	http://www.textpad.com/	Textpad
[37]	http://www.microsoft.com/windows/ie/downloads/default.msp	IE
[38]	http://www.mozilla.com/firefox/	Firefox
[39]	http://channels.netscape.com/ns/browsers/default.jsp	Netscape
[40]	http://www.opera.com/download/	Opera
[41]	http://www.chami.com/html-kit/download/	HTML-Kit
[42]	http://validator.w3.org/file-upload.html	HTML
[43]	http://www.apple.com/quicktime/download/	QuickTime
[44]	http://www.w3.org/TR/html401/	HTML 4.01
[45]	http://www.w3.org/TR/CSS21/	CSS 2.1
[46]	http://jigsaw.w3.org/css-validator/	CSS
[47]	http://www.ecma-international.org/publications/standards/Ecma-262.htm	ECMA
[48]	http://www.w3.org/TR/DOM-Level-3-Core	DOM
[49]	http://www.zvon.org/index.php?nav_id=3	DOM
[50]	http://www.w3.org/DOM/	DOM
[51]	http://dmoz.org/Computers/Programming/Internet/W3C_DOM/	DOM
[52]	http://zvon.org/xxl/DOM2reference/Output/index.html	DOM
[53]	http://www.php.net/download-docs.php	PHP
[54]	http://httpd.apache.org/download.cgi	Apache
[55]	http://www.php.net/downloads.php	PHP
[56]	http://dev.mysql.com/downloads/mysql/5.0.html	MYSQL
[57]	http://dev.mysql.com/downloads/gui-tools/5.0.html	MYSQL
[58]	http://www.gimp.org/	Gimp
[59]	http://www.pdf995.com	Pdf995
[60]	http://audacity.sourceforge.net/	Audacity
[61]	http://www.httrack.com/page/2/en/index.html	Httrack
[62]	http://www.aptrio.com/Multimedia/Multimedia/hypercam-12929.html	HYpercam
[63]	http://virtualdub.sourceforge.net/	Virtualdub
[64]	http://www.poyoli.de/mobile/pdasim.htm	PDA
[65]	http://arzimagina.awardspace.com/noticias_amplia.php?id=10#	PDA
[66]	http://pda.startpagina.nl/	PDA
[67]	http://developer.mozilla.org/it/docs/AJAX:Iniziare	AJAX