

---

A proposito di Internet e Computers

# I contenuti e gli strumenti per trattarli

Michele Fucito  
Aprile 2016

- Il termine "**contenuto**" viene oggi generalmente utilizzato per indicare qualsiasi possibile "incarnazione" di una "**proprietà intellettuale**".
- Esiste cioè una stretta relazione con il concetto di "proprietà intellettuale" che indica "i contenuti di novità" che si possono trovare in una certa espressione o realizzazione.
- Con una tale definizione si rimuove quindi la dipendenza dal formato fisico di una certa "**cosa**" puntando invece l'attenzione su quanto di quella cosa esprime "**valore**".
- Questo passaggio, apparentemente solo di nomenclatura, in realtà è essenziale in quanto oggi, grazie al progresso tecnologico, è possibile prescindere dalla forma fisica mantenendo "il valore" in forme e modalità diverse, in funzione degli utilizzi che si vuole fare di quel "contenuto".
  
- Qualche esempio può essere interessante:
  - Un libro era tradizionalmente un oggetto contenente le idee dell'autore, espresse con un suo linguaggio , accessoriando il testo con delle illustrazioni ....
  - il testo/bozza veniva passato ad un "editore" che creava un oggetto fisico, "il libro" , che veniva stampato su carta, rilegato, fatto arrivare sugli scaffali delle librerie...
  - chi voleva accedere ai contenuti del libro si recava in libreria e acquistava una copia del libro: da quel momento il "cliente" era in grado di conoscere "i contenuti" che l'autore aveva voluto esprimere.
  - l'autore "protegeva" la sua "proprietà intellettuale" con il famoso "copyright"
  - chi acquistava un libro non era autorizzato a farne delle copie nè a distribuirle...

- Vediamo oggi come si è modificato lo scenario:
  - Un autore esprime le sue idee creando con un computer un "documento" nel quale inserisce una serie di elementi testuali, immagini, riferimenti, etc. che si concretano in un "file" conservato sulla memoria di massa del suo computer.
  - Volendo l'autore può direttamente diffondere tramite i canali "internet" il documento che sostanzia la sua opera di ingegno senza ulteriori passaggi...
  - Se però l'autore vuole rendere fruibile la sua opera anche a chi preferisce avere in mano il bel "libro tradizionale" , i files contenenti il libro vengono passati ad un editore il quale provvede a creare un libro tradizionale che successivamente viene stampato e distribuito nelle librerie.
- E' evidente che quello che cambia rispetto al passato è "**la catena di delivery**" del "**contenuto**" ovvero delle idee espresse nel libro, fermo restando le idee stesse.
  - nel caso tradizionale si usava la catena
    - **autore -> editore -> libraio -> consumatore**
  - nel caso attuale la catena diventa duplice:
    - **autore -> editore -> libraio -> consumatore**
    - **autore -> internet -> consumatore**
- Si nota che accanto alla catena di delivery tradizionale si ha oggi una catena di delivery "**immateriale**" che presenta una serie di caratteristiche interessanti:
  - molto più veloce: minimizzazione del "time-to-market"
  - assenza di passaggi intermedi: molto più economica
  - molto più "gestibile" ad es. archiviazione, trasmissione, consultazione, etc.

- Volendo potremmo fare altri esempi:
  - le opere musicali o cinematografiche: si passa dal disco/pizza di pellicola al CD/DVD o semplicemente ai files scaricati da internet
  - le opere grafiche: si passa dalle stampe ai files distribuiti tramite vari supporti o via internet
  - etc. etc.
- Esistono esempi ancora più sbalorditivi:
  - il progetto di un circuito elettronico ---> diventa un file in un opportuno formato che fornito in ingresso ad una certa macchina produce un "chip" utilizzabile
  - il progetto di un pezzo meccanico --> diventa un file che dato in pasto ad una opportuna macchina ( es. stampante 3D o macchina a controllo numerico) produce il pezzo meccanico finito.
  - etc. etc.
- Tutto questo pone o lascia inalterati una serie di temi:
  - come proteggere la "proprietà intellettuale"
  - come rendere realmente fruibile il contenuto da parte dei potenziali utilizzatori
  - come conservare in modo sicuro i contenuti
  - come spostare nello spazio i contenuti ( dove va a finire la libreria o la biblioteca )
  - con quali strumenti creare, modificare o accedere ai contenuti

- Un elemento chiave di tutte le problematiche è "il formato" da utilizzare per rappresentare un certo contenuto, preso atto che esso è ormai "immateriale"
- Per "**formato**" si intende tutta una serie di "**regole**" e "**metodologie**" atte ad esprimere in maniera completa ed esauriente "il contenuto", consentendo di implementare le necessarie azioni di **protezione della "proprietà intellettuale"** e di "**accessibilità**" allo stesso.
- Si capisce allora che la creazione o la scelta di un "formato" di rappresentazione di un contenuto è un tema estremamente critico in funzione della tipologia di "proprietà intellettuale" presente nel contenuto stesso.
- Ovviamente più strette sono le esigenze, più complessi diventano i formati e più difficoltoso diventa accedere ai contenuti.
- E' anche evidente che molto spesso esistono esigenze assolutamente diverse:
  - un poesia o un pezzo di musica fatto da un esordiente ....
  - una canzone di un big della musica....
  - un film di prima visione ....
  - un testo di una relazione tecnica ...
  - gli appunti di un corso all'università...
- In pratica esiste una enormità di formati che spesso nascono anche per esigenze diverse e spesso occulte...
  - formati di documenti ( MS word e simili proprietari, formati aperti, etc.)
  - volontà di legare un formato ad uno "strumento" per ostacolare la concorrenza
  - etc. etc.

## Il ruolo degli "strumenti" ....

---

- Quando di parla di "**strumenti**" con riferimento ai contenuti, ormai ci si riferisce sempre a "**programmi SW**" e a "piattaforme HW" su cui tali programmi possono funzionare.
- Molto spesso accade che un certo "strumento" sia strettamente legato ad una certa "piattaforma HW", ma sempre più spesso si assiste alla "**indipendenza dalla piattaforma HW**" proprio a voler sottolineare la concentrazione del valore di tali strumenti nel solo SW.
- Esistono una molteplicità di strumenti che si differenziano per la tipologia di contenuti che consentono di trattare e/o per la tipologia di trattamento che consentono di realizzare su tali contenuti.
- Spesso si trovano strumenti diversi che realizzano esattamente le stesse funzioni con solo delle piccole sfumature spesso limitate all'aspetto "visivo" e alla "facilità di uso".
- Molto spesso si assiste alla comparsa di strumenti che si differenziano solo per la presenza di funzioni accessorie spesso del tutto "esoteriche" o create al solo scopo di battere la concorrenza.
- Molto spesso ci si trova di fronte a problemi di "formati proprietari" che limitano le possibilità di manipolazione dei contenuti costringendo ad usare un certo tool .
- Per questi motivi esiste ormai da parecchio tempo una forte spinta verso l'adozione, soprattutto nelle "**amministrazioni pubbliche**", di "**formati aperti**" ovvero non proprietari e coperti da standard internazionali accessibili a tutti: è un modo per minimizzare le barriere di accesso ai contenuti e **abbassare i costi favorendo la concorrenza tra fornitori di "strumenti" per il trattamento dei contenuti stessi.**

- Esistono varie tipologie di strumenti:
  - strumenti di tipo "office" : consentono di creare, modificare, visionare documenti di testo, arricchiti di elementi aggiuntivi quali ad es. immagini, tabelle, etc.
    - word processor : gestione testi
    - spreadsheet : specifico per tabelle o fogli di calcolo
    - presentazioni: per realizzare slide show o presentazioni
    - database: per gestire dati in maniera articolata
  - strumenti di grafica e manipolazione video: consentono di sviluppare e gestire progetti grafici dai più semplici ai più complessi, integrando foto, video, etc.
  - strumenti di CAD ( computer aided design) per lo sviluppo e la gestione di progetti complessi relativi a varie discipline quali ad esempio meccanica, elettronica, etc.
  - strumenti ERP per la gestione dei flussi di lavoro nelle imprese
  - strumenti di simulazione per lo sviluppo e la gestione di progetti complessi spesso di natura scientifica quali ad es. la progettazione elettronica, il testing di assiemi complessi etc.
- Esistono poi strumenti "generici" il cui unico scopo è quello di "contenere" altri strumenti rendendoli facilmente interoperabili in modo da consentire un facile accesso alle informazioni indipendentemente dal loro formato: sono la classe dei "**Browser**"
  - **Con il termine "browser" si intende in genere uno strumento che rende possibile l'accesso tramite internet ad opportuni "siti" di informazione su cui sono disponibili una serie di contenuti.**
  - A seconda del contenuto da accedere, il browser attiva un opportuno "plug-in" che consente di gestire il contenuto scelto.

## Come identificare i contenuti ....

---

- Atteso che ogni contenuto si riduce ad un "file", sorgono una serie di problemi:
  - quale è la tipologia di contenuto
  - quale è il formato del contenuto
  - quali strumenti esistono per accedere al contenuto e per quali finalità
  - dove si trova il contenuto
  - che regole esistono per poter accedere al contenuto
  - chi mi garantisce sull'autenticità del contenuto
  - come faccio ad avere i diritti di accesso ai contenuti
  - etc. etc.
- L'avvento di internet ha fatto esplodere i temi di cui sopra: la soluzione che è stata trovata è quella di "inventare" un modo per identificare in maniera (quasi) univoca ogni possibile contenuto tramite un **URI (Universal Resource Indicator)**.
- In questo modo è anche possibile indicare agevolmente il riferimento ad un generico contenuto tramite un **URL ( Universal Resource Locator)**.
- La struttura di questi identificativi è molto semplice... anche se spesso criptica !
  - **indicazione di una modalità di accesso** ( es. http, ftp, rtps, etc.)
  - **un indirizzo internet di un "sito" presso cui il contenuto è accessibile**
  - **un indice gerarchico del contenuto all'interno del "sito"**
  - **un indicatore del "formato" del contenuto**
- Qualche esempio :
  - <http://tempid.altervista.org/docs/LERN/SRI.pdf>
  - <http://atom-ecology.russgeorge.net/2016/04/04/homlid-olafsson-rydberg-fusion/SRI.pdf>

- Nel seguito approfondiremo la struttura della rete internet: per ora ci interessa solo vedere la rete internet come un insieme di "luoghi" a cui è possibile "accedere" con delle modalità operative appropriate ovvero con un certo "protocollo"
- Ogni "sito" internet quindi dovrà avere ovviamente un "indirizzo" utilizzabile tramite gli strumenti appropriati per poter accedere al sito stesso.
- L'idea di fondo che sta dietro un **indirizzo** è quella di esprimere in maniera sintetica una serie di **elementi atti ad individuare il sito**:
  - la nazione o la tipologia di organizzazione cui appartiene
  - l'organizzazione o l'individuo cui il sito appartiene
  - l'unità funzionale o logistica all'interno dell'organizzazione
  - il gruppo operativo di appartenenza
  - la macchina fisica che contiene i "files" che sostanziano il contenuto.
- Qualche esempio:
  - [www.repubblica.it](http://www.repubblica.it)
  - [napoli.repubblica.it](http://napoli.repubblica.it)
  - [www.nasa.gov](http://www.nasa.gov)
  - [www.business.alcatel.com](http://www.business.alcatel.com)
  - [www.tg3.rai.it](http://www.tg3.rai.it)
  - [longo.docenti.unina.it](http://longo.docenti.unina.it)
  - [sl7.intesitalia.com](http://sl7.intesitalia.com)
- La **modalità di gestione degli indirizzi** viene definita "**a dominio**": ovvero esiste un concetto di gerarchia all'interno di altre gerarchie: in questo modo diventa facile a colpo d'occhio capire un certo sito che "panni veste" !!!!!

- Ogni sito internet deve avere una sua **struttura interna di organizzazione dei contenuti**: ovvero un vero e proprio indice, con dei capitoli, dei sottoparagrafi, etc.
- Per semplificare l'individuazione di un preciso contenuto all'interno di un sito sarà quindi necessario fornire i nomi dei capitoli, paragrafo, etc... per semplificare si utilizza il concetto di "**percorso**" ("**path**" in inglese).
- Un "percorso" altro non è che una sequenza di nomi separati dal carattere "/"
- Qualche esempio:
  - /dl/rainews/articoli/Intercettazioni-Renzi-Il-governo-non-ci-mettera-mano.html
  - /2016/04/04/homlid-olafsson-rydberg-fusion/index.html
- Il percorso va letto da sinistra a destra nel senso di andare sempre più in dettaglio; il pezzo di percorso più a destra in genere contiene una indicazione del tipo di contenuto a cui il percorso conduce.
- L'**organizzazione dei contenuti** all'interno di un sito segue quindi ancora una volta il concetto di "**gerarchia**" in modo da raggruppare agevolmente i contenuti in base a dei criteri facilmente intuibili dal potenziale utilizzatore del sito.
- L'insieme dei vari elementi che compongono un URL sono quindi tali da individuare in maniera univoca un contenuto e le modalità per accedere allo stesso.

# Protocolli... è qualcosa che si mangia ?????

---

- La parte più criptica di ogni indirizzo di contenuto è la prima parte: il **protocollo**
- Si tratta di un indicativo, in genere un acronimo, che serve ad individuare il modo e quindi gli strumenti da utilizzare per l'accesso al contenuto.
- Qualche esempio:
  - HTTP : Hypertext Transfer Protocol
  - HTTPS : Hypertext Transfer Protocol Secure
  - FTP: File Transfer Protocol
  - SMTP : Simple Mail Transfer Protocol
  - MMS : Microsoft Media Services
  - RTSP : Real Time Streaming Protocol
- A seconda del tipo di protocollo sarà possibile fare certe cose: ad es.
  - RTSP consente di vedere un contenuto video in modalità streaming
  - FTP consente di "scaricare" un file sul proprio computer relativo al contenuto
  - HTTP consente di visualizzare il contenuto all'interno di un opportuno Browser
- Spesso lo strumento base per accedere ad un qualsiasi contenuto è un browser che automaticamente sceglie di usare dei "plug-in" interni a seconda del tipo di protocollo richiesto per accedere ad un certo contenuto.

- La scelta di un formato in cui rendere disponibile un contenuto è un tema chiave: infatti la struttura di un formato:
  - determina quali tipi di azioni si possono fare sul contenuto
  - che livello di qualità si può ottenere
  - che funzioni di protezione del contenuto si possono utilizzare
- Alcuni esempi:
  - **formato pdf** : è un formato di "descrizione di pagina" pensato per l'uso con le stampanti; consente di sfruttare il massimo delle caratteristiche delle stampanti; consente di limitare le azioni possibili ( es. solo display), e l'accesso ai contenuti su base credenziali.
  - **formati jpg, png, gif, etc** : orientati ai contenuti di tipo immagine/foto, non supportano meccanismi di protezione o di limitazione degli accessi, hanno limiti di qualità ( compressione con perdita)
  - **formati mpg, mp3, mp4, mkv, etc.** : orientati ai contenuti audio/video multimediali, non supportano meccanismi di protezione o di limitazione degli accessi, hanno limiti di qualità ( usano compressione con perdita)
  - **formati zip, rar, z, 7z, etc**: orientati alla archiviazione dei dati; supportano meccanismi di protezione tramite credenziali, crittografia, e non hanno limiti di qualità ( usano compressione senza perdita)
  - **formati doc, xls, docx, ppt, pptx, odt, etc**: orientati alla creazione di documenti multimediali, consentono spesso protezione dei contenuti tramite credenziali e limitazione delle funzioni possibili; non hanno limiti di qualità in quanto non usano meccanismi di compressione con perdita.

- **formati html, htm**: orientati alla descrizione di contenuti multimediali, consentono agevoli modifiche dei contenuti, non supportano nativamente meccanismi di protezione del contenuto o di limitazione degli accessi.
- Esistono poi una miriade di formati specifici di particolari applicazioni che sono pensati per usi in connessione con strumenti ( tools) precisi e non per la diffusione generica di un contenuto.
- Un formato molto interessante è l'**html (Hyper Text Markup Language)** : esso utilizza un linguaggio di "**markup**" per individuare e referenziare contenuti internet e per **formattare le modalità di presentazione all'interno di un "browser"**
- E' un linguaggio "**estensibile**" in modo da potersi agevolmente adattare alla rappresentazione di nuovi contenuti o a poter implementare nuove modalità di presentazione dei contenuti
- E' definito da una serie di standards completamente aperti e "vendor independent" e quindi rappresenta una sorta di "lingua franca" per la descrizione dei contenuti.
- E' un linguaggio "testuale" e come tale si presta ad essere trattato anche con tools generici per realizzare modifiche .
- La sintassi del linguaggio si basa sul concetto di "**tag**" per individuare gli elementi di contenuto o gli attributi di presentazione : ad es.
  - **<img> http://www.pippo.com/esempio/immagine.png </img>**
  - **<A http://www.pippo.com/esempio/pagina.html> esempio di pagina </A>**

## A proposito di "browsers" ....

---

- Esistono numerosi Browsers in genere disponibili gratuitamente; ad esempio:
  - Google Chrome: disponibile su moltissime piattaforme e allineato agli standards
  - Firefox: strumento "open source" disponibile su moltissime piattaforme e allineato agli standards.
  - Internet Explorer: strumento proprietario Microsoft usato dalla stessa come principale veicolo di protezione del proprio mercato contro la concorrenza.
- Esistono poi una serie di numerosi derivati che, sfruttando le piattaforme aperte fornite da Firefox e Chrome, hanno creato dei browser alternativi privilegiando alcuni aspetti quali ad es. la velocità, la richiesta di risorse, funzionalità aggiuntive.
- Un discorso a parte è quello dei "**plug-in**": in genere estendono le funzionalità di un browser e forniscono funzionalità aggiuntive; alcuni esempi:
  - **Java virtual machine**: consente di eseguire applicazioni scritte nel linguaggio java; implementa un ambiente esecutivo molto sofisticato garantendo l'indipendenza dalla piattaforma HW/SW
  - **Flash extensions**: consentono di eseguire contenuti che utilizzano il linguaggio flash; la soluzione è stata oggetto di una enormità di bachi SW e il maggior veicolo di propagazione di virus, per cui è in fase di eliminazione dal mercato
  - **plugins multimediali**: sono orientati alla decodifica di vari formati: ad es: silverlight, quicktime, etc.
  - **estensioni funzionali**: ad es. bookmark manager, giochi, temi alternativi, etc.

## A proposito di "pacchetti office"....

---

- Esistono numerosi pacchetti per il trattamento dei testi e dei documenti: alcuni esempi significativi:
  - **Microsoft Office**: sicuramente il pacchetto più conosciuto; è un ambiente proprietario dal costo significativo e uno dei principali veicoli di protezione del mercato Microsoft. Ha dei costi esorbitanti per l'utente non professionale e richiede molte risorse di calcolo ( ovvero HW abbastanza potente)
  - **Open Office / Libre Office**: ambiente aperto e gratuito allineato agli standard e compatibile con i formati proprietari Microsoft; rappresenta il maggior competitor del prodotto Microsoft Office. Ha il vantaggio di girare su numerose piattaforme HW/SW. Supporta i formati aperti. E' abbastanza "pesante" come richiesta di risorse di calcolo.
  - **SoftMaker** : è un ambiente proprietario alternativo a Microsoft Office ma strettamente compatibile con esso; ha il vantaggio di estendere il supporto ai formati aperti ed è molto più "leggero" del prodotto Microsoft. Ha un costo decisamente basso e alla portata di tutti.
- Un tema che si pone usando questi pacchetti è quali formati usare per i documenti:
  - sicuramente l'uso di formati aperti dovrebbe essere preferito.....
  - ma purtroppo il dominio Microsoft degli anni pregressi è un dato di fatto....
  - nel frattempo che i nuovi formati aperti prendano piede conviene usare i formati Microsoft di NON ultima versione: ad es.
    - formato .doc o .docx versione antecedente al 2010
    - formato .xls o .xlsx versione antecedente al 2010
    - formato .ppt o .pptx versione antecedente al 2010

---

Grazie per l'attenzione !