

Scegliere una licenza per i dati aperti

Formez_{PA}



Questo materiale didattico è stato realizzato da Formez PA nel *Progetto PerformancePA*, Ambito A Linea 1, in convenzione con il Dipartimento della Funzione Pubblica, organismo intermedio del Programma Operativo Nazionale Governance e Azioni di Sistema (PON GAS), Asse E Capacità istituzionale. Il PON GAS è cofinanziato dal Fondo Sociale Europeo ed è a titolarità del Ministero del Lavoro e delle Politiche Sociali.

L'opera è distribuita con Licenza [Creative Commons Attribuzione - Condividi allo stesso modo 4.0 Internazionale](https://creativecommons.org/licenses/by-sa/4.0/).



Autore: Simone Aliprandi

Creatore: Formez PA

Diritti: Dipartimento della Funzione Pubblica

Data: Ottobre 2015

Scegliere una licenza per i dati aperti

Quando si parla di open data non deve sfuggire che il fatto che i dati siano aperti non implica che non debbano essere coperti da licenza d'uso una licenza d'uso che determini le modalità con le quali l'amministrazione possa utilizzarli citando la fonte piuttosto che non citandola potendoli modificare piuttosto che non potendolo fare potendoli utilizzare o distribuire nell'ottica di consentirne un uso un riuso commerciale sono tutte caratteristiche che devono essere determinate e governate attraverso delle licenza d'uso. Licenza d'uso specifica, certo, pensate per favorire e supportare lo sviluppo di un ecosistema fatto di informazione libera e aperta stiamo parlando delle licenze "creative commons" che sono ormai uno standard in quest'ambito. Quali sono le caratteristiche di queste licenze, quali punti di forza, quali le cose da sapere per la PA che intende utilizzarle? Ne parliamo con l'avvocato Simone Aliprandi, uno dei principali esperti italiani del settore.

Perché è importante parlare di licenze per gli open data? Quali sono i fattori da considerare?

È importante parlare di licenze per l'open data perché un dato non licenziato, cioè senza una licenza open, non è un dato open, è la licenza in sé che rende il contenuto o il dato che stiamo pubblicando open, cioè libero dai normali tradizionali vincoli che il copyright normalmente applica sulle opere creative e sulle banche dati. Quindi, la questione da considerare è questa, benché si creda spesso il contrario per disinformazione di leggenda metropolitana cioè che qualsiasi cosa che si trova online è sostanzialmente libera perché se è stata messa lì allora, qualcuno gradiva che se ne facesse libero uso, in realtà non è così, bisogna considerare che, qualsiasi contenuto che non avremo creato noi stessi, è da considerare come un contenuto "closed by default" ed è quindi chiuso per definizione perché c'è qualcuno che ha creato quel contenuto e che quindi detiene dei diritti d'autore su questo contenuto.

Lo stesso discorso vale con alcune precisazioni e differenze che non possiamo qua approfondire anche per le banche dati quindi se non troviamo una licenza che ci dice che possiamo liberamente utilizzare una banca dati e un data set non possiamo legittimamente considerarlo come un data set open. Ecco perché è molto importante spingere affinché si alzi la consapevolezza da parte titolari dei diritti quindi in questo caso le pubbliche

amministrazioni, sul tema delle licenze: applicare correttamente le licenze e scegliere le licenze corrette aiuta a rendere open i dati.

Che cos'è in breve il principio dell'open by default applicato ai dati e documenti della Pubblica Amministrazione?

Il cosiddetto principio open by default è una delle più grosse innovazioni in questo campo e nella legislazione italiana perché, benché dal punto di vista del diritto d'autore del puro diritto d'autore, quindi della Legge 633 del 1941 non sia cambiato nulla, il nostro legislatore, nel 2012 ha inserito questo nuovo principio all'interno del codice dell'amministrazione digitale, il così detto CAD e cosa fa, a cosa serve questo principio? Serve a sfruttare, l'inerzia delle Pubbliche Amministrazioni nell'utilizzo delle licenze open è interessante come cosa, le licenze open esistono già da 15 anni circa.

La spinta verso l'open data esiste già dal 2008/09, il dibattito scientifico si è aperto già da un po' di anni eppure si è registrato che le pubbliche amministrazioni per una questione di inerzia di poca voglia di fare questo passo innovativo, non utilizzavano le licenze open allora il legislatore, scaltramente ha pensato di invertire il meccanismo e dire che, nel caso in cui, sia una pubblica amministrazione a pubblicare un dataset o un documento perché si parla di dati e documenti e non applichi nessun tipo di licenza non mette nessun tipo di licenza proprietaria allora automaticamente il dato e il contenuto deve essere considerato open ai sensi della definizione che viene fornita da articolo 68 del CAD che è un altro articolo a cui si fa riferimento.

State attenti, nelle pubbliche amministrazioni in cui non viene indicata una licenza, in realtà il legislatore sta dicendo con una norma che è l'articolo 52 del CAD a tutti gli utilizzatori di questo data set che lo possono utilizzare liberamente perché non c'è una licenza e quindi non c'è nessun tipo di condizione o di restrizione allora lo si considera open by default questo è un po' lo scopo. Io consiglio sempre e comunque alle pubbliche amministrazioni di utilizzare le licenze, perché avere una licenza è sempre e comunque un atto di maggiore chiarezza agli occhi dell'utente.

Partendo proprio dall'importanza di utilizzare e scegliere una licenza come atto di maggiore chiarezza nei confronti dell'utente, cosa sono e come funzionano le licenze Creative Commons e come scegliere la licenza più adatta alle esigenze della pubblica amministrazione?

Le licenze Creative Commons altro non sono che le più diffuse, le più utilizzate licenze "open content" cioè per contenuti aperti per contenuti liberi; nascono nei primi anni 2000 sull'onda della diffusione di internet come fenomeno di massa, il progetto Creative Commons voleva mettere a disposizione di tutti gli utenti della rete che avessero attività creative di qualsiasi tipo degli strumenti legali per poter rilasciare le loro opere in un regime più libero rispetto a quello che era il classico copyright cioè tutti i diritti riservati, quel "all rights reserved". Creative Commons ha un moto che invece fa un po' il verso a questo modello che è "some rights reserved" cioè alcuni diritti riservati. Basti pensare, giusto per dire un nome molto il richiamo, Wikipedia è quel tipo di opera liberamente condivisibile e anche modificare la parte di tutti proprio perché utilizza una licenza creative commons e nello specifico la By-SA la Attribution Share Alike.

Come scelta della licenza più opportuna per l'open data soprattutto nell'ambito di dati rilasciati dalla Pubblica Amministrazione dal mio punto di vista credo che non ci siano molti dubbi anche perché ci sono indicazioni chiare da parte del legislatore italiano cioè, io sceglierei tra le 6 licenze creative commons le due più semplici la più semplice forse è la scelta migliore, è la BY, quella che richiede solamente che venga riconosciuta correttamente la paternità dell'opera e quindi venga richiamata la Pubblica Amministrazione o comunque l'ente che ha rilasciato il data set come fonte originaria, ma nessun altro tipo di restrizione, ad esempio poco opportune sono le restrizioni come la "no deriver" indicare come clausola la non possibilità di fare opere derivate cioè di non modificare il data set perché un data set nasce proprio per essere riutilizzato reimpastato e mischiato insieme ad altri dati e quindi non avrebbe molto senso e anche la clausola "non commercial" cioè quella che vieta la realizzazione di opere commerciali l'uso commerciale da parte degli utenti di questo data set.

Ovviamente anche in questo caso la Pubblica Amministrazione non è in un'ottica di mercato, non è sullo stesso piano di aziende che devono fare business e magari con i data set e quindi vedo poco opportuna l'applicazione di clausole non commercial da parte di enti pubblici, se poi consideriamo che in alcuni paesi del mondo importanti come gli Stati Uniti i dati pubblici di origine governativa sono di per sé fuori dal copy right cioè di pubblico dominio una licenza che ci avvicini il più possibile a un pubblico dominio rende la vita più facile a tutti anche alla Pubblica Amministrazione che deve rilasciare il dato perché effettivamente non le crea

nessun tipo di vantaggio economico quello di applicare clausole limitative ma solamente deve avere più scocciature da gestire dal punto di vista legale.

La versione 4.0 delle licenze Creative Commons ha introdotto una serie di importanti novità di cosa si tratta e quali vantaggi offre?

Abbiamo detto che le licenze Creative Commons sono sei, si tratta infatti di una gamma di 6 licenze, 6 opzioni dalla più restrittiva alla più libera, tra le quali appunto il creatore, l'autore di un'opera può attingere, può scegliere, per rilasciare liberamente la sua opera. Non solo, le licenze Creative Commons dalla loro nascita, quindi dai primi anni 2000 ad oggi, hanno subito un'evoluzione, hanno avuto delle generazioni siamo arrivati alla 4.0, quando si dice licenza 4.0 è perché siamo arrivati alla quarta generazione di licenze Creative Commons e questo perché, perché le licenze subiscono ovviamente variazioni e migliorie a seconda delle segnalazioni, delle esigenze, anche dal cambiamento del mondo della comunicazione, è ovvio che internet nel 2001 era diversa dall'internet di oggi.

La più grossa novità che le 4.0, la generazione 4.0 pone, è relativa al famoso diritto sui generis. Il diritto "sui generis" è uno strano diritto che si applica alle banche dati, ai data base, e che esiste solamente nel contesto europeo, dell'unione europea, perché è un diritto che ha le sue radici in una direttiva europea del '96 e che abbiamo solo noi europei; il motivo per cui le Creative Commons fino alla 4.0 non avevano questo diritto menzionato è molto semplice, essendo licenze che provengono dall'ordinamento americano, gli americani non conoscono questo diritto e non avevano trovato opportuno, non avevano pensato di occuparsene all'interno delle licenze e quando le licenze hanno avuto successo e hanno cominciato a diffondersi anche al di fuori dei confini degli Stati Uniti, qualcuno ha fatto notare che era necessario considerare anche questo aspetto perché si rischiava di non licenziare correttamente i dataset, cioè applicare una licenza Creative Commons a una banca dati senza le corrette e dovute clausole relative al diritto sui generis, creava un problema, lasciava un buco, allora Creative Commons dopo una serie di studi e approfondimenti, ha rilasciato le 4.0 che trattano invece correttamente anche il diritto "sui generis" quindi, pur provenendo da un ordinamento americano possono essere applicate anche in ambito europeo senza problemi e finalmente riescono a licenziare correttamente anche questo diritto.

L'altra novità relativa alle 4.0 è legata al discorso del così detto "porting": fino alla 3.0 Creative Commons si è occupata di fare un lavoro, un lavoro veramente pesante di adattamento delle licenze ai vari contesti dei vari ordinamenti giuridici cioè, in ogni stato, in ogni ordinamento in cui c'era un progetto Creative Commons attivo, i giuristi Creative Commons si

occupavano della traduzione in lingua ma anche dell'adattamento all'ordinamento locale cosa che però creava, oltre a un dispendio di risorse, anche problemi di differenza di interpretazione tra una versione, ad esempio una versione giapponese, rispetto alla versione italiana, rispetto alla versione spagnola. Creative Commons ha fatto marcia indietro su questa scelta e ha deciso che le 4.0 vengono rilasciate in un'unica versione che venne chiamata "international" che poi comunque ha delle traduzioni, traduzioni meramente linguistiche, per essere più comprensibile a persone che non parlano inglese, però la licenza è identica, uguale a se stessa in tutte le nazioni quindi non c'è più il problema, il dubbio che si possono creare differenze e discrasie nell'interpretazione.

Che cosa è la Creative Commons Zero, e che cosa implica rilasciare dati in pubblico dominio?

CC Zero è un'ulteriore strumento messo a disposizione da Creative Commons, abbiamo detto che ci sono 6 licenze, ma questo è appunto uno strumento un po' diverso perché la licenza tecnicamente è , come dice la parola stessa, dal latino "licere", una autorizzazione cioè, il titolare dei diritti ha comunque i diritti, dichiara di averli e di volerli ancora esercitare e ci sta dando delle autorizzazioni a volte molto libere, tante autorizzazioni, ma che sono sempre condizionate al rispetto del diritto d'autore e quindi al rispetto delle condizioni che il titolare dei diritti sta dando attraverso la licenza.

La CC Zero, è detta in inglese un "atto di rinuncia", sarebbe una sorta di liberatoria ma qualcosa di ancora più potente di quella che noi in Europa chiamiamo liberatoria dei diritti perché è un atto di rinuncia totale definitivo per l'esercizio del diritto d'autore è come paradossalmente se l'autore dichiarasse di essere già morto da più di 70 anni, passatemi la battuta per rendere l'idea. In questo modo l'utente apponendo questa dichiarazione alla sua opera, se voi andate a leggere il testo lo capite sta dichiarando di non volere più da quel momento esercitare nessun diritto d'autore su quest'opera quindi è un passo ancora più libero ancora più radicale rispetto all'utilizzo di una delle sei licenze Creative Commons.

È come, io lo chiamo rilascio in pubblico dominio in una sorta di pubblico dominio artificiale perché il vero pubblico dominio arriva solo quando sono scaduti tutti i diritti e quindi sono passati i famosi 70 anni dalla morte dell'autore oppure 15 anni dalla pubblicazione di un dataset per quanto riguarda il diritto sui generis; in questo caso l'utente è ancora in vita, c'è ancora comunque un ente che detiene un diritto d'autore o sui generis sul data set però, per sua scelta, decide di liberarsene prima, di rilasciare liberamente l'opera questo ha una potenzialità mostruosa perché da quel momento chiunque può farne qualsiasi uso con

massima libertà e senza doversi preoccupare di dover nemmeno rendere attribuzione al titolare dei diritti. È una scelta che alcune pubbliche amministrazioni hanno fatto è una scelta che non fa altro che ripercorrere quello che in alcuni ordinamenti giuridici già avviene per legge che i dati e i contenuti delle pubbliche amministrazioni in alcuni casi sono già di per sé in pubblico dominio è un po' come quello che avviene da noi col famoso open data by default però appunto lì è la scelta del titolare dei diritti di rinunciare ai diritti che fa scattare il tutto. È un'altra opzione e consideratela e prendetela in esame.

Esporre dati in formato aperto: qualche suggerimento per la pagina "note legali" dei siti web pubblici che li espongono?

Come suggerimento per correttamente applicare le licenze e quindi correttamente esporre un data set sul web in modalità open direi che c'è un unico e molto semplice criterio, che è quello della estrema chiarezza. Noi dobbiamo mettere il più possibile nelle condizioni il nostro lettore, l'utente che arriva sul nostro sito web di capire quali sono i termini d'uso del nostro data set della nostra opera, del nostro documento questo perché un documento non licenziato correttamente, cioè non licenziato chiaramente dev'essere sempre trattato come un documento sotto copyright perché qualcuno lo ha creato e qualcuno potrà sempre vantare un diritto d'autore o un diritto connesso o un diritto sui generis. Se vogliamo renderlo veramente libero dobbiamo essere molto chiari e dire "signori questa è la licenza applicata e questa è una licenza open perché ha queste determinate caratteristiche".

Se poi andiamo a leggere le linee guida di Creative Commons, Creative Commons fa un passo in più cioè ci dà delle indicazioni molto chiare, molto tecniche, per applicare correttamente le loro licenze. Basta andare sul sito Creative Commons, soprattutto quello in versione originale, .org americano, c'è un widget che ci guida nella scelta della licenza ma non solo ci guida a scegliere le sei licenze alla fine della nostra scelta, quando abbiamo capito qual è la licenza più opportuna per i nostri scopi e per i nostri usi il widget ci fornisce anche una stringa di codice html che noi possiamo prelevare ed "embeddare" cioè nascondere infilare all'interno del codice html della nostra pagina web.

Questo cosa fa, non solo fa comparire un disclaimer nel testo della pagina web con scritto "... quest'opera è licenziata sotto licenza Creative Commons ..." e inserisce il link come indicato dalle linee guida ma essendo fornito anche di metadati, permette ai motori di ricerca di trovare la nostra opera, di trovare il data set di individuare che quel data set è licenziato con quello specifico tipo di licenza, questa è una cosa molto importante molto potente perché chiunque va su Google, pensate ad esempio alla ricerca avanzata di Google se voi andate in

fondo, nella ricerca avanzata c'è una voce che si chiama "diritti di utilizzo" che ci fa capire come Google sia attrezzato per riconoscere questi metadati e quindi poter dire all'utente sta cercando un data set oppure un brano musicale oppure un libro che quell'opera ha quelle caratteristiche di licenza ha una licenza di questo tipo, permette ad esempio anche l'uso commerciale permette ad esempio anche la modifica dell'opera, questa cosa è molto potente, soprattutto in un'ottica di web semantico, e quindi io consiglio il più possibile di seguire queste indicazioni.

Al di là di questo quello che è richiesto è di essere chiari e cioè fare sì che nel disclaimer ci sia chiaramente chi è il titolare dei diritti quindi la pubblica amministrazione o il soggetto che espone i dati in modo molto chiaro, l'anno in cui parte il copyright quindi 2004, 05, 2012, 2015 a seconda di quando è stato pubblicato il contenuto e chiaramente l'indicazione della licenza scelta; dire Creative Commons non basta perché sono 6 quindi devo dire quale di queste 6, state attenti alle abbreviazioni che a volte possono confondere le idee, vi ricordo soprattutto, molto importante, l'indicazione del link della licenza mettendo il link siamo sicuri che l'utente non può sbagliare perché basta fare click e viene mandato alla pagina di Creative Commons in cui c'è il testo della licenza integrale, quindi l'abbiamo messo nelle piene condizioni di informarsi su quali sono i termini d'uso della mia opera del mio data set questo è esattamente ciò che è richiesto e consigliato.