

**Linee guida provinciali
per la valorizzazione del patrimonio informativo pubblico**

Dicembre 2015

Le presenti linee guida - condivise con gli enti che già partecipano alla piattaforma dati.trentino.it. - rappresentano lo stato dell'arte della stessa alla data dell'approvazione del documento, sono rivolte prevalentemente a chi opera nel servizio pubblico e sono finalizzate a rendere più efficiente e di qualità il processo di produzione dei dati della pubblica amministrazione, fornendo le conoscenze di base su temi attuali del dibattito sul tema al fine di permettere la governance del processo di messa a disposizione, da parte dei titolari dei dati, delle informazioni di propria competenza in modalità open data. E' un documento di consultazione e di lavoro, che affronta la tematica dai diversi punti di vista: legale, procedurale, organizzativo, tecnologico, indicando i riferimenti per le attività specifiche degli operatori coinvolti nel processo di apertura dei dati.

Il documento è stato condiviso con la comunità open data del Trentino nell'autunno 2015 (<http://www.innovazione.provincia.tn.it/eventi/pagina69.html>).

Sotto il profilo organizzativo, con la deliberazione n. 1509 del 2015 la Giunta provinciale ha assegnato la competenza di coordinamento in materia di Open Data al Servizio Supporto alla Direzione generale e ICT della Provincia autonoma di Trento, che ne coordina le azioni nell'ambito delle politiche definite dall'Agenda digitale.

Questo documento si basa sul lavoro svolto dal gruppo di progetto *Open Data in Trentino* nei tre anni di attività e, a partire dalla primavera 2015, sul lavoro redazionale di Francesca Gleria, Eleonora Bassi, Francesca De Chiara, Marco Combetto, Stefano Leucci, Maurizio Napolitano, Luca Paolazzi, Roberto Cibir, David Leoni.

Indice generale

0. Executive summary.....	5
1. Introduzione.....	7
2. Finalità, destinatari e alcune definizioni utili.....	10
2.1 Finalità del documento.....	10
2.1.1 Perché un punto di accesso unico territoriale.....	11
2.2 Destinatari del documento.....	13
2.3 Definizioni.....	13
3. Il contesto giuridico.....	17
3.1. Quadro di sintesi normativa.....	18
3.1.1 Normativa sul riutilizzo delle informazioni del settore pubblico.....	18
3.1.2 Normative su domini specifici.....	20
3.1.3 Altre normative rilevanti in materia di open data.....	21
3.1.4 Normative di contesto.....	24
3.3 Modalità di individuazione dei dati oggetto di riutilizzo.....	28
3.4 Licenze associate ai dati.....	30
3.5 Le richieste di riutilizzo.....	32
3.6 Aspetti relativi al costo dei dati.....	33
3.7 Accordi in esclusiva e divieto di discriminazione.....	34
4. Modello operativo per la produzione e il riutilizzo dei dati.....	35
4.1 Cosa rende di qualità un dataset.....	36
4.1.1 Il catalogo dati.trentino.it e i certificati di qualità dei dati.....	38
4.2 Modello operativo per la piattaforma dati.trentino.it.....	40
4.2.1 Censimento delle basi dati.....	42
4.2.2 Analisi Giuridica.....	43
4.2.3 Analisi qualità, valutazioni costi/benefici, azioni di bonifica.....	43
4.2.4 Strutturazione del dato/bonifica.....	45
4.2.5 Politiche di accesso e licenze d'uso.....	46
4.2.6 Pubblicazione in open data.....	46
4.2.7 Metadattazione.....	47
4.2.8 Gestione e mantenimento del dato.....	47
4.3 Cosa vuol dire riutilizzo dei dati.....	48
4.4 Indicazioni generali per la realizzazione dei sistemi informativi.....	49
4.4.1 Censimento delle basi dati informative e dei servizi informativi degli enti.....	50
4.4.2 Analisi e progettazione dei sistemi informativi.....	51
4.4.3 Georeferenziazione dei dati.....	52
4.5 Indicazioni operative per i dati di tipo aperto nei capitolati di gara.....	54

4.6 Modello organizzativo.....	56
4.6.1 Team Open Data.....	56
4.6.2 Referente Open Data nei dipartimenti.....	57
4.6.3 Titolari dei dati.....	57
4.6.4 Referenti tecnici.....	58
4.6.5 Raccordo con le strutture provinciali, con il sistema trentino della pubblica amministrazione digitale e con la Community Open Data.....	58
5. Modelli e standard tecnici di riferimento.....	60
5.1 Standard tecnici e modello “5 stars” per produrre e pubblicare dati aperti.....	60
5.2 Formati aperti.....	63
5.2.1 Formati di tipo aperto per documenti.....	63
5.2.2 Formati di tipo aperto per dati “strutturati generici”.....	64
5.2.3 Formati di tipo aperto per dati geografici.....	66
5.2.4 Formati di tipo proprietario “aperti de facto”.....	68
5.3 Modello per la metadattazione.....	69
5.3.1 Modello di riferimento.....	69
5.3.2 Verso uno standard unico per la metadattazione: DCAT-AP.....	71
5.3.3 Elenco metadati dati.trentino.it.....	71
5.4 Altri standard, vocabolari e ontologie.....	74
5.5 Alcuni standard di riferimento su specifici ambiti settoriali.....	76
5.6 Considerazioni sui Linked Data.....	77
5.7 Standard nella struttura dei contenuti della comunicazione web del Sistema Trentino.....	78

0. Executive summary

I dati delle pubbliche amministrazioni sono **un patrimonio della collettività** costruito dal servizio pubblico. Tale patrimonio di informazioni ha un alto valore sociale, poiché dal riutilizzo di questi dati si abilitano prodotti e servizi digitali in grado di migliorare sia la **qualità e l'innovazione dei servizi offerti ai cittadini** che l'**interoperabilità interna** fra le strutture di una stessa amministrazione e tra amministrazioni diverse. L'aumento della standardizzazione, e quindi della qualità dei dati, può portare alla creazione di circoli virtuosi fra gli attori del sistema territoriale (servizio pubblico, imprese, ricerca e cittadini) avviando processi di innovazione sociale, nel solco promosso da Europa 2020¹ al fine di abilitare la crescita, ora anche in vista di quanto necessario allo sviluppo di un unico mercato digitale europeo².

È compito del servizio pubblico, nel suo insieme, non sprecare questo patrimonio, averne cura e valorizzarlo creando tutte le condizioni possibili per gestirlo in modo efficiente a beneficio della comunità.

Il valore del patrimonio informativo pubblico si genera **nel momento stesso in cui il dato viene creato**, quando il processo amministrativo viene reso operativo in un flusso di informazioni governate da strumenti informatici. Il processo avrà effetti positivi quanto più gli attori coinvolti e il territorio nel suo insieme saranno capaci di sviluppare **cultura del dato**, cioè quell'insieme di conoscenze e pratiche che producono dati secondo standard tecnici e ricchi di documentazione, in un contesto di sviluppo che supera i confini locali e si rivolge a mercati e comunità nazionali ed europee. In breve, occorre garantire processi di qualità per avere **dati di alta qualità**, gli unici capaci di innescare modelli di riutilizzo compatibili con la domanda di mercato.

Lo sviluppo tecnologico derivato dalla sempre maggiore capacità di calcolo e dalla disponibilità di tecnologie mobili ha cambiato le aspettative dei cittadini circa la fruibilità delle informazioni del servizio pubblico e ha aperto spazi nuovi per lo sviluppo di servizi prima sconosciuti.

Per realizzare questi obiettivi, in linea con il quadro nazionale ed europeo di sviluppo delle pratiche degli open data, è importante **definire dei processi** che permettano di garantire **legittimità, tempestività, sostenibilità e qualità nei dati che si offrono al riutilizzo**. Vanno quindi curati vocabolari, metadati e classificazioni, individuati gli standard nazionali ed europei, definiti i criteri attraverso cui guidare le scelte dei titolari dei dati **tenendo bilanciati sia il piano di legittimità sia l'opportunità di crescita economica** e di incremento dell'efficacia dell'azione amministrativa.

Questo documento è **modellato e prende spunto dalle "Linee guida nazionali per la valorizzazione del patrimonio informativo pubblico (anno 2014)"³**, aggiornato con le attività dei gruppi di lavoro AgID tuttora in corso. Un raccordo produttivo fra i diversi livelli di governo su questo profondo cambiamento organizzativo in atto nel servizio pubblico è assolutamente necessario, pertanto del documento nazionale vengono recepite le linee guida principali e sviluppate le specifiche normative,

1 http://ec.europa.eu/europe2020/index_it.htm

2 <http://ec.europa.eu/digital-agenda/en/digital-single-market>

3 http://www.agid.gov.it/sites/default/files/linee_guida/patrimoniopubblicolg2014_v0.7finale.pdf

le tecniche e le procedure organizzative **in modo adatto al contesto del sistema trentino**, garantendo uno **sguardo puntato alle pratiche oggi avviate dall'Unione Europea**⁴.

Lo scopo è soprattutto quello di offrire **uno strumento operativo per costruire e rafforzare le comunità di pratiche della pubblica amministrazione sulla cultura del dato raccolta attorno alla piattaforma dati.trentino.it**.

Il documento è organizzato in cinque capitoli.

Dopo un'introduzione che definisce la cornice generale delle attività, nel secondo capitolo si inquadrano finalità e destinatari e si fornisce una tabella per orientare le pubbliche amministrazioni su un uso corretto delle definizioni; il terzo capitolo affronta le tematiche giuridiche; il quarto capitolo affronta temi organizzativi e di processo con attenzione anche ad alcuni aspetti di riutilizzo e alle modalità di adeguamento dei sistemi informativi ai principi di riutilizzo e standardizzazione dei dati, anche tramite indicazioni da inserire nei capitolati di gara. Nel quinto ed ultimo capitolo si definiscono questioni tecniche calate nella realtà del "fare open data".

4 <http://www.europeandataportal.eu/>

1. Introduzione

Con questo documento si **aggiornano le precedenti linee guida in materia, approvate con deliberazione della Giunta provinciale n. 2858 del 27 dicembre 2012**⁵, facendo tesoro di quanto accaduto dal 2012 ad oggi.

A livello europeo nel 2013 è stata aggiornata la direttiva PSI 2003/98/CE del 17 Novembre 2003⁶. Il tema degli open data, inoltre, è stato operativamente sviluppato nei programmi Europa2020⁷ sia come pilastro per il mercato unico digitale (azione 3), sia come tema trasversale circa la necessità di una maggior standardizzazione ed interoperabilità (azione 26), sia come strumento di base per uno sviluppo dell'ICT teso a migliorare i servizi (azione 89). Infine, il terzo pilastro della strategia per la creazione del mercato unico digitale europeo⁸ colloca il tema tra le azioni necessarie per la realizzazione di un sistema di eGovernment abilitante un mercato di servizi innovativi⁹ (azione 16). Infine, in occasione dell'European Data Forum 2015, l'Unione Europea ha lanciato l'European Data Portal come supporto operativo ad integrare e standardizzare i portali open data europei e che al suo lancio ha riunito in un accesso unico multilinguistico più di 250.000 data set provenienti dai portali dei Paesi membri.

A livello nazionale il gruppo di lavoro open data dell'Agenzia per l'Italia Digitale¹⁰ ha definito le attuali "Linee guida nazionali per la valorizzazione del patrimonio informativo pubblico (anno 2014)" e "l'Agenda nazionale per la valorizzazione del patrimonio informativo pubblico (anno 2014)"¹¹: tali documenti offrono oggi alla comunità open data obiettivi definiti, un percorso per raggiungerli, gli elenchi di basi di dati di interesse nazionale, strumenti per definire protocolli di interscambio di dati. Vengono inoltre riordinate tematiche relative a modalità organizzative, metadatezione, formati, qualità del dato, ontologie e processi. Questi documenti sono un riferimento istituzionale attorno al quale strutturare un dialogo fattivo nella comunità open data che mette assieme i diversi livelli di governance nazionale, regionale e degli enti locali. In questo modo si creano strumenti per valorizzare il patrimonio informativo del sistema Paese in modo standardizzato e coordinato.

Recentemente è stato anche rivisto il portale nazionale dati.gov.it¹² quale accesso unico ai dati aperti del sistema pubblico nazionale. Nel mese di marzo 2015 la Presidenza del Consiglio dei Ministri ha pubblicato il documento "La strategia per la crescita digitale"¹³ che fra le tante azioni individua alcuni passaggi critici in merito agli open data, per inserire queste attività in un quadro più coordinato e

5 <http://www.innovazione.provincia.tn.it/notizie/pagina482.html>

6 <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2003:345:0090:0096:IT:PDF>

7 http://ec.europa.eu/europe2020/index_it.htm

8 http://europa.eu/rapid/press-release_IP-15-4919_en.htm

9 Utile citare anche gli sviluppi del gruppo EIRA (European interoperability reference architecture) finalizzato a sviluppare un sistema di architettura delle informazioni in grado di sviluppare un modello che renda i sistemi di eGovernment interoperabili su scala europea. <https://joinup.ec.europa.eu/asset/eia/description>

10 <http://www.agid.gov.it/>

11 <http://www.agid.gov.it/agenda-digitale/open-data/dati-pubblici-condivisione>

12 <http://www.dati.gov.it/>

13 http://www.governo.it/GovernoInforma/documenti/piano_crescita_digitale.pdf

governato di azioni a livello di sistema Paese. Infine, nel mese di agosto 2015 è stata approvata la legge delega n. 124 del 7 agosto 2015 di riorganizzazione della pubblica amministrazione¹⁴ che impegna lo Stato a rivedere il CAD - Codice dell'amministrazione digitale - alla luce dei principi della Carta della cittadinanza digitale che, all'art. 1, sottolinea come obiettivo delle azioni di riorganizzazione dell'amministrazione pubblica la necessità di "garantire ai cittadini e alle imprese, anche attraverso l'utilizzo delle tecnologie dell'informazione e della comunicazione, il diritto di accedere a tutti i dati, i documenti ed i servizi di loro interesse in modalità digitale".

Sono stati compiuti inoltre importanti interventi di chiarimento in merito agli aspetti di **legittimità dei processi di apertura dei dati** da parte del Garante per la protezione dei dati personali. Del 2014 è infatti la pubblicazione del documento "La trasparenza sui siti web della PA"¹⁵. Altre misure di supporto relative ad aspetti di trasparenza finalizzati a garantire le azioni preventive di attività di corruzione sono state proposte dall'Autorità Nazionale AntiCorruzione (ANAC)¹⁶, che a partire dal 2013 ha definito azioni concrete e tassative di pubblicazione di informazioni in formato aperto nella specifica sezione "Amministrazione trasparente" di ogni portale della pubblica amministrazione e dei suoi enti.

In questi anni il **sistema Trentino** ha fatto molti passi concreti sul tema, in attuazione delle azioni previste nel Piano di miglioramento 2012 - 2016. L'obiettivo si è focalizzato sulle azioni di avvio e accompagnamento del processo di cambiamento legato alla diffusione delle nuove tecnologie mobili e alle modifiche in atto nei mercati ICT, per effetto dello sviluppo degli open data e dei big data, che portasse a valorizzare i dati delle pubbliche amministrazioni come infrastruttura abilitante l'innovazione.

Nel 2013 è stata aperta la piattaforma dati.trentino.it che oggi è il **catalogo da cui accedere in modo coordinato ai 1.500 dataset pubblicati fino ad oggi dal sistema trentino**. Ad oggi sono 32 le strutture provinciali che pubblicano dati nel catalogo, assieme ad enti strumentali come Informatica Trentina Spa e Trentino Network; la piattaforma è **costruita in modo federato** e si collega, ad oggi, anche ai dati dei Comuni di Storo, Campodenno, Trento e Rovereto, della Regione Trentino Alto Adige/Südtirol, dell'Azienda Provinciale per i Servizi Sanitari e di quattro APT locali. Si è inoltre proceduto ad un **allineamento tecnologico tra il portale istituzionale provinciale e ComunWeb**¹⁷, la piattaforma che il Consorzio dei Comuni offre ai suoi consorziati e che ha incorporato nei suoi sistemi una particolare **attenzione al riutilizzo dei dati che transitano sui portali istituzionali secondo i principi dell'open data**. Alcune strutture provinciali hanno inoltre affrontato e sperimentato interessanti progetti verticali, primo fra tutti l'Istituto di statistica della Provincia di Trento (ISPAT) che ha in corso di conclusione un importante progetto (Lod4Stat) di apertura dei suoi dati in modalità linked¹⁸. Da citare poi il Servizio Turismo, che ha collaborato con i suoi dati alla

14 http://www.gazzettaufficiale.it/atto/stampa/serie_generale/originario nelle province autonome di Trento e Bolzano la legge è applicabile nei vincoli dei rispettivi statuti che stabiliscono competenza primaria nella gestione del personale, ciò non toglie che il riferimento ai principi veicolati dalla legge delega indicano una tendenza in atto che questi territori potranno realizzare con modalità speciali ma all'interno di un flusso di trasformazione la cui direzione è delineata.

15 <http://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/3488002>

16 <http://www.anticorruzione.it/portal/public/classic/AmministrazioneTrasparente>

17 <http://www.comunitrentini.it/Aree/Innovazione>

18 http://www.innovazione.provincia.tn.it/azioni/competitivita/-progetti_ricerca_innovazione/pagina104.html

sperimentazione di una piattaforma per i linked data nell'ambito del progetto europeo Fusepool F3 assieme alla regione Toscana, nonché il Servizio Trasporti pubblici che ha reso disponibili i suoi dati e il suo network per lo sviluppo di un sistema di bigliettazione su mobile capace di integrare le informazioni di più modalità di trasporto. Infine, un'ultima azione avviata più recentemente ma di particolare importanza è la **collaborazione con la Provincia autonoma di Bolzano, grazie anche alla "funzione ponte" scaturita dalla collaborazione con la Regione Trentino Alto Adige/Südtirol¹⁹**, per giungere a due piattaforme gemelle di pubblicazione dei dati aperti e sviluppare sinergie in ottica europea per la qualità dei dati resi riutilizzabili nei rispettivi territori.

A partire dall'autunno del 2013 la Provincia ha inoltre provveduto alla definizione delle azioni per la prevenzione della corruzione e ha realizzato la sezione "Amministrazione trasparente" all'interno del sito istituzionale. Annualmente, come previsto dal Piano di prevenzione della corruzione 2015 - 2017 (deliberazione della Giunta provinciale n. 19/2015) la struttura provinciale competente in materia di open data comunica lo stato di avanzamento del processo di apertura dei dati alle strutture competenti in materia di trasparenza e anticorruzione.

Il sistema trentino oggi, rispetto al 2012, ha dunque **la possibilità di confrontarsi con un quadro di maggior certezza e chiarezza sullo sviluppo della materia, sia sul versante europeo che su quello nazionale**. Su questi temi è determinante il **coordinamento e la standardizzazione**: open data significa innanzitutto riutilizzo dei dati, il valore dei dati come viene prodotto negli uffici e nei sistemi informativi è infatti correlato con la possibilità che le informazioni "aperte" possano essere utilizzate come materia prima per costruire servizi e quindi valore per **nuovi, diffusi ed innovativi mercati**. È dunque necessaria una rete di condivisione delle attuali buone pratiche, per poter passare anche in Trentino dalla conclusione della fase di avvio del processo di apertura dei dati alla messa a regime dello stesso, sia per gli aspetti organizzativi e procedurali interni all'amministrazione che per favorire il costante allineamento alla rete che definisce il processo di valorizzazione del patrimonio informativo nazionale e dell'Unione europea.

Questo documento è un prodotto costruito anche con quella comunità che in Trentino opera alla luce delle nuove tecnologie, che essendo estremamente trasversali e pervasive **finiscono per impattare in moltissimi processi e richiedono dunque** chiarezza di obiettivi, di metodi, di confini normativi e di modalità di coordinamento **su temi nuovi** e per certi aspetti ancora da definire in modo strutturato.

19 In questa cornice, la Regione Trentino Alto Adige/Südtirol nel mese di ottobre 2015 ha organizzato un intervento formativo - in collaborazione fra il Progetto Open Data in Trentino e la struttura competente per gli open data della Provincia autonoma di Bolzano - rivolto ai dirigenti di entrambe le Province (<http://www.innovazione.provincia.tn.it/notizie/pagina632.html>).

2. Finalità, destinatari e alcune definizioni utili

2.1 Finalità del documento

Questo documento ha l'obiettivo di allineare quanto realizzato in Trentino in questi anni agli sviluppi nazionali ed europei in materia di valorizzazione del patrimonio informativo pubblico, oltre che di definire strumenti efficaci per mettere a sistema il processo di produzione di dati di qualità nel sistema pubblico trentino.

In Trentino il processo di apertura dei dati della pubblica amministrazione si è sviluppato attraverso la definizione di **due strumenti - risorse**:

- **il gruppo Open Data in Trentino²⁰**, sorto dalla collaborazione fra Provincia autonoma di Trento, la società Informatica Trentina e - nei primi tre anni di attività - l'associazione TrentoRise²¹ (FBK²² e DISI²³), che ha avviato il progetto, secondo quanto indicato nel Piano di miglioramento 2012 - 2016, e ha supportato la Provincia per accelerare il processo di apertura dei dati. Con le presenti Linee guida, **nella nuova cornice di governo del progetto**, la Provincia intende **mettere a sistema quanto realizzato concretamente nella fase di avvio (oggi sostanzialmente conclusa), i nuovi obiettivi da perseguire e il nuovo modello organizzativo, date le risorse oggi disponibili**;
- **la piattaforma dati.trentino.it²⁴**, strumento e **luogo di aggregazione e di apprendimento**, nella pratica, del cambiamento organizzativo e di cultura interna agli uffici connesso con il paradigma open data. Nel portale confluisce l'intera catena di azioni e conoscenze che, partendo dall'offerta di dati da parte delle strutture del servizio pubblico in modo governato e legittimo, ha facilitato successivamente alcune esperienze di riutilizzo, supportando anche la definizione dei primi standard di qualità per formati, metadati, modellazione. La piattaforma è in gestione alla società Informatica Trentina, che ne garantisce la manutenzione e gli sviluppi²⁵.

Questo documento **si occupa di tematiche trasversali**, non approfondisce dunque il tema di come il paradigma open data sia da declinare all'interno dei diversi domini tematici (beni culturali, ambiente, turismo, ecc.). Offre piuttosto strumenti che si intendono utili a realizzare **un modello**

20 <http://www.innovazione.provincia.tn.it/opendata>

21 http://www.innovazione.provincia.tn.it/documentazione/in_trentino/pagina182.html

22 Fondazione Bruno Kessler - Center for Information and Communication Technology

23 Dipartimento di Ingegneria e Scienza dell'Informazione - Università di Trento

24 <http://dati.trentino.it/>

25 Recentemente la Provincia autonoma di Bolzano ha pubblicato un proprio catalogo di dati frutto anche della collaborazione con dati.trentino.it. La piattaforma della Provincia autonoma di Bolzano è infatti una versione aggiornata di quella trentina, che a breve a sua volta potrà riutilizzare quella stessa piattaforma (creando due piattaforme gemelle in grado di condividere sviluppi, come anche di mantenere le proprie specifiche autonomie).

generale di valorizzazione del patrimonio informativo pubblico attento alla qualità dei dati. In fase di revisione delle linee guida del 2012 sono pervenute una serie di sollecitazioni in merito **alla necessità di contestualizzare la qualità dei dati all'interno dei diversi ambiti di applicazione.** Il tema è cruciale per la “messa a sistema del processo”, in quanto significa poter analizzare dentro ogni specifica struttura o ufficio come tale materia impatti sulle pratiche, sulla normativa e sui modelli di crescita economica connessi, ma riguarda anche una delle attività principali da programmare per i prossimi anni, una volta definiti i criteri di qualità generali e allineati agli standard a quelli nazionali ed europei.

2.1.1 Perché un punto di accesso unico territoriale

Il secondo punto degli otto principi dell'Open Government Data di Sunlight Foundation²⁶ introduce il concetto di **dati primari**²⁷; tale caratteristica ha come conseguenza il fatto che **i dati aperti debbano stare il più vicino possibile a chi ne ha titolarità**, cioè a chi li ha prodotti o li ha fatti produrre. Tale accortezza garantisce che i saperi, incorporati nelle pratiche di chi crea quegli archivi, non vengano persi o fraintesi al momento del riutilizzo. **È quindi buona pratica realizzare una pagina “open data” dentro i propri portali** e in questa pagina fornire documentazione circa il **senso con cui interpretare quelle informazioni.** Se però i dati aperti restassero frammentati in più portali si incorrerebbe nel **rischio di frammentare “la cultura del dato” in differenti modalità operative, standard, interpretazioni della norma.**

In un mondo complesso e tecnologicamente assistito è invece importante definire **modalità facili e di base per dialogare e semplificare la complessità.**

Per questo motivo si è realizzata una soluzione in grado di offrire servizi trasversali a chi vi aderisca: la Provincia ha scelto di utilizzare la **piattaforma CKAN**²⁸ che garantisce un'ampia diffusione fra le piattaforme open source a livello internazionale e che è oggi anche lo strumento di catalogazione e di accesso ai data set dell'European Data Portal. **Il Trentino ha inoltre da subito integrato la piattaforma CKAN con la tecnologia che il Consorzio dei Comuni trentini ha realizzato per gestire i portali dei Comuni (ComunWeb) e a breve anche il portale istituzionale della Provincia utilizzerà la stessa tecnologia sottostante: sarà così possibile generare automaticamente dati, che grazie ad un lavoro di allineamento sulle classificazioni, saranno “nativi open data” e facilmente integrabili con la piattaforma CKAN di dati.trentino.it dalle piattaforme ComunWeb e dal nuovo portale della Provincia.**

La scelta di questa piattaforma è del resto coerente con il modello di governance che la Provincia si è data con la L.P. 3/2006, la quale dispone che **“la Provincia in attuazione dei principi di sussidiarietà, adeguatezza, differenziazione, efficienza, qualità, economicità e semplificazione istituzionale promuove [...] l'esercizio in forma associata di funzioni e servizi da parte dei comuni, tenendo conto**

26 <http://opengovdata.org/>

27 <http://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf> M-13-13

28 <http://ckan.org/>

delle rispettive dimensioni territoriali, al fine di superare la frammentarietà, attuare obiettivi di coesione territoriale, elevare il livello di qualità delle prestazioni e ridurre complessivamente gli oneri organizzativi, procedurali e finanziari in funzione del rafforzamento dell'efficacia delle politiche pubbliche." In altre parole, un punto unico di accesso al patrimonio informativo pubblico del sistema territoriale e la messa in comune dei saperi connessi al paradigma open data **si configura come strumento per promuovere la realizzazione, con lo specifico contributo di tutti, di un modello che rafforzi l'efficacia delle politiche pubbliche.**

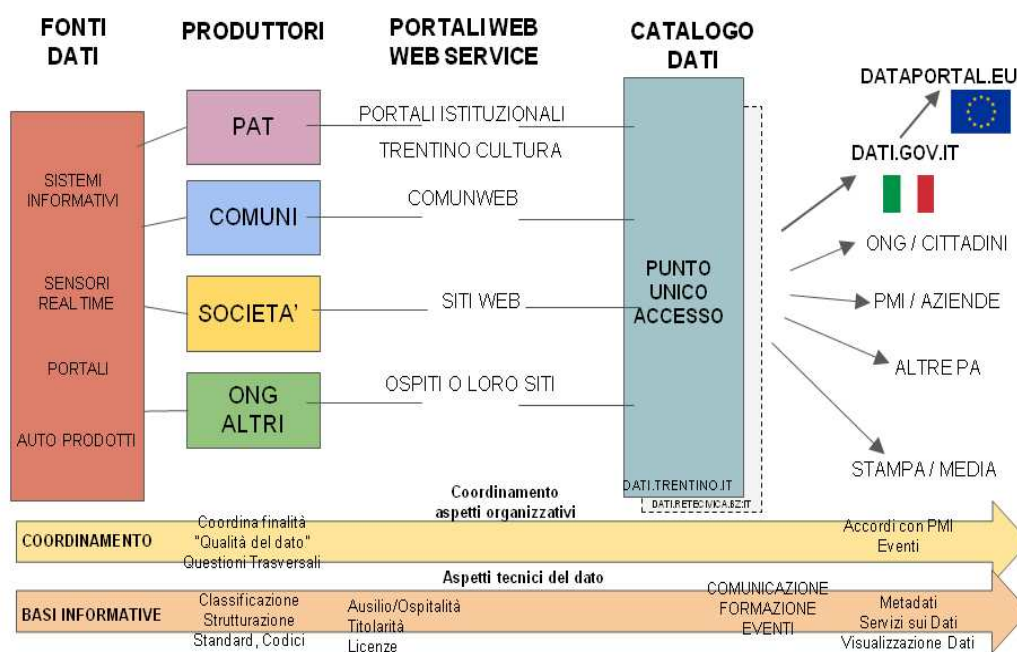


Figura 1 Schema illustrativo delle relazioni fra le varie fonti di dati, i vari soggetti che ne hanno titolarità e l'utilità di una piattaforma punto unico di accesso da cui soggetti terzi esterni al sistema locale possono avere accesso a dati standardizzati in modo trasversale e rilancia i dati "locali" tramite piattaforme europee capaci di raccogliere l'esperienza open data dei 27 Paesi membri e tradurla in linguaggi comuni. Nelle **freccie alla base dello schema due funzioni che sono descritte nel paragrafo 4.6.5**: la prima riguarda il coordinamento fra gli attori del sistema (funzioni di governo alte); la seconda riguarda il livello di integrazione di basi dati, controllo aggiornamento normativo, ecc. (funzione di garanzia della qualità del dato a livello di Team Open Data).

Un punto unico di accesso è oggi inoltre un **passaggio necessario anche in termini di coordinamento nazionale ed europeo**: il portale dati.gov.it è il punto unico di accesso per i dati aperti del servizio pubblico nel nostro Paese, ed è il punto di riferimento per l'European Data Portal, la piattaforma sulla quale già oggi convergono **quasi 300.000 data set provenienti dalla complessità linguistica,**

normativa, di pratiche amministrative dei 27 stati membri. Quanto più a livello di singoli territori saranno individuati standard comuni nei codici, nella metadattazione, nell'uso di vocabolari, tanto più i singoli dataset acquisteranno valore all'interno del Mercato Unico Digitale europeo.

L'adesione alla piattaforma dati.trentino.it impegna quanti vi aderiscano **a costruire un confronto e una collaborazione per trovare soluzioni** che garantiscano una situazione dove le opportunità vengono dalla valorizzazione **del sistema territoriale seguendo una modalità strategica win-win.** Per rendere sostenibile nel tempo questo processo si è definito, attraverso le pratiche di questi primi anni di progetto, il modello organizzativo descritto nel capitolo 4.

2.2 Destinatari del documento

Sono destinatari delle presenti linee guida tutti le organizzazioni che aderiscono o aderiranno a dati.trentino.it.

Questo documento, **previsto dalla L.P. 16/2012²⁹**, è rivolto principalmente a tutti gli enti che aderiscono al SINET (Sistema informativo elettronico trentino), ai Comuni, alle Comunità e alle società di sistema provinciali. Ma poiché in questa strategia la quantità e la diversità dei dati aumenta il valore generale del processo, si è deciso di rendere disponibile la piattaforma anche ad altri enti quali la Regione Trentino Alto Adige/Südtirol e ad altre associazioni e organizzazioni che, di volta in volta e sulla base di specifici accordi, vorranno aderire alla stessa, con l'obiettivo di mettere a disposizione del sistema territoriale strumenti utili per rendere il Trentino un territorio "ad alta cultura del dato" .

2.3 Definizioni

La Tabella 1 propone un glossario in cui sono riportate le definizioni, con rispettiva fonte, di una serie di concetti incontrati nel corso di questi anni di lavoro sui temi in questione. La finalità del glossario è di **rendere meno ambigue alcune categorie** che apparentemente possono sembrare appartenenti ai medesimi campi di significato ma che, **poiché definite dentro domini differenti, sono invece diverse nella loro essenza.** La tabella è anche uno dei modi per descrivere il campo semantico degli open data.

Tabella 1. Riassunto delle definizioni di concetti centrali o simili in ambito open data secondo la normativa o standard internazionali condivisi.

Concetto	Definizione	Fonte
Public Sector Information	"E' l'informazione in possesso degli enti pubblici, ivi compresi gli organismi di diritto pubblico (cfr. DIR-2003/98/CE relativa al riutilizzo dell'informazione del settore pubblico ss.mm.ii. e il d.lgs. 36/2003). L'informazione pubblica si caratterizza per essere di tipo statico o di tipo	Dir-2003/98/CE; d. Lgs. 36/2006; Linee Guida AgID

29 <http://tinyurl.com/pwp7csw>

(PSI)	dinamico. L'informazione statica è rappresentata dal contenuto informativo in possesso della Pubblica Amministrazione (ad esempio gli archivi dei beni culturali); l'informazione dinamica è invece prodotta dalle istituzioni pubbliche nello svolgimento dei propri compiti istituzionali (ad esempio i dati di bilancio di un ente)"	Open Data
Dato	"Rappresentazione fisica di fatti, eventi atomici, fenomeni oggettivi, informazioni atte alla comunicazione, interpretazione ed elaborazione da parte di esseri umani o mezzi automatici"	Linee Guida AgID Open Data 2014
Dati di tipo aperto (open data)	I dati tipicamente non riferibili a singole persone "che presentano le seguenti caratteristiche: 1) sono disponibili secondo i termini di una licenza che ne permetta l'utilizzo da parte di chiunque, anche per finalità commerciali, in formato disaggregato; 2) sono accessibili attraverso le tecnologie dell'informazione e della comunicazione, ivi comprese le reti telematiche pubbliche e private, in formati aperti ai sensi della lettera a), sono adatti all'utilizzo automatico da parte di programmi per elaboratori e sono provvisti dei relativi metadati; 3) sono resi disponibili gratuitamente attraverso le tecnologie dell'informazione e della comunicazione, ivi comprese le reti telematiche pubbliche e private, oppure sono resi disponibili ai costi marginali sostenuti per la loro riproduzione e divulgazione"	CAD art. 68, c. 3, lett. b); ripreso anche dalle Linee Guida AgID Open Data;
Open source	"Quella categoria di software che può essere liberamente usato, cambiato e condiviso (allo stesso modo o in forma modificata) da chiunque. Il software open source è il prodotto di molte persone e viene distribuito tramite licenze che rispettano la Open Source Definition" . Open Data e Open Source sono due cose differenti che riguardano una i dati, l'altra i software	http://opensource.org/
Dataset	(1)"una collezione di dati, generalmente riguardanti una stessa organizzazione, che vengono erogati e gestiti congiuntamente"; (2)"insieme di dati strutturati in forma relazionale"	(1) Linee Guida AgID Open Data; (2) Linee Guida AgID Interoperabilità
Documento	"a) qualsiasi contenuto, a prescindere dal suo supporto (testo su supporto cartaceo o elettronico, registrazione sonora, visiva o audiovisiva ³⁰); b) qualsiasi parte di tale contenuto"	Direttiva 2003/98/CE art. 2 comma 3 ³¹
Banca di dati	"Raccolte di opere, dati o altri elementi indipendenti sistematicamente o metodicamente disposti ed individualmente accessibili mediante mezzi elettronici o in altro modo"	art. 2, n. 9, L. 33/1941
Base di dati di interesse nazionale	"Si definisce base di dati di interesse nazionale l'insieme delle informazioni raccolte e gestite digitalmente dalle pubbliche amministrazioni, omogenee per tipologia e contenuto e la cui conoscenza è utilizzabile dalle pubbliche amministrazioni, anche per fini statistici, per l'esercizio delle proprie funzioni e nel rispetto delle competenze e delle normative vigenti."	CAD Sezione 2, art. 60
Titolare del dato	"la pubblica amministrazione o l'organismo di diritto pubblico che ha originariamente formato per uso proprio o commissionato ad altro soggetto pubblico o privato il documento che rappresenta il dato o che ne ha la disponibilità"	art. 2, c. 1, lett. i, D. Lgs. 36/2006
Titolare del trattamento di dati personali	"la persona fisica, la persona giuridica, la pubblica amministrazione e qualsiasi altro ente, associazione od organismo cui competono, anche unitamente ad altro titolare, le decisioni in ordine alla finalità, alle modalità del trattamento di dati personali e agli strumenti utilizzati, ivi compreso il profilo della sicurezza"	art. 4, comma 1, lett. f, D. Lgs. 196/2003
Trattamento di dati personali	"qualunque operazione o complesso di operazioni effettuati anche senza l'ausilio di strumenti elettronici concernenti la raccolta, la registrazione, l'organizzazione, la conservazione, la consultazione, l'elaborazione, la modificazione, la selezione, l'estrazione, il raffronto, l'utilizzo, l'interconnessione, il blocco, la comunicazione, la diffusione, la cancellazione e la distruzione di dati, anche se non registrati in una banca di dati".	art. 4, comma 1, lett. a, D. Lgs. 196/2003

30 E' bene notare che ognuna di tali codifiche (su supporto cartaceo o elettronico) è da considerarsi comunque come un dato.

31 Sul piano ontologico questa definizione è problematica perché porta a negare un "criterio di integrità" per un documento che comporta un rischio di perdita di integrità del documento.

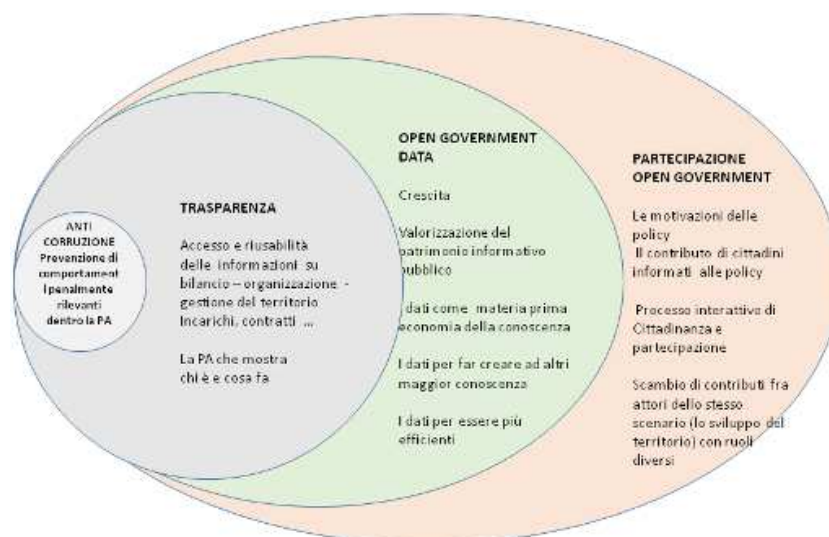
Dato pubblico	"il dato conoscibile da chiunque". E' pubblicabile come open data.	art. 1, c. 1, lett. n, CAD e art. 2, c. 1, lett. d), D. Lgs. 36/2006
Dato a conoscibilità limitata	"dati la cui conoscibilità è riservata per legge o regolamento a specifici soggetti o categorie di soggetti". Solitamente non sono pubblicabili come open data.	Linee Guida AgID Interoperabilità
Dato personale	"qualunque informazione relativa a persona fisica, persona giuridica, ente od associazione, identificati o identificabili, anche indirettamente, mediante riferimento a qualsiasi altra informazione, ivi compreso un numero di identificazione personale". Non sono pubblicabili come open data a parte casi eccezionali (es. obbligo di legge o di regolamento).	art. 4, c. 1, l. b, Codice Privacy
Dati sensibili	"I dati personali idonei a rivelare l'origine razziale ed etnica, le convinzioni religiose, filosofiche o di altro genere, le opinioni politiche, l'adesione a partiti, sindacati, associazioni od organizzazioni a carattere religioso, filosofico, politico o sindacale, nonché i dati personali idonei a rivelare lo stato di salute e la vita sessuale." Non sono pubblicabili come open data.	art. 4, c. 1, l. d, Codice Privacy
Anonimizzazione	"L'elaborazione dei dati che comprende informazioni personali in modo che le persone non siano più identificabili dai dati risultanti. L'anonimizzazione consente ai dati di essere pubblicati senza violare il principio della protezione dei dati. Le tecniche principali sono l'aggregazione e de-identificazione. Si deve prestare attenzione per evitare la fuga di dati che si potrebbe tradurre nel fatto che venga compromessa la vita privata delle persone.	OKFN - Open data Handbook – Glossario
Dato statistico	"dato proveniente da uno studio o da una sorgente amministrativa, utilizzati per produrre statistiche e/o dati che comprendono tali statistiche". Pubblicabili come open data ad alcune condizioni.	Linee Guida AgID Interoperabilità
Metadato	(1)"L'Iso (International organization for standardization) definisce i metadati come "dati che descrivono e definiscono altri dati in un determinato contesto". Il contesto riguarda le condizioni in cui avviene il trattamento dei dati." (2)"Nel caso specifico dei metadati descrittivi, essi costituiscono un livello di informazioni che corredano i dati, ne esplicitano alcune caratteristiche e ne rendono più semplice l'identificazione. Per agevolare la reperibilità dei dati e la loro interoperabilità è importante utilizzare elementi descrittivi come titolo, descrizione, link, indicazione della licenza, periodo di validità, ente gestore, formato, ecc."	1) ISTAT http://www3.istat.it/servizi/studenti/valoredati/Cap3/Cap3_3.htm 2) FAQ dati.trentino.it
Formato di dati di tipo aperto	"un formato di dati reso pubblico, documentato esaurientemente e neutro rispetto agli strumenti tecnologici necessari per la fruizione dei dati stessi"	art. 68, c. 3, lett. a), CAD
Standard formale aperto	(1)"uno standard che è stato definito in forma scritta, precisando in dettaglio i requisiti per assicurare l'interoperabilità del software." Lo standard può dunque essere considerato aperto quando: (2)"è adottato e mantenuto da un'organizzazione non-profit ed il suo sviluppo avviene sulle basi di un processo decisionale aperto e a disposizione di tutti gli interlocutori interessati (decisioni consensuali, a maggioranza ecc.); lo standard è pubblicato e il documento di specifiche è disponibile liberamente oppure ad un costo nominale. Deve essere possibile farne copie, riusarlo e distribuirlo liberamente senza alcun costo aggiuntivo; eventuali diritti di copyright, brevetti o marchi registrati sono irrevocabilmente concessi sotto forma di royalty-free; non è presente alcun vincolo al riuso, alla modifica e all'estensione dello standard"	(1)Direttiva 2003/98/CE (e s.m.i.) art. 2, c. 1, n. 8; (2)Final European Interoperability Framework - November 2004 http://ec.europa.eu/idabc/en/document/3473/5585.html
Formato leggibile meccanicamente	"un formato di file strutturato in modo tale da consentire alle applicazioni software di individuare, riconoscere ed estrarre facilmente dati specifici, comprese dichiarazioni individuali di fatto e la loro struttura interna"	Direttiva 2003/98/CE (e s.m.i.)
Licenza standard per il riutilizzo	"il contratto, o altro strumento negoziale, redatto ove possibile in forma elettronica, nel quale sono definite le modalità di riutilizzo dei documenti delle pubbliche amministrazioni o degli organismi di diritto pubblico". Si tratta in generale di licenze condivisa, elaborata e approvata da comunità di giuristi nel web che definisce le condizioni e le modalità di riutilizzo da parte di chiunque anche per scopi commerciali (conformemente all'art. 8 del D. Lgs. 36/2006, all'art. 9 della L. P. 16/2012 e all'art. 68 CAD). La PAT fa uso delle licenze Creative Commons Zero (CC0) e Creative Commons Attribuzione (CC-BY)	art. 2, c. 1, lett. h, D. Lgs. 36/2006; art. 8 del D. Lgs. 36/2006; art. 9 L. P. 16/2012; art. 68 CAD

Accessibilità dei sistemi informatici	"la capacità dei sistemi informatici, nelle forme e nei limiti consentiti dalle conoscenze tecnologiche, di erogare servizi e fornire informazioni fruibili, senza discriminazioni, anche da parte di coloro che a causa di disabilità necessitano di tecnologie assistive o configurazioni particolari "	legge n.4 9 gennaio 2004, art. 2, a)
Accessibilità del dato	"Richiama l'aspetto tecnologico del formato aperto in cui il dato viene reso disponibile e richiede l'effettiva presenza dei metadati relativi al dato medesimo."	Linee Guida AgID Open Data
Diritto di accesso civico	"Il diritto di chiunque di richiedere la pubblicazione di documenti, informazioni o dati per i quali sussistono specifici obblighi di trasparenza, nei casi in cui sia stata omessa la loro pubblicazione. Pertanto, l'accesso civico si configura come rimedio alla mancata pubblicazione, obbligatoria per legge, di documenti, informazioni o dati sul sito istituzionale. "	cfr. art. 5 del d.lgs. n. 33/2013; e * FAQ ANAC
Diritto di accesso ad atti e documenti amministrativi	(1)"Il diritto degli interessati a prendere visione e ad estrarre copia di documenti amministrativi. Ne sono titolari tutti i cittadini, società e associazioni, compresi quelli portatori di interessi pubblici o diffusi, che abbiano un interesse diretto, concreto ed attuale, corrispondente ad una situazione giuridicamente tutelata e collegata al documento di cui si chiede l'accesso. " Mentre (2)"l'accesso civico introduce una legittimazione generalizzata a richiedere la pubblicazione di documenti, informazioni o dati per i quali sussiste l'obbligo di pubblicazione da parte delle pubbliche amministrazioni ai sensi della normativa vigente, [...] il diritto di accesso agli atti è [invece] finalizzato alla protezione di un interesse giuridico particolare , può essere esercitato solo da soggetti portatori di tali interessi e ha per oggetto atti e documenti individuati. "	(1) legge n. 241/1990 e LP 23/1992. (2) FAQ ANAC
Disponibilità	"la possibilità di accedere ai dati senza restrizioni non riconducibili a esplicite norme di legge"	art 1, c. 1, lett. o, CAD
Fruibilità	(1)"La possibilità di utilizzare il dato anche trasferendolo nei sistemi informativi automatizzati di un'altra amministrazione"; (2)"Il trasferimento di un dato da un sistema informativo ad un altro non modifica la titolarità del dato" Quindi è necessario valutare sempre la titolarità originaria dei dati che si intende aprire.	(1) CAD, art. 1, c. 1, lett. f) (2) CAD art. 58 lett. c)
Riutilizzo	"l'uso di documenti in possesso di enti pubblici da parte di persone fisiche o giuridiche a fini commerciali o non commerciali diversi dallo scopo iniziale nell'ambito dei compiti di servizio pubblico per i quali i documenti sono stati prodotti. Lo scambio di documenti tra enti pubblici esclusivamente in adempimento dei loro compiti di servizio pubblico non costituisce riutilizzo"	Direttiva 2003/98/CE art. 2, comma 1, 4)
Interoperabilità	"In ambito informatico, la capacità di sistemi differenti e autonomi di cooperare e di scambiare informazioni in maniera automatica , sulla base di regole comunemente condivise"	Linee Guida AgID Open Data
Interoperabilità semantica	"La capacità di elaborare informazioni da fonti esterne o secondarie senza perdere il reale significato delle informazioni stesse nel processo di elaborazione"	Linee Guida AgID Open Data
Protocollo di rete aperto	"La definizione formale, di pubblico dominio ed esente da vincoli di copyright, di marchio e di brevetto, che descrive le modalità di interazione che, due o più apparecchiature elettroniche collegate tra loro , devono rispettare per operare particolari funzionalità di elaborazione necessarie all'espletamento di un certo servizio di rete."	PaT - Linee Guida per il riutilizzo e la diffusione dei dati pubblici - del. G. P. n. 2858 27/12/2012
Ontologia	"In ambito informatico una rappresentazione formale e condivisa dei concetti e delle mutue relazioni che caratterizzano un certo dominio di conoscenza "	Linee guida AgID Open Data

3. Il contesto giuridico

Prima di entrare nel dettaglio del contesto giuridico di riferimento è utile presentare uno schema di interazione fra diversi temi che sono all'ordine del giorno all'interno degli uffici anche per effetto dell'innovazione normativa che si illustra di seguito (**Figura 2**). Secondo questo modo di osservare, gli open data sono uno dei temi relativi alla trasformazione del rapporto fra servizio pubblico e cittadini. Essi si collocano come strumento per rendere riutilizzabile “la materia prima necessaria alla conoscenza”, sia essa finalizzata alla trasparenza, all'anticorruzione, alla comunicazione o alla crescita economica. Gli open data si collocano **ad un livello trasversale e tecnico** per le azioni descritte come “anticorruzione”, “trasparenza”, “crescita economica”, “efficienza della pubblica amministrazione”, “partecipazione”.

Figura 2 Come si collocano gli open data su alcuni importanti processi di cambiamento della pubblica amministrazione in corso



Questo capitolo, dopo aver introdotto le coordinate giuridiche necessarie per orientarsi nella disciplina del riutilizzo dell'informazione pubblica, approfondisce gli aspetti più importanti legati alla traduzione operativa del processo di apertura dei dati pubblici, quali la titolarità dei dati, le licenze, l'individuazione dei dati che possono essere pubblicati come open data, le richieste di riutilizzo, le questioni legate ai costi e agli eventuali accordi di esclusiva. Tramite l'esperienza maturata dal Progetto Open Data in questi anni, e grazie all'analisi critica delle migliori pratiche in costante collegamento ai network di discussione nazionali ed europei, le soluzioni qui descritte sono aggiornate allo stato dell'arte disponibile al momento della definizione delle presenti linee guida.

3.1. Quadro di sintesi normativa

Il panorama giuridico di condivisione e valorizzazione dell'informazione del settore pubblico è molto vario e opera su più livelli: parte dalle normative europee, passa attraverso il recepimento nazionale e viene tradotto operativamente negli enti pubblici con strumenti giuridici che individuano specifici obiettivi e abilitano processi interni ed esterni.

Le normative sul riutilizzo dell'informazione del settore pubblico si sono sviluppate partendo da domini verticali, in particolare dall'informazione ambientale (2003) e geografica (2007). Da questi due grandi settori, e da altri specifici interventi, si è delineata una cornice giuridica di più ampio respiro, confluita nella Direttiva PSI (del 2006, modificata nel 2013), utile a qualificare la disciplina del riutilizzo dell'informazione pubblica come un istituto giuridico applicabile a tutti i dati pubblici, salvaguardando specifici diritti e interessi.

Ad oggi, la valorizzazione dei dati pubblici è una delle azioni chiave nell'agenda digitale europea, Europa2020, e quindi del Digital Single Market, in quanto infrastrutturale alla crescita economica e alla costruzione di servizi innovativi per i cittadini.

Questo paragrafo **analizza il contesto normativo e approfondisce ad uso pratico i temi e gli strumenti** che permettono di condividere e valorizzare l'informazione del settore pubblico **in modo legittimo**.

3.1.1 Normativa sul riutilizzo delle informazioni del settore pubblico

Di seguito sono riassunte le direttive europee e i rispettivi recepimenti nazionali e provinciali di più ampio respiro e infrastrutturali rispetto ai domini verticali, che verranno analizzati in seguito. Sono qui contenute le **definizioni di alto livello** sul tema del riutilizzo e della valorizzazione del patrimonio informativo pubblico, nonché i **principi e le salvaguardie** che devono essere sempre tenuti in considerazione in ogni dominio applicativo.

Europa	Italia	Trentino
<u>Direttiva 2003/98/CE (PSI)</u>	<u>Decreto Legislativo 36/06</u> "Attuazione della direttiva 2003/98/CE relativa al riutilizzo di documenti nel settore pubblico"	<u>Legge provinciale 16/12</u> "Disposizioni per la promozione della società dell'informazione e della amministrazione digitale e per la diffusione del software libero e dei formati di dati aperti"; <u>Deliberazione della Giunta provinciale 2858/12</u> "Approvazione delle "Linee guida per il riutilizzo e la diffusione dei dati pubblici" in attuazione legge provinciale 27 luglio 2012, n. 16, articolo 9"
	<u>Decreto Legge 179/2012</u> "Ulteriori misure urgenti per la crescita del Paese" conv. in <u>Legge 221/2012</u> , art. 9	

Direttiva 2013/37/UE (PSI)	Decreto Legislativo 102/2015 di attuazione della direttiva 2013/37/UE che modifica la Direttiva 2003/98/CE, relativa al riutilizzo dell'informazione del settore pubblico	
-------------------------------	--	--

Europa

Direttiva 2003/98/CE “Relativa al riutilizzo dell'informazione del settore pubblico”: getta le basi normative per l'avvio dei processi di apertura dei dati e dei documenti pubblici nelle diverse amministrazioni dei Paesi Membri allo scopo di favorirne il riutilizzo da parte di chiunque e per qualunque finalità, **fissando regole minime** in materia di accesso, licenze, regole di non discriminazione tra riutilizzatori e principi di tariffazione.

Direttiva 2013/37/UE (nuova PSI): modifica la direttiva 2003/98/CE relativa al riutilizzo dell'informazione del settore pubblico e stabilisce che l'apertura di dati e documenti pubblici per il riutilizzo **non è più una facoltà degli enti pubblici, ma è un obbligo**. La direttiva ha ampliato i settori soggetti alla normativa sul riutilizzo anche al settore dei beni culturali, ha fissato regole sui formati per il riutilizzo e ha stabilito il principio del costo marginale in materia di tariffazione.

Italia

Decreto Legislativo 36/2006 “Attuazione della direttiva 2003/98/CE relativa al riutilizzo di documenti nel settore pubblico” e s.m.i.: ha recepito nell'ordinamento italiano la direttiva 2003/98/CE ed è stato modificato nel 2015 per implementare le modifiche previste dalla Direttiva 2013/37/UE (nuova PSI).

Decreto Legislativo 82/2005 “Codice dell'Amministrazione digitale” e s.m.i.: le norme di riferimento in materia di **open data sono contenute agli artt. 50, 52 e 68**, come innovati in particolare dall'art. 9 del *decreto legge 18 ottobre 2012, n. 179 “Ulteriori misure urgenti per la crescita del Paese”, convertito con legge 17 dicembre 2012, n. 221*, stabilendo il principio di open by default per i dati pubblici e fornendo le definizioni di dati aperti e di formati aperti (cfr. tabella 1).

Decreto Legislativo 179/2012 “Ulteriori misure urgenti per la crescita del Paese” e s.m.i.: prevede innovazioni importanti alla cui base sono collocati i dati aperti, modificando le leggi sopra citate. Oltre all'art. 9 di modifica degli articoli del Codice dell'Amministrazione Digitale, di fondamentale importanza è **l'articolo 20, rubricato “Comunità intelligenti”**, che definisce un ruolo di coordinamento centrale degli standard tecnici e organizzativi utili allo sviluppo dell'innovazione sul territorio.

Decreto Legislativo 102/2015 di attuazione della direttiva 2013/37/UE che modifica la direttiva 2003/98/CE, relativa al riutilizzo dell'informazione del settore pubblico: interviene a recepire nell'ordinamento italiano le innovazioni sul tema del riutilizzo dell'informazione pubblica che

l'Unione Europea ha definito nella Direttiva 2013/37/UE. Dopo aver allineato le definizioni con il CAD, abilita il riutilizzo di alcune tipologie di dati culturali e demanda ai ministeri competenti, al Ministro dell'Economia e delle Finanze e all'Agenzia per l'Italia Digitale la stesura delle linee guida per gli aspetti del costo dei dati (in particolare, riguardo le tariffe e le relative modalità di versamento).

Trentino

Legge provinciale 16/2012 "Disposizioni per la promozione della società dell'informazione e dell'amministrazione digitale e per la diffusione del software libero e dei formati di dati aperti": assicura la messa a disposizione ai fini del riutilizzo dei dati pubblici e dei documenti di cui è titolare, in attuazione della legislazione italiana ed europea sul riutilizzo dell'informazione del settore pubblico.

Deliberazione della Giunta provinciale n. 2858/2012 di approvazione delle "Linee guida per il riutilizzo e la diffusione dei dati pubblici" in attuazione legge provinciale 27 luglio 2012, n. 16, articolo 9": adotta le Linee guida per il riutilizzo e la diffusione di dati pubblici per la Provincia Autonoma di Trento e predispone gli strumenti tecnici e organizzativi per supportare la politica di apertura dei dati pubblici del Sistema trentino. Le Linee guida del 2012 hanno stabilito gli standard giuridici e tecnologici per la pubblicazione degli open data provinciali.

3.1.2 Normative su domini specifici

A partire dal 2003, una serie di direttive europee hanno abilitato la valorizzazione di specifici domini di dati particolarmente utili in quanto maggiormente rilevanti all'interno delle banche dati pubbliche e strategiche rispetto alla costruzione di un'economia di servizi innovativi. L'obiettivo è la definizione di standard condivisi in Europa, affinché i dati e le informazioni siano più interoperabili e riutilizzabili entro un mercato unico digitale.

Europa	Italia	Trentino
Direttiva 2003/4/CE "sull'accesso del pubblico all'informazione ambientale e che abroga la direttiva 90/313/CEE del Consiglio"	Decreto Legislativo 195/2005 "Attuazione della direttiva 2003/4/CE sull'accesso del pubblico all'informazione ambientale"	
Direttiva 2007/2/CE "che istituisce un'infrastruttura per l'informazione territoriale nella Comunità europea (Inspire)"	Decreto Legislativo 32/2010 "Attuazione della direttiva 2007/2/CE, che istituisce un'infrastruttura per l'informazione territoriale nella Comunità europea (INSPIRE)"	Deliberazione della Giunta provinciale 195/2012 "Autorizzazione al rilascio di alcuni dati del Sistema Informativo"

		Ambiente e Territorio (SIAT) secondo il paradigma degli Open Government Data (OGD)"
Direttiva 2010/40/UE "sul quadro generale per la diffusione dei sistemi di trasporto intelligenti nel settore del trasporto stradale e nelle interfacce con altri modi di trasporto"	Decreto Ministero delle Infrastrutture e dei Trasporti del 1 febbraio 2013 "Diffusione dei sistemi di trasporto intelligenti (ITS) in Italia"	

Direttiva 2003/4/CE "sull'accesso del pubblico all'informazione ambientale", recepita dall'ordinamento Italiano con il D. Lgs. 19 agosto 2005, n. 195. La normativa è collegata all'istituto dell'accesso, ma esonera il richiedente dal dimostrare un interesse specifico. Inoltre, presenta la necessità di costruire cataloghi pubblici dell'informazione ambientale, espone alcuni principi in tema di tariffazione e relativi alla qualità di questi dati.

Direttiva 2007/2/CE "che istituisce un'infrastruttura per l'informazione territoriale nella Comunità europea (INSPIRE)": ha fissato standard tecnici per l'accessibilità, la fruibilità e la riutilizzabilità dell'informazione territoriale. La direttiva INSPIRE è stata recepita dall'ordinamento Italiano con il D. Lgs. 27 gennaio 2010, n. 32, e attuata dalla Provincia Autonoma di Trento con la *deliberazione della Giunta Provinciale 17 febbraio 2012, n. 195, "Autorizzazione al rilascio di alcuni dati del Sistema Informativo Ambiente e Territorio (SIAT) secondo il paradigma degli Open Government Data (OGD)"*.

Direttiva 2010/40/UE "sul quadro generale per la diffusione dei sistemi di trasporto intelligenti nel settore del trasporto stradale e nelle interfacce con altri modi di trasporto": disciplina le modalità della diffusione dei sistemi di trasporto intelligenti nel settore del trasporto stradale e nelle interfacce con altri modi di trasporto. Per i dati relativi a questa materia la direttiva impone i requisiti e il rispetto della normativa sul riutilizzo dei dati pubblici. La direttiva è stata recepita dall'ordinamento Italiano con il *decreto del Ministero delle Infrastrutture e dei Trasporti del 1 febbraio 2013* (Diffusione dei sistemi di trasporto intelligenti (ITS) in Italia); in attuazione dell'art. 8 del decreto legge 179/2012, sono adottate le direttive con cui vengono stabiliti i requisiti per la diffusione, la progettazione, la realizzazione degli ITS, per assicurare disponibilità di informazioni gratuite di base e l'aggiornamento delle informazioni infrastrutturali e dei dati di traffico, nonché le azioni per favorirne lo sviluppo sul territorio nazionale in modo coordinato, integrato e coerente con le politiche e le attività in essere a livello nazionale e comunitario.

3.1.3 Altre normative rilevanti in materia di open data

L'informazione, quale bene primario alla base della maggior parte delle attività di un ente pubblico, è disciplinata in svariati campi di interesse collettivo. In questo senso, molte normative iniziano a prevedere più o meno esplicitamente connessioni rilevanti verso la possibilità di riutilizzare

l'informazione prodotta anche per finalità diverse rispetto al tema specifico che mira a disciplinare la normativa settoriale. Sono in seguito elencate le normative che risultano avere un impatto più rilevante sulla disciplina del riutilizzo dell'informazione pubblica.

Italia	Trentino
<p><u>Decreto Legge 83/2012</u> "misure urgenti per la crescita del Paese" (conv. in Legge 134/2012), art. 18 ("Amministrazione aperta")</p>	<p><u>Legge provinciale 25/12</u> "Disposizioni per la formazione del bilancio annuale 2013 e pluriennale 2013-2015 della Provincia autonoma di Trento (legge finanziaria provinciale 2013)", art. 32);</p> <p><u>Deliberazione della Giunta provinciale n. 2675/2013</u> "Prime disposizioni per l'attuazione degli adempimenti previsti dall'articolo 31 bis della legge provinciale 30 novembre 1992, n. 23 - Principi per la democratizzazione, la semplificazione e la partecipazione all'azione amministrativa provinciale e norme in materia di procedimento amministrativo - in materia di amministrazione aperta";</p> <p><u>Legge provinciale 9/2013 art. 13 c. 4</u> "Ulteriori interventi a sostegno del sistema economico e delle famiglie";</p>
<p><u>Legge 190/2012</u> "Disposizioni per la prevenzione e la repressione della corruzione e dell'illegalità nella pubblica amministrazione"</p>	<p><u>Legge provinciale 10/2012</u> "Interventi urgenti per favorire la crescita e la competitività del Trentino" (art. 4 bis: fissa le modalità per ottemperare agli obblighi di trasparenza in capo alle stazioni appaltanti);</p> <p><u>Legge provinciale 9/2013 art. 13 c. 4</u> "Ulteriori interventi a sostegno del sistema economico e delle famiglie";</p> <p><u>Deliberazione della Giunta provinciale n. 133/2014</u> "Piano di prevenzione della corruzione 2014 - 2016 della Provincia autonoma di Trento";</p>
<p><u>Decreto Legislativo 33/13</u> "Riordino della disciplina riguardante gli obblighi di pubblicità, trasparenza e diffusione di informazioni da parte delle pubbliche amministrazioni";</p>	<p><u>Legge provinciale 4/2014</u> "Disposizioni riguardanti gli obblighi di pubblicità, trasparenza e diffusione di informazioni e modificazione delle legge provinciale 28 marzo 2013, n. 5".</p>

Italia

Decreto Legge 83/2012 "misure urgenti per la crescita del Paese" (Decreto Sviluppo, convertito in Legge 134/2012), art. 18 ("Amministrazione aperta"): ha fissato i primi obblighi di pubblicazione come open data di dati pubblici.

Legge 190/2012 "Disposizioni per la prevenzione e la repressione della corruzione e dell'illegalità nella pubblica amministrazione": ha individuato ulteriori dati pubblici da pubblicarsi come open data. L'open data è qui considerato come principio abilitante la trasparenza e come strumento per prevenire azioni di corruzione.

Decreto Legislativo 33/2013 "Riordino della disciplina riguardante gli obblighi di pubblicità, trasparenza e diffusione di informazioni da parte delle pubbliche amministrazioni": ha riordinato e

innovato la disciplina della trasparenza amministrativa e ha stabilito che tutti i dati e i documenti a pubblicazione obbligatoria devono essere pubblicati in formati aperti e con licenze che ne permettano il riutilizzo.

Trentino

Legge provinciale 10/2012 *“Interventi urgenti per favorire la crescita e la competitività del Trentino” (Piano di miglioramento)*: include la politica open data tra le iniziative di innovazione della pubblica amministrazione per la crescita e l’efficienza della pubblica amministrazione.

Legge provinciale 9/2013 *“Ulteriori interventi a sostegno del sistema economico e delle famiglie”*: all'art. 13 disciplina l'adeguamento della normativa in tema di amministrazione aperta e obblighi open data presenti nel Decreto Legge 83/2012 (decreto sviluppo) e nella Legge 190/2012 (legge anticorruzione).

Deliberazione della Giunta provinciale n. 2675/2013 *“Prime disposizioni per l’attuazione degli adempimenti previsti dall’articolo 31 bis della legge provinciale 30 novembre 1992, n. 23 - Principi per la democratizzazione, la semplificazione e la partecipazione all’azione amministrativa provinciale e norme in materia di procedimento amministrativo - in materia di amministrazione aperta”*. Questo provvedimento attua a livello provinciale l'art. 18 del decreto sviluppo, dedicato alla pubblicazione in formato aperto dei dati inerenti alla concessione delle sovvenzioni, contributi, sussidi ed ausili finanziari alle imprese e l'attribuzione dei corrispettivi e dei compensi a persone, professionisti, imprese ed enti privati e comunque di vantaggi economici di qualunque genere di cui all'articolo 12 della legge 7 agosto 1990, n. 241 ad enti pubblici e privati.

Legge provinciale 25/2012 *“Disposizioni per la formazione del bilancio annuale 2013 e pluriennale 2013-2015 della Provincia autonoma di Trento (legge finanziaria provinciale 2013) - art. 32*”: articolo modificativo degli articoli 3, 5, 8, 16 quater, 16 quinquies, 23, 31, introduttivo degli articoli 31 bis, 33 bis e abrogativo dell'art. 42 della legge provinciale sull'attività amministrativa. Viene garantita la più ampia conoscibilità degli atti pubblici e degli altri atti soggetti a pubblicità per effetto di specifiche previsioni di legge tramite i portali istituzionali. Viene inoltre statuito l'obbligo di pubblicare concessioni di sovvenzioni, contributi, sussidi e ausili finanziari e all'attribuzione di corrispettivi e compensi a persone, professionisti, imprese ed enti privati e comunque di vantaggi economici di qualunque genere a enti pubblici e privati, come previsto dall'articolo 18 del decreto legge 22 giugno 2012, n. 83 (Misure urgenti per la crescita del Paese), convertito, con modificazioni, dalla legge 7 agosto 2012, n. 134.

Deliberazione della Giunta provinciale n. 133/2014 *“Piano di prevenzione della corruzione 2014-2016 della Provincia autonoma di Trento”*: la deliberazione approva il Piano triennale di prevenzione della corruzione, come da disposizioni nazionali contenute nella L. 190/2012 e recepite in Trentino dalla legge provinciale n. 9/2013. L’art. 19 dispone un coordinamento pro attivo tra il Responsabile anticorruzione, il Responsabile per la trasparenza e il Progetto Open Data in Trentino realizzato tramite una relazione annuale della struttura competente in materia di open data sullo stato di avanzamento del progetto, da inoltrare alle strutture competenti in materia di trasparenza e anticorruzione.

Legge provinciale 4/2014 “Disposizioni riguardanti gli obblighi di pubblicità, trasparenza e diffusione di informazioni e modificazione delle legge provinciale 28 marzo 2013, n. 5”: disciplina gli obblighi di trasparenza concernenti l'organizzazione e l'attività della Provincia, delle sue agenzie e dei suoi enti strumentali pubblici e privati previsti dall'articolo 33 della legge provinciale 3 giugno 2006, n. 3. La legge recepisce e sistematizza gli obblighi di pubblicità, trasparenza e diffusione dei dati, dei documenti e delle informazioni individuati dal decreto legislativo 14 marzo 2013, n. 33 (Riordino della disciplina riguardante gli obblighi di pubblicità, trasparenza e diffusione di informazioni da parte delle pubbliche amministrazioni). Tutti i dati pubblicati tramite i siti istituzionali in ottemperanza alla presente legge dovranno essere pubblicati come dati aperti.

3.1.4 Normative di contesto

Infine, è necessario tenere in considerazione molte normative che disciplinano la circolazione dell'informazione in modo da tutelare diritti soggettivi e interessi legittimi dei cittadini nei rapporti tra questi e gli enti pubblici e con la società. Oltre al regime dell'accesso amministrativo che resta materia differente ma che è utile valutare alla luce delle novità introdotta in seguito alla facoltà di riutilizzo in forma massiva e per scopi anche commerciali dell'informazione del settore pubblico, è fondamentale tenere in considerazione le problematiche inerenti la protezione dei dati personali e la gestione dei diritti di proprietà intellettuale, così come regolati dalle specifiche normative di settore.

Italia	Trentino
<u>Legge 633/1941</u> - Protezione del diritto d'autore e di altri diritti connessi al suo esercizio	
<u>Legge 241/1990</u> - “Nuove norme in materia di procedimento amministrativo e di diritto di accesso ai documenti amministrativi”	<u>Legge provinciale 23/1992</u> - “Principi per la democratizzazione, la semplificazione e la partecipazione all'azione amministrativa provinciale e norme in materia di procedimento amministrativo” <u>Decreto del Presidente della Provincia 17-97/Leg</u> “Regolamento per la disciplina delle modalità di esercizio e dei casi di esclusione del diritto di accesso ai documenti amministrativi”
<u>Decreto Legislativo 322/1989</u> . “Norme sul Sistema statistico nazionale e sulla riorganizzazione dell'Istituto nazionale di statistica”	
<u>Decreto Legislativo 196/2003</u> - “Codice in	<u>Deliberazione della Giunta provinciale n. 1081/2013</u> “D.lgs. n. 196 del 30 giugno 2003 (Codice in materia di protezione dei dati personali). Approvazione nuove disposizioni attuative della disciplina

materia di protezione dei dati personali”	della protezione dei dati personali”
Decreto Legislativo 163/2006 “Codice dei contratti pubblici relativi a lavori, servizi e forniture in attuazione delle direttive 2004/17/CE e 2004/18/CE”	

Italia

Legge 633/1941, “Protezione del diritto d’autore e di altri diritti connessi al suo esercizio”, normativa salvaguardata dall’articolo 4 della normativa sul riutilizzo dell’informazione pubblica (D. Lgs. 36/2006), a tutela del diritto d’autore e del diritto sui generis sulle banche dati, sulle opere dell’ingegno e sugli altri diritti connessi.

Legge 241/1990, “Nuove norme in materia di procedimento amministrativo e di diritto di accesso ai documenti amministrativi”: norma fondamentale per il funzionamento del sistema pubblico italiano, rileva in quanto fondante il diritto d’accesso, salvaguardato dall’articolo 4 della normativa sul riutilizzo dell’informazione pubblica (D. Lgs. 36/2006).

Decreto Legislativo 322/1989 - “Norme sul Sistema statistico nazionale e sulla riorganizzazione dell’Istituto nazionale di statistica.” Questa normativa è utile per individuare criteri di qualità dei dati statistici e per chiarire modalità e finalità del trattamento.

Decreto Legislativo 196/2003 “Codice in materia di protezione dei dati personali”. Codice italiano privacy, recepisce la direttiva 95/46/CE. Normativa salvaguardata dall’articolo 4 della normativa sul riutilizzo dell’informazione pubblica (D. Lgs. 36/2006), sia a tutela degli interessi privati dei soggetti a cui i dati personali si riferiscono, sia a fondamento della diffusione di dati personali di interesse pubblico.

Decreto Legislativo 163/2006 “Codice dei contratti pubblici relativi a lavori, servizi e forniture in attuazione delle direttive 2004/17/CE e 2004/18/CE”, normativa sui contratti pubblici per cui è prevista la pubblicazione come dati di tipo aperto.

Trentino

Legge provinciale 23/1992 “Principi per la democratizzazione, la semplificazione e la partecipazione all’azione amministrativa provinciale e norme in materia di procedimento amministrativo”, recepimento provinciale della Legge 241/90.

Decreto del Presidente della Provincia 5 luglio 2007, n. 17-97/Leg “Regolamento per la disciplina delle modalità di esercizio e dei casi di esclusione del diritto di accesso ai documenti amministrativi”, recepimento provinciale del regolamento attuativo della Legge 241/90.

Deliberazione della Giunta provinciale n. 1081/2013, “D.lgs. n. 196 del 30 giugno 2003 (Codice in materia di protezione dei dati personali)”, che sistematizza i concetti del corretto trattamento dei dati personali all’interno dell’organizzazione provinciale.

Si approfondiscono ora, in sinergia con le Linee Guida nazionali e le migliori pratiche europee, alcuni passaggi cruciali derivati dall’interpretazione del quadro normativo appena descritto. Le seguenti soluzioni sono declinate come passaggi operativi necessari alla pubblicazione dei dati nel Catalogo.

3.2 Titolarità dei dati e regimi di responsabilità

Fonti

- “I poteri e le facoltà connessi al riutilizzo spettano unicamente al titolare del dato” (art. 5, c. 5, D. Lgs. n. 36/2006)
- “titolare del dato” è “la pubblica amministrazione o l'organismo di diritto pubblico che ha originariamente formato per uso proprio o commissionato ad altro soggetto pubblico o privato il documento che rappresenta il dato o che ne ha la disponibilità” (art. 2, lett. i), D. Lgs. n. 36/2006)
- i riutilizzatori non diventano proprietari del dato, ma solo fruitori (art. 58, c. 1, D. Lgs. n. 82/2005)
- “la diffusione [di dati personali] da parte di un soggetto pubblico è ammessa unicamente quando è prevista da una norma di legge o di regolamento” (art. 19, c. 3, D. Lgs. 196/2003)
- la responsabilità dei dati pubblicati è delle singole strutture (art. 18, 1ter L.P. 7/1997 e della deliberazione della Giunta provinciale n. 1081/2013)

La Provincia Autonoma di Trento, le pubbliche amministrazioni e gli altri soggetti aderenti al catalogo dei dati aperti del Trentino provvedono alla pubblicazione per finalità di riutilizzo dei dati di cui sono titolari o di cui hanno la piena disponibilità in conformità alla normativa in materia di diffusione dell’informazione del settore pubblico (L.P. n. 16/2012, art. 9; Del. Giunta provinciale n. 2858/2012; D. Lgs. 36/2006 e s.m.i.; D. Lgs. n. 82/2005 e s.m.i.) e ad ulteriori normative di settore.

La pubblicazione di tali dati è accompagnata dalla loro esposizione sul catalogo dei dati aperti del Trentino.

I dati e i documenti che li rappresentano, resi disponibili come dati di tipo aperto ai sensi dell'art. 68 del D. Lgs. 82/2005, sono pubblicati dai soggetti che ne hanno la titolarità o la piena disponibilità, "in quanto i poteri e le facoltà connessi al riutilizzo spettano unicamente al titolare del dato" (art. 5, c. 5, D. Lgs. n. 36/2006).

A tali fini, "titolare del dato" è "la pubblica amministrazione o l'organismo di diritto pubblico che ha originariamente formato per uso proprio o commissionato ad altro soggetto pubblico o privato il documento che rappresenta il dato o che ne ha la disponibilità" (dall'art. 2, lett. i), D. Lgs. n. 36/2006 e s.m.i.); altre eventuali amministrazioni che utilizzino il dato non diventano titolari del dato medesimo, essendone soltanto fruitrici, ai sensi dell'art. 58, c. 1, D. Lgs. n. 82/2005. La pubblicazione di dati come dati di tipo aperto non modifica la titolarità dei dati rilasciati.

L'individuazione e la decisione circa la diffusione delle informazioni come dati di tipo aperto di cui è titolare la Provincia autonoma di Trento è stabilita in capo ai dirigenti generali e ai dirigenti di servizio della Provincia, in base alle competenze assegnate (cfr. art. 18, c.1 ter della L.P. n. 7/1997; Del. Giunta provinciale n. 1081/2013).

Nell'ipotesi in cui la titolarità sui dati sia condivisa tra diversi soggetti (pubblici o privati), la pubblicazione come dati di tipo aperto è consentita solo se chi la effettua ha la piena disponibilità dei dati o previo accordo tra i soggetti titolari in merito alla pubblicazione (in conformità alla Legge n. 633/1941 e s.m.i. "Legge sul Diritto d'Autore").

Nel caso in cui il dataset contenga dati personali, è necessario tener presente il ruolo che la Provincia, come Titolare del trattamento dei dati personali (ex art. 28, D. Lgs. 196/2003), ha affidato ai singoli Dirigenti tramite la preposizione di questi a Responsabili del trattamento dei dati personali relativi alle proprie funzioni. Essi dispongono, quindi, del trattamento di cui sono responsabili e ne verificano la correttezza in conformità con quanto stabilito dalla disciplina in materia di trattamento dei dati personali.

L'attività di pubblicazione **di dati personali** contenuti in dati di tipo aperto nel Catalogo integra un trattamento di diffusione di dati personali (art. 4, comma 1, lett. m, D.Lgs. 196/2003) e, in quanto tale, **è consentita solo se espressamente prevista da una norma di legge o di regolamento** (art. 19, comma 3, D.Lgs. 196/2003). In questa situazione è poi necessaria un'attenta analisi di contesto anche normativo, in particolare considerando quanto disposto dal Provvedimento del Garante per la protezione dei dati personali intitolato "Linee guida in materia di trattamento di dati personali, contenuti anche in atti e documenti amministrativi, effettuato per finalità di pubblicità e trasparenza sul web da soggetti pubblici e da altri enti obbligati" (doc. web 3134436, del 15 maggio 2014)³².

Per questi motivi, vista la complessità giuridica delle diverse situazioni prospettate e **l'elevato rischio - con conseguenti sanzioni - di trattamento illecito di dati personali** (art 167, D. Lgs. 196/2003), si consiglia ai Dirigenti responsabili di coinvolgere la struttura competente in materia di privacy.

I dati ed i documenti pubblicati come dati di tipo aperto e presenti sul Catalogo sono stati raccolti, trattati, pubblicati e aggiornati ai sensi degli artt. 50 ss. del D. Lgs. 82/2005 e di ulteriori disposizioni di legge individuate per specifiche tipologie di dati e sono pubblicati autonomamente dalla struttura

32 <http://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/3134436>

che ne ha la responsabilità (ai sensi dell'art. 18, 1ter L.P. 7/1997 e della Del. della Giunta provinciale n. 1081/2013). Ne consegue che la Provincia Autonoma di Trento non è responsabile della pubblicazione dei dati presenti sul catalogo dati.trentino.it di cui non abbia la titolarità o la piena disponibilità.

I dati sono messi a disposizione di chiunque voglia riutilizzarli per qualunque scopo.

Chi ne effettua il download manuale o tramite API³³, si assume la piena responsabilità sull'utilizzo dei dati e sulle conseguenze che possano derivarne a terzi. Pertanto, la Provincia Autonoma di Trento non sarà responsabile per eventuali rielaborazioni e riagggregazioni illecite di terzi riutilizzatori.

3.3 Modalità di individuazione dei dati oggetto di riutilizzo

Fonti

- entro il 31 marzo di ogni anno, ogni Dipartimento o struttura provinciale comunicherà il piano della pubblicazione dei dati alla struttura competente in materia di open data per pianificarne la pubblicazione sul catalogo dati.trentino.it. (Linee Guida per il riutilizzo e la diffusione dei dati pubblici - deliberazione della Giunta provinciale n. 2858/2012)
- alcuni vincoli impediscono la pubblicazione di dati per finalità di riutilizzo (D. Lgs. 36 del 2006, e s.m.i.)
- sono esclusi i dati presenti nell'Anagrafe Tributaria, e i documenti contenenti dati personali che non sono conoscibili da chiunque o la cui conoscibilità è subordinata al rispetto di determinati limiti e modalità, in base alle leggi, ai regolamenti e alla normativa dell'Unione Europea, nonché quelli il cui riutilizzo è incompatibile con gli scopi originari del trattamento (art. 11, D. Lgs. 30 giugno 2003, n. 196; D. Lgs. 36/2006)
- per agevolare la trasparenza amministrativa e rendere riutilizzabili il maggior numero possibile di dati, e per meglio garantire la protezione dei dati personali o coperti da segreto, la Provincia Autonoma di Trento favorirà la pubblicazione di dati aggregati o anonimizzati, associando agli stessi apposite licenze standard per il riutilizzo dei dati (art. 8 del D.Lgs 36/2006, Linee guida in materia di trattamento di dati personali, contenuti anche in atti e documenti amministrativi, effettuato per finalità di pubblicità e trasparenza sul web da soggetti pubblici e da altri enti obbligati, Doc. Web. 3134436).

33 API Application Programming Interface: in informatica, si indica ogni insieme di procedure disponibili per il programmatore, di solito raggruppate a formare un insieme di strumenti specifici, per realizzare un determinato compito all'interno di un certo programma. Permettono, dunque, la comunicazione i dati "a basso livello", cioè tra software che processano i dati per diverse finalità, mostrando all'utente solo il risultato necessario.

Ogni Ente titolare dei dati deve individuare i dati pubblici che intende pubblicare come dati di tipo aperto. Tale individuazione verrà compiuta in raccordo con la struttura competente in materia di open data ed in coerenza con le scadenze previste dalla Giunta provinciale per la programmazione di settore ICT, volta ad individuare gli interventi da realizzare nell'anno di riferimento, secondo quanto previsto dall'art. 21 della L.P. 16/2012.

Ogni Dipartimento o struttura provinciale comunicherà tali dati alla struttura competente in materia di open data per pianificarne la pubblicazione sul catalogo dati.trentino.it.

Gli obiettivi di accessibilità così individuati saranno inoltre pubblicati sul sito web della Provincia Autonoma di Trento entro il 31 marzo di ogni anno, in conformità a quanto disposto dall'art. 9, comma 7 del D.L. 18 ottobre 2012, n. 179, come convertito con Legge 17 dicembre 2012, n. 221.

In conformità a quanto stabilito nel D.Lgs. 36 del 2006, e s.m.i., **non saranno oggetto di riutilizzo i documenti e i dati detenuti per finalità che esulano dall'ambito dei compiti istituzionali della Pubblica Amministrazione, né i documenti, non contenenti dati pubblici, il cui accesso è disciplinato dalla legislazione sul procedimento amministrativo e sul diritto di accesso ai documenti amministrativi.**

Inoltre, non saranno rilasciati come dati di tipo aperto per il riutilizzo i dati che potrebbero violare:

- la sicurezza pubblica, la difesa nazionale, lo svolgimento di indagini penali o disciplinari;
- il diritto di terzi al segreto statistico, industriale e commerciale, o altri vincoli di segretezza fissati in obblighi di legge;
- i diritti di proprietà intellettuale di terze parti;
- il diritto alla protezione dei dati personali.

A quest'ultimo proposito, si ricorda che saranno esclusi i dati presenti nell'Anagrafe Tributaria, e i documenti contenenti dati personali che non sono conoscibili da chiunque o la cui conoscibilità è subordinata al rispetto di determinati limiti e modalità, in base alle leggi, ai regolamenti e alla normativa dell'Unione Europea, nonché quelli il cui riutilizzo è incompatibile con gli scopi originari del trattamento (in applicazione del D. Lgs. 30 giugno 2003, n. 196 - art. 11).

Inoltre, per agevolare la trasparenza amministrativa e rendere riutilizzabili il maggior numero possibile di dati, e per meglio garantire la protezione dei dati personali o coperti da segreto, la Provincia Autonoma di Trento favorirà, ove possibile, la pubblicazione di dati aggregati o anonimizzati, in modo da non consentire alcuna identificazione, nemmeno indiretta, dei soggetti a cui tali dati si riferiscono, coerentemente con la normativa vigente in materia. A tali documenti saranno associate apposite licenze standard per il riutilizzo dei dati, così come previsto dall'art. 8 del D.Lgs 36/2006, e in applicazione delle raccomandazioni del Garante per la Protezione dei Dati Personali.

Alcune categorie di dati particolarmente utili allo sviluppo di servizi sono disciplinati da direttive e leggi di settore che suggeriscono approcci più verticali all'apertura. E' il caso dei dati ambientali, dei

dati relativi ai servizi di trasporto intelligente e ai dati relativi ai beni culturali, così come disciplinati dalle normative di settore (si veda sintesi normativa al punto 4.1).

Periodicamente e con cadenza almeno annuale e stabilita in relazione alla tipologia di dati, gli enti e le strutture titolari dei dati **provvederanno all'aggiornamento dei dati e dei metadati già pubblicati** sul catalogo dati.trentino.it e disponibili per il riutilizzo.

La Provincia, inoltre, intende supportare sul piano formativo, giuridico e tecnologico, attraverso la struttura competente in materia di open data, gli enti che aderiscono al catalogo dati.trentino.it nell'intero processo di formazione dei dati, in modo da garantirne la coerenza con gli standard necessari alla loro piena fruibilità e apertura.

3.4 Licenze associate ai dati

Fonti

- “In alcuni casi i documenti sono riutilizzati senza che sia prevista una licenza; in altri, è rilasciata una licenza che impone al suo titolare condizioni di riutilizzo riguardanti questioni quali la responsabilità, l'uso corretto dei documenti, la garanzia di non alterazione e la citazione della fonte. Se gli enti pubblici autorizzano su licenza il riutilizzo di documenti, le relative condizioni dovrebbero essere eque e trasparenti. In tale contesto può rivelarsi importante anche la disponibilità online di licenze standard. Gli Stati membri dovrebbero pertanto provvedere a che siano disponibili licenze standard” (Considerando 17, Direttiva PSI (2003/98/CE))
- Il titolare del dato predispose le licenze standard per il riutilizzo in considerazione delle caratteristiche, nel rispetto degli eventuali diritti di terzi e della disciplina sulla protezione dei dati personali (art. 8, D. Lgs. 36/2006)
- “Le licenze per il riutilizzo dei dati pubblici e dei documenti contenenti dati pubblici, predisposte in conformità al decreto legislativo n. 36 del 2006, devono consentire la più ampia e libera utilizzazione gratuita, anche per fini commerciali” (art. 9, comma 5, L.P. 16/2012).

Le licenze per il riutilizzo definiscono le condizioni e le modalità di riutilizzo dei dati pubblici e dei documenti contenenti dati pubblici, consentendone la più ampia e libera utilizzazione gratuita, anche per fini commerciali, in conformità all'art. 8 del D. Lgs. 36/2006 e s.m.i., e all'art. 9 della L. P. 27 luglio 2012, n. 16.

Le licenze di riferimento individuate dal Progetto Open Data nel catalogo dati.trentino.it per il riutilizzo dei dati sono le seguenti:

- **Creative Commons Zero (CC0)**

- **Creative Commons Attribuzione (CC-BY) 4.0.**

Gli strumenti giuridici di rilascio **Creative Commons Zero** e **Creative Commons Attribuzione** sono stati individuati in ragione dell'ampiezza del riutilizzo concesso dalla licenza, anche per fini commerciali; del loro alto livello di interoperabilità con altri modelli di licenze standard; della loro facilità di comprensione e soprattutto diffusione fra il pubblico.

Le licenze Creative Commons si caratterizzano per un linguaggio semplice e facilmente comprensibile da parte degli utenti e garantiscono la redistribuibilità del dato, **nonché un livello di diffusione e di conoscenza a livello nazionale, europeo ed internazionale.**

Inoltre, grazie alla loro diffusione sulla rete Internet, anche nell'ambito di iniziative volte alla messa a disposizione di documenti e dati pubblici di altre amministrazioni in Italia, in Europa e nel mondo, rappresentano ad oggi **lo standard di fatto per la licenza di diritti di proprietà intellettuale online. La loro diffusione minimizza le barriere al riuso**, riducendo i costi per i riutilizzatori e massimizzando la probabilità che i dati dell'amministrazione provinciale possano essere combinati con altri dati pubblici e con dati generati dagli utenti della rete o da operatori privati.

Tale scelta si pone in continuità con quella fissata dalla deliberazione della Giunta provinciale n. 195 del 17 febbraio 2012, relativa all'apertura dei dati territoriali, i quali sono rilasciati con licenza CC0.

Le eventuali versioni successive verranno monitorate al fine di valutare adeguatamente l'eventuale aggiornamento.

A partire dall'approvazione delle presenti Linee Guida, verrà adottata la Licenza Creative Commons Attribuzione (CC-BY) 4.0 in sostituzione della precedente versione (2.5). I dati già pubblicati con la precedente versione non subiranno variazioni di licenza, salvo diversa decisione da parte del titolare del dato.

Al momento dell'individuazione del dataset da pubblicare in formato aperto, la struttura che pubblica il dato, in raccordo con la struttura provinciale competente in materia di open data, indicherà il tipo di licenza con cui verrà rilasciato, in conformità alle indicazioni riportate al paragrafo 3.4 di questo stesso documento.

Il riutilizzo presuppone l'accettazione della licenza associata al dato o al documento d'interesse. L'accettazione potrà essere implicita od esplicita, in relazione alla natura del dato o del documento, secondo quanto indicato sul Catalogo dei dati aperti del Trentino.

Conformemente a quanto disposto dall'art. 52, comma 2 del Codice dell'Amministrazione Digitale (c.d. Open by Default), i dati già pubblicati sui siti web della Provincia Autonoma di Trento, qualora non sia indicata una specifica licenza, si intendono rilasciati **“come dati di tipo aperto” e quindi sono disponibili per il riutilizzo “come se forniti di una licenza”** che in tal caso non può prevedere alcuna specificazione e di fatto il riutilizzo si allinea ad una licenza Licenza Creative Commons Zero (CC0) **ad eccezione dei casi in cui la pubblicazione riguardi dati personali.**

Qualora ricorrano **giustificati motivi** tali da orientare la scelta verso una licenza standard diversa da quelle di riferimento, la specifica licenza, per quanto possibile, **deve rispettare il principio di**

consentire la più ampia e libera utilizzazione gratuita, anche per fini commerciali e con finalità di lucro.

In particolare, e di concerto con la struttura competente, in caso **di presenza di dati personali** all'interno dei dataset, **devono essere predisposte apposite note di licenza utili al corretto trattamento** di questi dati.

Le altre Pubbliche Amministrazioni ed enti aderenti al catalogo dei dati aperti del Trentino potranno rilasciare i propri dati accompagnandoli eventualmente con altre licenze standard per il riutilizzo, dopo averne valutato l'opportunità e il livello di compatibilità con le licenze CC0 e CCBY 4.0 con la struttura provinciale competente in materia di open data.

Per attribuire una licenza ai dati da pubblicare è sufficiente citare la licenza (nome e versione) che si intende adottare nella pagina web in cui si pubblicano i dati, nel disclaimer del sito o in altra collocazione, ma comunque **in modo tale che sia chiaro per chi utilizza il dataset con quale licenza esso è distribuito**.

Per quanto concerne l'attribuzione, il licenziatario dovrà provvedere ad una menzione adeguata, rispetto al mezzo di comunicazione o supporto utilizzato, di:

- autore originale e/o titolare dei diritti;
- terze parti designate, se esistenti;
- la descrizione/titolo del documento o del dataset, se indicato dal licenziante;
- nella misura in cui ciò sia ragionevolmente possibile, l'Uniform Resource Identifier (URI) che il Licenziante specifichi dover essere associato con il documento oggetto di riutilizzo.

3.5 Le richieste di riutilizzo

Fonti

- gli enti predispongono canali per ricevere le richieste di riutilizzo dei dati: "i soggetti che intendono riutilizzare dati delle pubbliche amministrazioni o degli organismi di diritto pubblico presentano apposita richiesta secondo le modalità stabilite dal titolare del dato con proprio provvedimento". La risposta deve pervenire "entro il termine di trenta giorni, prorogabile di ulteriori trenta giorni nel caso in cui le richieste siano numerose o complesse" (art. 5, D. Lgs. 36/2006 e s.m.i.).

Con riferimento ai dati pubblici non ancora resi disponibili, i soggetti interessati al riutilizzo di dati e di documenti contenenti dati, diversi od ulteriori rispetto a quelli già licenziati come dati di tipo

aperto sul catalogo dei dati aperti del Trentino o su altri siti istituzionali provinciali, possono presentare formale richiesta di riutilizzo ai sensi dell'art. 5 del D. Lgs. 36 del 2006 e s.m.i..

Le richieste dovranno essere inviate **all'indirizzo e-mail info@dati.trentino.it** avvalendosi, eventualmente, del modulo scaricabile dalla medesima sezione del Catalogo.

Entro cinque giorni lavorativi la richiesta è reinoltrata alla struttura titolare dei dati e documenti richiesti.

La struttura titolare dei dati richiesti **comunica - via pec - l'avvio del procedimento** al richiedente **e per conoscenza a info@dati.trentino.it**.

La struttura titolare, previa consultazione, se necessaria, con il Team Open Data, **provvede ad accogliere o a rigettare la richiesta, nel termine di trenta giorni**, che potrà essere prorogato di ulteriori trenta giorni qualora le richieste siano numerose o complesse, così come previsto dall'art. 5 del D. Lgs. 36 del 2006.

In caso di accoglimento della richiesta di riutilizzo, la struttura titolare dei dati mette a disposizione i dati e i documenti sul catalogo dei dati aperti del Trentino e darà contestuale comunicazione al richiedente quei dati e/o documenti e al Team Open Data.

In caso di impossibilità ad accogliere la richiesta di riutilizzo, la struttura titolare, sentito anche il Team Open Data, **comunicherà la motivazione del rifiuto tramite pec al richiedente e per conoscenza al Team Open Data**.

3.6 Aspetti relativi al costo dei dati

Fonti

- “i dati sono resi disponibili gratuitamente oppure, qualora per il riutilizzo di documenti sia richiesto un corrispettivo, quest'ultimo e' limitato ai costi effettivi sostenuti per la loro riproduzione, messa a disposizione e divulgazione” (art. 7, c. 1, D.Lgs. 36/2006, come modificato da D.Lgs. 102/2015).
- “l'Agenzia per l'Italia digitale determina, su proposta motivata del titolare del dato, le tariffe standard corrispondenti ai costi effettivi previsti nel comma 1 e provvede alla pubblicazione delle stesse sul proprio sito istituzionale” (art. 7, c. 2, D.Lgs. 36/2006, come modificato da D. Lgs. 102/2015).
- “con decreto del Ministro dei beni e delle attività culturali e del turismo, di concerto con il Ministro dell'economia e delle finanze e sentita l'Agenzia per l'Italia digitale, da adottarsi entro il 15 settembre 2015, sono determinati i criteri generali per la determinazione delle tariffe e delle relative modalità di versamento” (art. 7, c. 4, D.Lgs. 36/2006, come modificato da D.Lgs. 102/2015).

Al fine di favorire la diffusione delle informazioni e di accelerare i processi di innovazione legati al riutilizzo di dati pubblici, la Provincia Autonoma di Trento intende mettere a disposizione i dati pubblici di cui è nella disponibilità, secondo il principio di gratuità e, pertanto, di norma, senza l'applicazione di tariffe o corrispettivi, come previsto dall'art. 9, c. 3 della legge provinciale 27 luglio 2012, n. 16 e dall'art. 7 del D.Lgs. 36/2006 e s.m.i.

In casi specifici ed eccezionali individuati dagli enti titolari dei dati, ed in presenza di costi effettivi legati alla riproduzione, messa a disposizione e divulgazione dei dati richiesti, e sentita l'Agenzia per l'Italia Digitale, potranno essere applicate specifiche tariffe determinate dagli enti stessi in base ad un principio di sostenibilità ed in conformità a quanto previsto dall'art. 7 del D.Lgs. 36/2006 e s.m.i.

Le eventuali tariffe saranno proposte dal titolare dei dati, valutate assieme alla struttura competente in materia di open Data, e definite in linea con quanto stabilito da regolamento nazionale³⁴.

3.7 Accordi in esclusiva e divieto di discriminazione

Fonti

- “tutti gli operatori potenziali sul mercato” devono poter riusare i documenti delle pubbliche amministrazioni e degli organismi di diritto pubblico. Gli accordi di esclusiva devono essere riesaminati periodicamente, almeno trimestralmente, e resi pubblici (art. 11, D.Lgs. 36/2006 e s.m.i.).

Gli accordi in esclusiva con gli operatori economici che immettono sul mercato prodotti a valore aggiunto basati sui dati oggetto di riutilizzo **sono ammessi solo ed esclusivamente quando l'accordo in questione risulti necessario per l'erogazione di un servizio di interesse pubblico**, secondo quanto previsto dall'art. 11 del D.Lgs. 36/2006 che **fissa il divieto di accordi di esclusiva**.

Il diritto di esclusiva eventualmente concesso dovrà comunque **essere oggetto di riesame con cadenza almeno triennale**.

34 Al momento della stesura di queste Linee guida, il regolamento è in elaborazione presso l'AgID. Questo documento definirà regole specifiche e parametri oggettivi per la definizione delle tariffe applicabili alle varie tipologie di dati, in linea con le previsioni normative in materia.

4. Modello operativo per la produzione e il riutilizzo dei dati

Il valore di un dato è **direttamente proporzionale alla sua qualità** ed è perciò necessario attivare i **processi** in grado di garantire questa qualità, attraverso **prassi sostenibili nel tempo dal momento in cui il dato si forma al momento in cui viene pubblicato e metadato, fino al momento in cui deve essere aggiornato**. In questo capitolo è definito il processo che la piattaforma dati.trentino.it ha sviluppato in modo allineato alle normative nazionali ed europee in materia di dati aperti per la valorizzazione del patrimonio informativo pubblico.

Passaggio chiave per realizzare questo processo è la **consapevolezza** da parte del servizio pubblico **del valore anche economico che hanno le informazioni che si producono negli uffici**. I processi individuati non sprecano questo valore e ne permettono la messa a disposizione a cittadini, imprese e al sistema della ricerca, dando impulso allo sviluppo di servizi innovativi.

L'efficacia del processo di apertura e valorizzazione del patrimonio informativo pubblico dipende **anche dal coinvolgimento diffuso - l'engagement - di soggetti esterni alla PA** e quindi dalla creazione di **un ambiente ecosistemico di laboratorio territoriale ad alta intensità di cultura del dato**.

Figura 3 Attori del laboratorio dell'ecosistema territoriale per la cultura del dato



Di seguito, dopo aver **definito cosa si intenda per qualità dei dati** e quale modello di certificazione il sistema trentino proponga alle strutture titolari dei dati, si descrive in dettaglio il modello operativo individuato per **garantire nel miglior modo possibile la qualità dei dati di chi aderisce al catalogo**

dati.trentino.it. Successivamente vengono fornite indicazioni affinché questa qualità possa essere raggiunta in **modo sistematico durante la fase di realizzazione dei sistemi** informativi, fase che nel settore pubblico si concretizza anche con la definizione di opportuni capitolati di gara. Infine, viene descritto **il modello organizzativo** individuato per sostenere tutto questo.

4.1 Cosa rende di qualità un dataset

La necessità di garantire attraverso standard internazionali la qualità dei dati è ribadita anche da AgID che, in particolare per **“l’elenco delle basi dati di interesse nazionale”**³⁵, ha promosso l’adozione dello **standard ISO/IEC 25012** che definisce un “Data Quality Model” generale, valido per i dati conservati in un formato strutturato, all’interno di un sistema informativo, pensato per essere interoperabile fuori dai propri contesti d’ufficio. Coerentemente con la certificazione ISO/IEC25012, AgID identifica un dato di qualità quando soddisfa i principi di **accuratezza, attualità, coerenza e completezza**, ovvero i seguenti requisiti³⁶:

- il dato ed i suoi attributi sono **accurati** quando **rappresentano correttamente ciò a cui si riferiscono**. Esistono molti esempi e/o “standard di fatto” per ogni oggetto o tratto si voglia rendere rappresentabile/rilevabile/misurabile: questi standard di solito hanno già affrontato problematiche relative all’accuratezza delle codifiche utilizzate per rappresentare qualcosa in modo tale che un computer possa elaborare quella informazione. **Si raccomanda**, in fase di definizione degli archivi e dei dataset, **di vagliare sempre se esistano standard o soluzioni già sperimentate su quanto si sta definendo per quei dataset**;
- il dato ed i suoi attributi si dicono **attuali o tempestivamente aggiornati** quando **riflettono la situazione corrente, con una frequenza di aggiornamento che dipende dal contesto cui si riferiscono**. I dati relativi ad alcune rilevazioni, si pensi al meteo, saranno attuali se in tempo reale; mentre altri dati potranno considerarsi attuali anche se aggiornati meno frequentemente, si pensi ai dati di un bilancio che hanno di solito aggiornamenti annuali o semestrali. In qualsiasi caso **il dataset è di qualità quando il dato viene aggiornato ogni volta vi siano variazioni rilevanti e quando viene rispettata la periodicità di aggiornamento dichiarata nei metadati**;

35 L’AgID con la determinazione commissariale (<http://tinyurl.com/p3gftgp>) ha definito un elenco di dati con particolare valore socio-economico e ha definito necessità prioritaria che questi siano coerenti con lo standard internazionale sulla qualità dei dati ISO/IEC 25012 "Data quality model". Si tratta nel dettaglio: l’anagrafe nazionale della popolazione residente; la banca dati nazionale dei contratti pubblici di cui all’articolo 62-bis; il casellario giudiziale; il registro delle imprese; gli archivi automatizzati in materia di immigrazione e di asilo. Questi dati riguardano informazioni necessarie a molte attività amministrative e quindi di fatto funzionali ad una interoperabilità dei sistemi, non si parla di open data, ma se per questi dati viene individuato uno standard di qualità **questo stesso è utile sia di riferimento per dati sul cui riutilizzo si vuole innestare nuovi modelli di business**, quindi attività economiche per le quali **ugualmente è prioritaria la garanzia della qualità**.

36 Agenda AgID e determinazione commissariale n. 68 /2013 <http://tinyurl.com/p3gftgp> e <http://tinyurl.com/nzvah6s>

- i dati e i suoi attributi si dicono **coerenti** quando **non presentano contraddittorietà rispetto ad altri dati, almeno nell'ambito dell'amministrazione che ne è titolare**. Nei cataloghi di dati aperti sono normalmente presenti più dataset su medesimi domini che, se pur coerenti all'interno dei singoli dataset, possono mostrare informazioni contraddittorie, o apparentemente tali, quando incrociate con quelle di altri dataset. Per la prima volta, recentemente, sono stati esposti pubblicamente, e quindi incrociati al di fuori dei luoghi di produzione di quei dati, alcuni dataset nazionali che hanno fatto emergere questa situazione ed hanno sollevato dibattito. Quello che si è appreso da questa esperienza è che **l'uso dei dati in contesti imprevisi** può fare emergere alcune loro problematiche: poiché aprire i dati significa **renderli disponibili per usi differenti** da quelli per cui sono stati raccolti **è molto probabile che usi non previsti** e nuovi incroci facciano **emergere situazioni da allineare**. Ciò evidenzia come il processo stesso di apertura dei dati sia **uno strumento attivo per aumentarne la qualità**. Sono gli stessi utilizzatori a segnalare problemi e a collaborare con il servizio pubblico, il quale **si impegna a rendere disponibili i propri dati garantendo il suo "best effort"** e a migliorare i propri archivi a seguito di segnalazioni. Un altro modo **per ridurre a monte questi errori** è l'utilizzo di ontologie e dizionari dei termini standard (cfr. paragrafo 5.4), utili, tra l'altro, **a ridurre la possibilità di incoerenze trasversali dovute ad esempio a interpretazioni diverse dello stesso termine**. Infatti basarsi **solo su controlli a valle in base all'uso che viene fatto delle informazioni** può comportare costi e conseguenze, a volte anche gravi, se si tratta di dati utilizzati in applicazioni critiche. In altre parole, il tema della coerenza interna dei dataset **è cruciale e va affrontato con pragmatismo** e capacità di valutare il rapporto costo/beneficio nel prevenire eventuali incoerenze anche sulla base della criticità del dataset in questione;
- un dataset risulta **completo** quando è **esaustivo per tutti i valori attesi e rispetto alle fonti che concorrono alla definizione del procedimento**. Nel caso di dati tabellari, ad esempio, si parla di completezza quando non sono presenti celle vuote (senza valore), che potrebbero essere dovute alla dimenticanza, all'errore nell'imputazione o semplicemente ad una documentazione superficiale del dato.

I requisiti sopra descritti trattano dimensioni di qualità **"interna ai singoli dataset"**, ma è altrettanto importante garantire una **"qualità esterna"** legata ai sistemi che producono, archiviano, inviano o "pescano" i dataset in modo da garantire nel tempo che le seguenti tre proprietà legate ai sistemi in sé siano soddisfatte:

- **disponibilità**: è importante poter garantire una volta reso disponibile un dataset un accesso continuo, o quando serve, a quei dati; è importante, in altri termini, che **i dati siano continuativamente disponibili** a chi li riutilizza secondo le modalità stabilite³⁷;
- **ripristinabilità**: è importante avere un piano che possa garantire un eventuale ripristino dell'integrità fisica e logica, laddove si verifici un incidente. Si tratta di definire le procedure

³⁷ Un riferimento sono le Linee Guida per il Disaster Recovery delle PA ai sensi del comma 3, lettera b) dell'art. 50-bis del DLgs. N. 82/2005 e s.m.i. e la circolare dicembre 2011, n. 58 DigitPA/AgID che introduce l'obbligo di stilare i piani di CO/DR e di aggiornarli almeno una volta all'anno.

e la soluzioni atte a garantire il recupero dell'informazione, qualora si generino delle interruzioni nella continuità del servizio;

- **accessibilità:** ove possibile è opportuno garantire l'accesso ai dati da parte di tutte le persone, incluse quelle svantaggiate.

Garantire la qualità dei sistemi che gestiscono i dati aperti vuole dire, in ultima analisi, diventare consapevoli che la decisione di sospendere, modificare o cancellare una pubblicazione di dati **può avere conseguenze critiche da non sottovalutare** per chi riutilizza quel dato.

Il riuso produce una **catena di valore** che si basa su una catena di servizi; il valore reale dei dati è determinato **da tutti gli anelli della catena dei servizi**. Sono quindi in questo senso **promosse tutte le azioni che spingano ad identificare delle forme nelle quali la qualità esterna connessa ai sistemi informativi sia garantita** tramite servizi **continuativi di qualità adeguata alla tipologia di dataset**.

Di seguito sono definiti il processo ed i relativi strumenti che il sistema trentino si è dato per garantire la qualità ai dati pubblicati nel catalogo dati.trentino.it.

Il primo passo è il **censimento** generale dei dataset potenzialmente disponibili all'interno dei sistemi informativi oggi gestiti dalla pubblica amministrazione; successivamente è necessaria la **verifica della loro legittimità, qualità e opportunità di pubblicazione** in modalità aperta; dopodiché si illustrano le azioni da mettere in pratica tramite il catalogo **affinché i dati del servizio pubblico siano facilmente riutilizzabili** da cittadini, organizzazioni e imprese; infine, si propone ai titolari dei dati la possibilità di riferirsi ad una **certificazione** internazionale di qualità.³⁸

4.1.1 Il catalogo dati.trentino.it e i certificati di qualità dei dati

La Provincia autonoma di Trento **promuove l'uso dei certificati di qualità dei dati ed in particolare dei certificati ODI - Open Data Institute**³⁹.

³⁸ Dalla consultazione pubblica di questo documento è emerso come sul **tema della qualità dei dati ogni ambito di attività abbia le sue peculiari criticità da controllare**. Un esempio in tale senso è stato proposto nel **settore dei beni culturali**, dove non è raro riscontrare disomogeneità nei criteri di catalogazione, errori di attribuzione, situazioni incerte riguardo alla titolarità del possesso delle opere. La natura densa e complessa di questi dati e i limiti delle risorse umane disponibili sono un ostacolo non da poco al rendere di pubblico dominio questo patrimonio. Si tratta di un settore che necessariamente utilizza prestatori d'opera ma raramente nei contratti si accenna alla possibilità per l'istituzione di disporre di licenze aperte per l'utilizzo della documentazione delle opere d'arte; la definizione di questi passaggi potrebbe essere un beneficio anche per gli stessi prestatori d'opera. In particolare, nel settore dei beni archeologici è in corso uno studio per proporre, appunto, delle specifiche linee guida di settore per la produzione di dati di qualità.

³⁹ L'Istituto è stato fondato nel 2012# da sir Tim Berners-Lee e da Nigel Shadbolt con l'obiettivo di promuovere globalmente la consapevolezza del valore che i dati hanno assunto per lo sviluppo economico di questo secolo anche attraverso la ricerca e lo definizione di pratiche da condividere. L'ODI si sta impegnando in modo specifico su vari aspetti che riguardano la qualità e la cultura del dato. L'istituto è rappresentato in

La **certificazione ODI**⁴⁰ si basa sul modello di qualità dei dati a “5 stelle” di Tim Berners Lee (cfr. paragrafo 5.1): essa permette, **attraverso una sequenza di domande**, di verificare **se le pratiche adottate** per aprire i dati siano volte a **garantire e sostenere nel tempo la loro qualità**, inoltre porta ad aggiungere informazioni **contestuali al dato molto utili per il riutilizzo**. In pratica **uno schema di validazione** che evidenzia quali dimensioni sia importante tenere sotto controllo nell’apertura dei dataset, offre consigli affinché siano prodotti dati più concretamente riutilizzabili da cittadini, organizzazioni e imprese. Seguendo queste indicazioni, si supera cioè la fase della “pubblicazione e rilascio”, indicando le caratteristiche che devono avere i dataset e in particolare i metadati, affinché tale rilascio sia utilizzabile massivamente⁴¹.

Oltre ad informazioni relative a **titolo, descrizione, chi pubblica il dato, URL e formato di rilascio** - informazioni o metadati di base già previste da AgID, dalla raccomandazione europea DCAT-AP, come anche da queste linee guida (cfr. capitolo 5) - la certificazione ODI introduce **altre tre dimensioni informative importanti**:

1. **una dimensione legale** dove la certificazione propone una scala di qualità a seconda del grado di complessità con cui vengono affrontate tematiche giuridiche. Per es. sul tema dell’anonimato, il titolare può scegliere una prima strategia di livello base **in cui non vengono pubblicati dati a rischio** ma può anche spingersi a livelli più elevati dove vengono individuati **processi di anonimizzazione specifici** per poter pubblicare in modo garantista dati altrimenti potenzialmente non in grado di tutelare il vincolo di anonimizzazione⁴²;
2. **una dimensione sociale** che individua e propone livelli diversi, sempre più profondi e articolati, **di interazione fra data producer e data user** in modo da spingere ad una sempre più stretta e pragmatica valorizzazione concreta del patrimonio informativo⁴³;
3. **una dimensione tecnica** che prevede diversi livelli di garanzia tecnologica a tutela delle altrettanto numerose dimensioni connesse alla qualità dei dati in generale.

Il sistema trentino **propone quindi i certificati ODI come ambiente di riferimento per sviluppare pratiche di produzione e pubblicazione dei dati aperti di qualità**.

Il sistema delle certificazioni lamenta, come esternalità negativa, il possibile rallentamento di alcuni processi, ma quanto indicato da un processo di certificazione può essere **utilizzato come modello a**

Italia dalla Fondazione Bruno Kessler.

40 Per ottenere la certificazione ODI ci si collega a questo link <http://certificates.theodi.org> e si effettua la registrazione. A questo punto ogni titolare dati può rispondere ad una sequenza di domande proposte che gli permettono di valutare la qualità dei propri dataset e di valutare le azioni necessarie per poterli migliorare.

41 Per approfondire questi temi si vedano le slide presentate a marzo 2015 da Ulrich Atz di ODI all’incontro dei mercoledì di NEXA sulla qualità dei data <http://theodi.github.io/presentations/training-law-best-practice.html#/22>

42 In Trentino il Progetto LOD4STAT della Provincia autonoma di Trento, sviluppato con l’Agenzia Statistica, ha lavorato in questa direzione per fornire un sistema automatico di anonimizzazione dei dati statistici;

43 Vanno, ad esempio, in questa direzione le attività organizzate dal Progetto Open Data per le PMI locali e le azioni coordinate con gli attori dell’ecosistema locale che a vario titolo interagiscono con le PMI locali con approccio di innovazione aperto.

cui tendere: un buon punto di riferimento per operare scelte e per garantire **la qualità dei dati**. La scelta in merito è compito del **titolare dei dati** come unico attore in grado di valutare, stimare l'eventuale valore in termini di domanda di riutilizzo dei propri dataset e decidere agli eventuali livelli di qualità del dato sia opportuno perseguire.

4.2 Modello operativo per la piattaforma dati.trentino.it

Entro il 31 marzo di ogni anno ogni titolare di dati comunica alla struttura competente per il Progetto Open Data in Trentino l'elenco dei dataset da valutare in merito alla pubblicazione. Sono consigliati come criteri di scelta: l'obbligatorietà, la disponibilità dei dati già in formato aperto perché oggetto di specifici progetti, la domanda reale o prevista sia da parte delle imprese e cittadini che da parte di altri soggetti della pubblica amministrazione, la pubblicazione nei portali.

A supporto di tale attività, di seguito si prefigurano le operazioni che si dovranno realizzare per pubblicare i dati, così da facilitare la redazione di un piano di attività annuale.

La **Tabella 2** riassume la sequenza di azioni nel tempo che conducono alla valorizzazione del patrimonio informativo pubblico, **anche oltre la pubblicazione dei dati in modalità aperta**.

Tali azioni rappresentano le attività che una struttura farà per giungere a pubblicare i propri dati, ma soprattutto per avere **piena consapevolezza del patrimonio informativo pubblico di cui dispone**.

Un processo, quindi, di **valorizzazione degli archivi del servizio pubblico definito** in tre importanti articoli del CAD: **l'articolo 50, 52, 58** sugli open data e **l'articolo 68**, i quali **assieme descrivono le norme per rendere interoperabili le banche dati (riutilizzo interno fra PA) e l'apertura dei dati (riutilizzo esterno)**.

Per questo motivo la prima azione si colloca a livello di sistemi informativi in generale ed è tesa a rilevare **il valore che le pubbliche amministrazioni hanno nei propri archivi, prima ancora di valutarne l'eventuale pubblicazione in modalità open data**. Porre attenzione al processo di produzione, conservazione e utilizzo dentro e fuori il sistema pubblico aumenta **la qualità generale dei dati, ne agevola il riutilizzo da parte del mercato, cittadini e mondo della ricerca**, ed è il presupposto di base anche per un **riutilizzo interno delle informazioni fra le pubbliche amministrazioni**.

Tabella 2 Che fare, dove operare e con che strumenti, per garantire standard di qualità nei dataset

<i>Dove</i>	<i>Sui sistemi/archivi</i>	<i>Sui singoli dataset</i>	<i>Sul catalogo</i>	<i>Sul catalogo (ricorsivamente)</i>	<i>In rete</i>
<i>Cosa</i>	<i>Censimento Sorgenti dati</i>				
		<i>Analisi giuridica</i>			
		<i>Analisi qualità/costi benefici e usi conosciuti</i>			
		<i>Strutturazione del dato/ Bonifica</i>			
		<i>Politiche di accesso e licenza</i>			
			<i>Pubblicazione in open data</i>		
			<i>Metadatazione</i>		
			<i>Assegnazione licenza uso</i>		
				<i>Gestione e mantenimento</i>	
					<i>Certificazione dato (ODI)</i>
				<i>Rilevare nel tempo se possibile il riutilizzo</i>	
<i>Con che cosa</i>	<i>Scheda rilevazione basi dati</i>	<i>Checklist</i>	<i>Catalogo dati.trentino.it</i>	<i>Catalogo dati.trentino.it</i>	--

Questo schema è una sintesi dell'esperienza maturata dalle strutture che hanno aderito al catalogo dati.trentino.it. Nel caso **di dati aperti attraverso servizi web**⁴⁴, le fasi di questo modello possono risultare aggregate e leggermente diverse, ma la definizione di questo processo è utile per individuare **passaggi essenziali al fine di arrivare ad una reale apertura dei dati**. Lo schema si allinea a quanto proposto anche dalle linee guida AgID. Si definisce la catena di azioni necessarie affinché i dati **possano passare da un uso meramente interno**, a volte anche non oltre il singolo ufficio o struttura di una stessa amministrazione, ad **essere interscambiabili, certi della loro qualità, con altre amministrazioni** e finalmente essere **pubblicati in modalità open data perché legittimi, strutturati, licenziati e metadati**.

44 Un sistema software creato per rendere interoperabili diversi sistemi di elaborazione all'interno di un contesto distribuito più o meno vasto, per es. una rete interna alla pubblica amministrazione o la rete internet stessa.

4.2.1 Censimento delle basi dati

Un censimento dei dati disponibili è necessario per valutare lo stato del “patrimonio informativo pubblico” di cui una struttura è titolare e per potere quindi valutare se e come rendere aperti quei dataset, seguendo anche quanto indicato dalle Linee guida AgID:

“Il percorso [...] parte quindi dal censimento dei dati esistenti attraverso, ove possibile, un obiettivo aziendale nell’ambito della programmazione strategica, condiviso da tutta l’amministrazione. Si raccomanda pertanto al responsabile Open Data di effettuare una ricognizione interna, alla luce della normativa vigente, in collaborazione con i responsabili delle basi di dati, al fine di individuare l’insieme di dati esistenti pubblicabili in formato aperto.”⁴⁵

Si tratta di una attività tesa a rendere esplicite dimensioni quali una possibile domanda di riutilizzo, eventuali obblighi di legge, eventuali costi - in termini di tempo o di servizi - compresi quelli di mantenimento e di ripristino eventuale in caso di problematiche tecnologiche rilevanti.

Per questa sua **funzione necessaria e prioritaria**, AgID ha già realizzato **la prima rilevazione delle basi dati delle pubbliche amministrazioni nazionali**⁴⁶, rilevazione che il sistema trentino sta utilizzando come punto di partenza e di riferimento per la “fase di censimento”.

Sia a livello micro di singola struttura amministrativa, sia a livello macro delle amministrazioni centrali il primo passo necessario è proprio quello della **valutazione sullo stato dell’arte degli archivi già disponibili**. Il Progetto Open Data metterà a disposizione un supporto per rilevare: il **processo di produzione interno delle basi dati o di estrazioni di queste** così da registrare gli attori, enti e/o società coinvolti nella produzione degli archivi; **dove il dato viene memorizzato** e quale sia la sua **destinazione**, ovvero se vi siano obblighi o richieste di trasmissione da parte di istituzioni centrali, di altre strutture interne o anche semplicemente un utilizzo da parte dei titolari per finalità di comunicazione od altro, così da **permettere il monitoraggio delle reti relazionali fra diversi livelli di governo e di trasmissione dei dati**. Inoltre è auspicabile che tali informazioni **siano integrate con gli schemi (logici e/o concettuali) dei dati utilizzati e i vocabolari impiegati**, quando esistenti, si tratta di una informazione preziosa per affrontare il problema dell’integrazione semantica delle varie basi di dati della PA.

45 <http://www.agid.gov.it/agenda-digitale/open-data/dati-pubblici-condivisione>.

46 Questo il link alla pagina del lavoro effettuato da AgID sulle Basi Dati della pubblica amministrazione <http://www.agid.gov.it/agenda-digitale/open-data/basi-dati-pa>

4.2.2 Analisi Giuridica

Con l'analisi giuridica si entra nel vivo delle **ricadute in termini legali** derivanti dal rendere **disponibili e facilmente elaborabili le informazioni prodotte dal servizio pubblico in modo massivo tramite i computer**. Non tutti i dati generati dalla pubblica amministrazione possono infatti essere pubblicati come open data. Si sono approfonditi tali profili nel capitolo 3 a proposito dei vincoli di privacy, di sicurezza, di segretezza. **I titolari dei dati**, secondo quanto indicato dall'Autorità garante della Privacy, supportati dalla struttura competente in materia, **pubblicheranno solo dati legittimi e andranno a definire il tipo di licenza per il riutilizzo**.

4.2.3 Analisi qualità, valutazioni costi/benefici, azioni di bonifica

L'analisi della qualità sui dataset che si considera legittimo pubblicare in modalità open è tesa ad individuare una serie di problematiche che spesso penalizzano i sistemi informativi, che oggi con la pubblicazione in modalità open si rendono più evidenti. Attorno a questi temi la piattaforma dati.trentino.it sta investendo per supportare nei prossimi anni un sistema che sia consapevole del valore dei dati, dei modi per tutelarli e di come meglio metterli a disposizione per la crescita.

*“La fase di valutazione della qualità dei dati è importante in tutti i sistemi informativi **indipendentemente dall'apertura dei dati**. Con l'adozione di politiche di apertura dei dati, la qualità dei dati assume un ruolo ancora più rilevante in quanto elemento per la certificazione della bontà dei dati forniti e soprattutto dell'appropriatezza rispetto all'utilizzo che del dato si vuole fare.”⁴⁷*

Spesso i dati prodotti dal servizio pubblico sono creati **con finalità di gestione con riutilizzo per lo più interno con modalità “caso per caso”**, attento a coprire il dettaglio del singolo caso - la specifica domanda di contributo da parte di un cittadino, la storia professionale di quel particolare dipendente, ecc. - più che la confrontabilità con altri dati anche all'interno di uno stesso dataset.

Ma “dare valore ai dati” significa poter **riutilizzare quelle informazioni anche oltre la registrazione e il controllo amministrativo del singolo caso**, pensando invece a collegare/incrociare quelle informazioni con altri dati. Inoltre confrontando fra loro dataset provenienti da titolari diversi, si riscontra una **grande varietà di modelli utilizzati** per compilare le tabelle: diversi i formati per le date, per gli indirizzi, per il telefono telefono ecc. Se la qualità è direttamente collegata alla

47 Linee Guida AgID 2014.

standardizzazione di codifiche, formati e modelli, la mancanza di standardizzazione rende complicata la comprensione da parte dei computer. Anche **una documentazione inesistente o troppo autoreferenziale** e il **non aggiornamento** dei dati nel tempo o rispetto ai cambiamenti del fenomeno rappresentato abbassano la qualità dei dati.

La **Tabella 3** descrive **alcuni degli errori nei dataset che maggiormente si rilevano come esperienza comune fra chi gestisce dataset**, raggruppati a **partire da quattro principali fonti di dati**: quelli prodotti manualmente nei propri uffici (tabelle di calcolo, archivi relazionali autoprodotti, ecc.); quelli prodotti dai sistemi informativi; quelli prodotti da sensori e infine i dati prodotti dai sistemi di CMS dei siti web, e per ognuna di queste si propongono alcune possibili azioni da intraprendere per prevenire e/o bonificare i dataset.

Tabella 3 Principali fonti di dati con gli errori maggiormente riscontrati e le possibili azioni da intraprendere per correggerli o evitarli alla fonte.

Tipologia dei dati in base all'origine	Tipologia di errori maggiormente riscontrati	Possibili azioni di bonifica
Prodotti manualmente	Errori input, incompletezza, inconsistenza nel tempo delle codifiche, duplicazione dati, problemi di aggiornamento, presenza di conoscenze tacite difficilmente comprensibili all'esterno	Codifiche internazionali, standard di imputazione per date, indirizzi, ecc., riferimenti ad ontologie, coordinate geografiche per localizzare eventi e oggetti, datazione della produzione/aggiornamento del dato
Prodotti dai sistemi informativi	Formati che non fanno riferimento a standard, codici identificativi autoreferenziali senza attenzione a codici nazionali e standard, limitata metadatazione, scarsa diffusione di servizio di interoperabilità	
Prodotti da sensori	Problemi di trasmissione/raccolta dati influenzati da eventi atmosferici (es. ghiaccio che blocca i sensori o nevicatae abbondanti che coprono i sensori), mancanza rete elettrica, deriva lenta del sensore, malfunzionamento del datalogger	Alert disfunzioni, utilizzo di standard per l'interoperabilità applicativa e la raccolta di dati in modo uniforme (es. protocollo Sensor Observation Service) Bonifica sugli storici,
Prodotti nei portali dai CMS	Mancanza di schemi che strutturino le informazioni - non aggiornamento - quando strutturati mancanza di vocabolari condivisi fra attori diversi	Utilizzare un CMS/DMS che strutturi in classi gli oggetti che vengono comunicati attraverso i portali (es. notizia, evento,

		luogo della cultura, progetto, atto normativo ... ⁴⁸⁾
--	--	--

Una volta **individuate le criticità**, possono essere avviate azioni più o meno impegnative di **bonifica** del dato, svolte dal titolare con il supporto del Progetto Open Data. E' il titolare, infatti a poter/dover valutare il rapporto tra costi e benefici di questa attività. Corretti i singoli dataset è però importante **intervenire sui processi di produzione**, analizzando le cause che creano errori ed intervenendo così a livello strutturale al fine di garantirne la qualità nel tempo. In questo senso è da interpretare quanto AgID afferma in merito al rapporto fra valutazione del costo necessario alla bonifica dei dati e loro valore oltre uno scenario di riuso immediato⁴⁹.

“L’apertura dei dati può essere uno stimolo importante per la conduzione di attività mirate di bonifica. Oltre all’utilizzo del dato all’interno dell’amministrazione, infatti, si possono aprire nuovi scenari di utilizzo che giustificano maggiormente gli investimenti connessi al processo di bonifica.”⁵⁰

4.2.4 Strutturazione del dato/bonifica

Una volta valutato che il dato può aumentare la sua qualità e quindi il suo valore attraverso delle modifiche che lo rendano riutilizzabile oltre il suo scopo originario, si tratta di apprendere alcune pratiche di base e avviare dei processi di “pulizia” più o meno automatizzati⁵¹. I computer hanno bisogno di **dati strutturati con indicazioni semplici, a volte ripetitive e soprattutto standardizzate**. Trasformare tabelle create per una lettura a misura d’uomo in tabelle utili ad una lettura da parte dei computer è una delle attività principali dell’avvio del processo di apertura dati che crea un certo impatto sul lavoro degli uffici della PA, per facilitare queste attività il Progetto Open Data in Trentino ha raccolto indicazioni e precauzioni nelle “Regole di scrittura dati”⁵² e nel “Tutorial CSV”. Nel caso invece in cui i dati siano collegati al catalogo dati.trentino.it attraverso un servizio web, si presuppone che l’attività di strutturazione/bonifica sia stata realizzata precedentemente, nel momento della realizzazione della base dati dalla quale il servizio web attinge i dati.

48 Si veda a questo proposito il progetto ComunWeb del Consorzio dei Comuni Trentini

49 Napolitano - Un calcolo per l’opendata (de.straba.us/2015/02/18/un_calcolo_per_lopendata/)

50 ibidem

51 un tool molto utilizzato a tale fine è Open Refine (<http://openrefine.org/>), ma diverse operazioni di pulizia possono essere fatte tramite gli strumenti di Open Office.

52 Regole di scrittura dati: <http://tinyurl.com/og6rcue>

4.2.5 Politiche di accesso e licenze d'uso

In questo passaggio si definisce se il dato che è stato bonificato, ed è ora quindi interoperabile all'interno delle strutture del servizio pubblico, possa essere reso anche fruibile in modalità open data. Inoltre viene decisa quale licenza sia opportuno assegnare in modo da disciplinarne il riutilizzo.

“Altro aspetto importante da considerare sono eventuali forme di aggregazione dei dati e restrizioni di accesso, che hanno anche un impatto sulla scelta della licenza [...]. Sebbene sia sconsigliato restringere l'accesso ai dati o procedere con la pubblicazione di aggregazioni degli stessi, esistono tuttavia casi in cui i dati possono essere diffusi (pensiamo ad esempio ai redditi) solo in forma anonima, ossia ad un livello di aggregazione tale da impedire di identificare le persone cui i dati si riferiscono. A tal fine, in questa fase devono essere definite delle adeguate politiche di accesso ai dati in cui sia indicato un profilo di accesso specifico per ogni dato, dettato dai diritti sull'informazione di base, dalle norme o dalle policy in atto.”⁵³

4.2.6 Pubblicazione in open data

Per la **pubblicazione degli open data** la Provincia ha messo a disposizione del sistema territoriale il catalogo dei dati aperti dati.trentino.it. Esso garantisce standard condivisi ed economie di scala per affrontare i cambiamenti organizzativi e di saperi derivanti da queste nuove attività e le funzioni che spettano al servizio pubblico. Per non disperdere i saperi legati ai dataset è opportuno che i dati aperti siano il più vicino possibile alla struttura che ne è titolare. In tal senso, la pubblicazione dei dati nel catalogo deve avvenire **preferibilmente attraverso link o processi automatici di harvesting (raccolta) agganciati ai processi**, organizzativi e/o informatici, **che li generano**.

Dati.trentino.it si configura, dunque, prevalentemente come un catalogo di metadati che descrive, standardizza, fornisce API e facilita quindi l'accesso di cittadini, organizzazioni e aziende ai dati di loro interesse **che sono tendenzialmente caricati nei siti delle strutture o nei sistemi informativi centrali del SINET**.

53 ibidem

4.2.7 Metadattazione

Con **metadattazione** ci si riferisce alla produzione di informazioni volte a descrivere un insieme di dati, così da rendere intelligibile il loro significato a chiunque e semplificandone la ricerca e il riutilizzo. **I metadati danno certezza al dato, al processo con cui è stato creato, alla sua validità e qualità.** Il catalogo dati.trentino.it fornisce uno schema minimo di metadati⁵⁴ a compilazione obbligatoria i cui dettagli sono forniti di seguito in questo capitolo. Per facilitare una corretta compilazione dei metadati nel catalogo dati.trentino.it è stato prodotto un video esplicativo⁵⁵.

Anche il gruppo open data formato da AgID sta lavorando sul tema della metadattazione di base e il catalogo dati.trentino.it è allineato al dibattito nazionale in corso sugli standard AgID e DCAT-AP.

Fra i metadati obbligatori vi è anche la descrizione del **tipo di licenza di riutilizzo** che in questo passaggio viene associata pubblicamente al dataset aperto nel catalogo. Si noti bene che qui si sta parlando di metadati associati a dati in modo esterno utilizzando il catalogo dati.trentino.it, è importante ricordare che, **nel caso di dati aperti utilizzando servizi web**, i metadati sono inseriti nel servizio web stesso e quindi di fatto **“nascono insieme ai dataset”** (cfr. paragrafo 5.4).

4.2.8 Gestione e mantenimento del dato

Il dato ed i metadati ad esso associati devono essere gestiti e documentati in **modo costante anche dopo la loro pubblicazione** questo per garantire un'adeguata qualità **sia in termini di valori che di formati che di sistemi informativi.**

La questione della sostenibilità dell'aggiornamento dei dati impatta sullo sviluppo economico, in termini di crescita che gli open data del servizio pubblico possono creare. Infatti, **i servizi innovativi** che il mercato può proporre ed autofinanziare attraverso modelli di business adeguati **necessitano di una continuità non episodica** nella fornitura delle informazioni che sono utilizzate per questi servizi. Si tratta, in altre parole, di **superare una fase di “lancio”, “sperimentazione”** di nuovi servizi e di **garantire un flusso costante e aggiornato di quei dati che sono di maggior interesse per il mercato e la ricerca.**

54 Al seguente link è disponibile lo schema minimo dei metadati presente nella versione precedente delle Linee Guida per gli open data adottate dalla PAT, <http://tinyurl.com/ouu474u>

55 <http://tinyurl.com/pu3fq3r>

4.3 Cosa vuol dire riutilizzo dei dati

L'obiettivo del processo di apertura del patrimonio informativo pubblico è la valorizzazione dei dati attraverso il loro riutilizzo da parte di soggetti esterni alla pubblica amministrazione, anche per scopi commerciali.

Le Direttive 2003/98/CE⁵⁶ e 2013/37/UE⁵⁷ individuano nel patrimonio informativo pubblico un'importante materia prima per prodotti e servizi imperniati sui contenuti digitali, il cui riutilizzo può contribuire alla crescita economica e alla creazione di nuove opportunità occupazionali. A tal fine le pubbliche amministrazioni devono provvedere (art. 50 CAD) affinché i dati ed i documenti che sono prodotti o che sono raccolti - attività fisiologica della PA - in modo riutilizzabile. **Per riutilizzo la normativa intende** l'uso del dato di titolarità di enti ed organismi di diritto pubblico da parte di persone fisiche o giuridiche, **a fini commerciali o non, diverso dallo scopo iniziale per il quale il dato è stato prodotto** nell'ambito dei fini istituzionali della pubblica amministrazione.

La Provincia autonoma di Trento, con la legge finanziaria 2012⁵⁸ ed il Piano di miglioramento 2012-2016, ha individuato nella messa a disposizione del patrimonio informativo uno strumento **per abilitare l'innovazione di prodotti e servizi, supportare la crescita economica e migliorare l'efficienza della pubblica amministrazione**, in linea con quanto previsto dalla direttiva EU: *“Per facilitare il riutilizzo gli enti pubblici dovrebbero, ove possibile e opportuno, mettere i loro documenti a disposizione, tramite formati aperti e leggibili meccanicamente, insieme ai rispettivi metadati, al miglior livello di precisione e di granularità, in un formato che garantisce l'interoperabilità”* (Direttiva 2013/37/UE).

Il compito della pubblica amministrazione non si limita però al mero rilascio del proprio patrimonio informativo, ma è anche quello di garantire - in un'ottica collaborativa e sussidiaria rispetto al sistema territoriale - processi amministrativi in grado di ridurre le barriere che ostacolano lo sviluppo di iniziative orientate al riutilizzo del dato pubblico.

In un processo sussidiario, **“spettano al settore pubblico** tutte quelle azioni che coinvolgono la creazione, la certificazione, l'aggiornamento, e più in generale la qualità del dato, così come è di competenza del pubblico soddisfare la domanda di innovazione sociale”, mentre **“spetta al privato la parte di uso e la creazione di plusvalore sul dato** [...] producendo servizi per la collettività a partire dai dati di tipo aperto messi a disposizione”⁵⁹.

E' chiaro che nella creazione di un sistema di valorizzazione dei dati del servizio pubblico esteso a tutto il sistema territoriale i compiti che spettano alla Provincia (e in generale alla pubblica amministrazione in forma trasversale) sono prevalentemente quelli **di aver cura e di valorizzare la qualità dei dati messi a disposizione**. Sarà poi la collaborazione fra aziende e sistema della ricerca ad individuare i servizi innovativi capaci di fare breccia all'interno di un nuovo mercato che vede nei dati la materia prima originaria.

56 <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2003:345:0090:0096:IT:PDF>

57 <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:175:0001:0008:IT:PDF>

58 Legge Finanziaria 2012 della Provincia autonoma di Trento, <http://tinyurl.com/ok9lwda>

59 ibidem

4.4 Indicazioni generali per la realizzazione dei sistemi informativi

E' opportuno prendere adeguatamente in considerazione il tema del riutilizzo dei dati **durante il processo di realizzazione e gestione dei sistemi informativi, in modo tale da garantire la loro sostenibilità con costi marginali bassi e in collegamento con il loro valore e il loro possibile riutilizzo**⁶⁰. Gli stessi gestori dei sistemi informativi del servizio pubblico sono oggi sempre più consapevoli del valore di quanto stanno gestendo: in questi anni la loro professione ha subito un graduale mutamento e dall'offerta di risposte orientate alla gestione, si è spostata alla messa in esercizio di soluzioni collegate alla programmazione e alla governance. Sempre più il valore da salvaguardare coincide con la ricchezza delle informazioni e quindi nella conoscenza a disposizione della pubblica amministrazione, e sempre meno con la mera gestione immediata. Una simile conoscenza si sviluppa completamente solo se contenuta in sistemi in grado di raccogliere, garantire e mettere a disposizione facilmente dati e informazioni di qualità, aggiornati, tempestivi, standardizzati ed interoperabili. Offrendo quindi la possibilità, ove si ritenga utile, di riutilizzo a fini commerciali fuori dalle strutture e dai contratti della pubblica amministrazione.

Secondo questo paradigma, in tutte le fasi di sviluppo e gestione dei sistemi **il valore da garantire è la capacità informativa del dato che viene raccolto, archiviato, incrociato con altri, arricchito, bonificato e connesso.**

Tali azioni possono essere, a vario livello, intraprese sia sui sistemi informativi in esercizio, agendo con delle ristrutturazioni puntuali, sia sui sistemi di nuova progettazione, prevedendone le opportune funzionalità di integrazione dei dati.

Si individuano in tale direzione **due momenti nella gestione dei sistemi informativi** in cui è possibile predisporre degli **interventi al fine di mettere a sistema** una serie di attività che danno valore al patrimonio informativo pubblico:

- la **fase di censimento e documentazione delle basi informative** e dei servizi informativi dei vari enti che aderiscono al sistema, in gestione presso e non solo gli enti pubblici e le società di sistema ad essi collegati (anche facendo riferimento a metodologie e buone pratiche adottate in tali settori, quali ad esempio ITIL⁶¹)
- le **fasi di analisi e progettazione dei sistemi informativi**, ove sono definite le specifiche tecniche che permettono la realizzazione dei sistemi informativi in collaborazione con i fornitori selezionati sul mercato con le procedure di procurement più adeguate.

60 Le linee guida open data di AgID, capitolo 11

61 <https://it.wikipedia.org/wiki/ITIL>

Inoltre costituisce un passaggio fondamentale l'attività di **georeferenziazione dei dati**, un passaggio molto importante per la valorizzazione dei dati, attribuendo in tale modo a un dato l'informazione relativa alla sua localizzazione geografica (tipicamente latitudine e longitudine secondo un particolare sistema geodetico di riferimento)

4.4.1 Censimento delle basi dati informative e dei servizi informativi degli enti

Come accennato nel paragrafo 4.2.1, nel 2014 l'AgID ha istituito per la prima volta il "Catalogo dei Dati della Pubblica Amministrazione". Ciò è avvenuto a seguito del processo di raccolta dell'elenco delle basi di dati effettuato in attuazione dell'art. 24-quater, comma 2, del D.L. 90/2014, convertito in Legge 11 agosto 2014, n. 114. L'elenco, disponibile nel sito **basidati.agid.gov.it** è uno strumento teso a facilitare la diffusione e la conoscenza dei dati della PA e a favorirne la condivisione secondo i principi dell'open data. Si tratta di un primo importante passo per creare sinergie, standard e codifiche condivise fra la pleora di banche dati del patrimonio informativo nazionale.

Anche il sistema pubblico trentino fornisce le proprie informazioni per questo censimento, con un processo interno ben definito **che garantisce il costante aggiornamento dell'elenco**; tale processo è il punto di riferimento di partenza per poter operare scelte per quanto riguarda i piani di apertura dei dati delle singole strutture, ma è anche un catalogo che **deve essere consultato nel momento in cui si sono progettati i nuovi sistemi informativi per tenere sotto controllo e operare opportune interazioni e sinergie**.

Il censimento può fare riferimento agli elenchi già esistenti per i sistemi informativi provinciali, tra i quali si indicano a titolo esemplificativo i seguenti:

- il sistema di gestione del SINET⁶²
- elenco basi di dati di Informatica Trentina⁶³ pubblicato da AgID
- elenco delle basi di dati della Provincia autonoma di Trento⁶⁴ pubblicato da AgID
- l'elenco degli applicativi in gestione e di proprietà di PAT⁶⁵

62 Delibera PAT n° 1458 del 7/7/2011

63 http://basidati.AgID.gov.it/catalogo/amm?code=its_022

64 http://basidati.AgID.gov.it/catalogo/amm?code=p_TN

65 Elenco applicativi in gestione e di proprietà PAT <http://tinyurl.com/q78nu4v>

4.4.2 Analisi e progettazione dei sistemi informativi

La pubblica amministrazione è impegnata in una attività di digitalizzazione e di generale ammodernamento della propria infrastruttura ICT (SINET) in collaborazione con Informatica Trentina, non solo con l'introduzione di nuove tecnologie ICT ma anche con il consolidamento delle infrastrutture informative attualmente disponibili in modo tale da aumentare la possibilità di riutilizzo dell'informazione e dei dati, sia verso l'esterno ma anche verso l'interno della amministrazione stessa.

E' importante che si preveda di affrontare gli aspetti relativi alla eventuale apertura dei dati, nelle prime fasi di progettazione di un nuovo sistema o di una sua evoluzione funzionale, **redigendo e pubblicando la documentazione relativa**, e lasciando a successive valutazioni (legali, di opportunità ed altro) la scelta di aprire o meno i dati. La documentazione dovrebbe includere non solo il nome ed una descrizione qualitativa della base informativa, ma prevedere anche informazioni di dettaglio quali: il modello dei dati (logico e/o concettuale) e i supporti tecnologici sui quali sono messi a disposizione, con i servizi di integrazione già eventualmente disponibili (servizi di interoperabilità applicativa). Inoltre, è auspicabile che siano anche documentati **i flussi coinvolti nella creazione della base informativa, sia in ingresso che in uscita, al fine di individuare in modo chiaro anche la titolarità del dato** e delle sue componenti e di valutarne i vincoli di sicurezza/privacy implicati. Tale attività non dovrebbe implicare rilevanti costi aggiuntivi alla fase di analisi e progettazione del sistema prevista nello sviluppo di un nuovo sistema informativo. Tale documentazione è altresì necessaria per valorizzare i dati anche nei casi di interoperabilità fra banche dati della PA, oltre che al riuso attraverso la pubblicazione in open data.

A livello di sistema informativo, **nella fase di progettazione del servizio, occorre** che siano previste, almeno ad un livello elementare, **delle funzioni di accesso, rilascio ed esportazione della base informativa e della relativa sua metadattazione**. Comunemente si elencano le seguenti modalità, alternative/complementari fra loro, di accesso alla base informativa:

- fornire una funzione di esportazione di tutta la base informativa;
- fornire una funzione di esportazione di una quantità limitata temporalmente dei dati;
- permettere un accesso mediante API (Application Programming Interface) lato backoffice/server;
 - mediante un servizio web con chiamate più strutturate (SOAP/XML)
 - mediante un servizio web con chiamate meno strutturate (REST/JSON)
 - mediante un servizio web e protocollo SPARQL ai dati RDF
- permettere l'accesso ad un file in formato aperto (es. XML/JSON/CSV) pubblicato periodicamente in un area appropriata di un sito web;

In tale senso, nella documentazione di analisi e progettazione deve essere sempre riportata una sezione “Aspetti implicati con la pubblicazione open data” nella quale è descritto come i dati siano compatibili al sistema open data del trentino, o le motivazioni che hanno portato a non poter rendere compatibili con le regole di pubblicazione open data descritte in questo documento.

4.4.3 Georeferenziazione dei dati

Le esperienze di successo sul riuso dei dati hanno dimostrato che il loro valore aumenta quando sono accompagnati dalle informazioni che li **posizionano in modo chiaro nello spazio geografico**.

Oggi le tecnologie mobili permettono di **operare selezioni mirate e personalizzate sulla base del luogo in cui le persone si trovano in un determinato momento**. E' evidente che se le informazioni disponibili per le infinite attività che possono interessare una persona sono “georeferenziate” queste possono essere riutilizzabili in servizi che incontrano la domanda della persona in quei luoghi e in quei momenti. Attraverso la georeferenziazione, quindi, le tecnologie mobili permettono non solo di far cercare le informazioni alle persone, ma anche di creare un incontro tra conoscenza/attività di interesse e persone usando lo spazio come criterio di ricerca.

Con il termine **georeferenziazione** si intende l'attribuzione a un dato di un'informazione relativa alla sua dislocazione geografica, e tale posizione è espressa in un particolare sistema geografico di riferimento. Un esempio sono le coordinate di un **punto** rappresentate con **latitudine o longitudine**, o tramite la descrizione di un **poligono** i cui vertici sono individuati da coordinate geografiche.

La georeferenziazione è un'operazione che può risultare semplice o più complessa in relazione al **grado di dettaglio** che si vuole raggiungere: la precisione per descrivere dove si trova una città può essere rappresentata da un punto la cui latitudine e longitudine sono all'interno dei confini comunali, mentre in un dataset contenente le informazioni sulla produzione fotovoltaica di un edificio, o ancora meglio di una superficie, potrebbe essere necessario anche indicare la forma dell'edificio con le lunghezze espresse in metri e con i punti che rappresentano i vertici che lo descrivono.

Tale obiettivo può essere perseguito in diversi modi all'interno del flusso di inserimento dei dati nel sistema informativo. Il percorso migliore è quello del rilevamento dei dati sul luogo fisico stesso, attraverso **l'uso di una apparecchiatura apposita** (es. navigatore GPS). In molti casi è però possibile

derivare la stessa informazione utilizzando, se esistente ed aggiornata, una cartografia digitale georeferenziata - anche attraverso il "ricalco" dell'oggetto - tramite il **completamento del dataset di informazioni che possono aiutare nell'individuare la posizione geografica**, come il numero civico, il nome della via, il codice di avviamento postale, il nome del comune, della regione, dello stato, ed avvalersi successivamente di un servizio di **geocoding**⁶⁶, anche se quest'ultimo servizio non è disponibile per qualsiasi oggetto geografico e/o toponimo.

I servizi di geocoding sono strumenti online che permettono di ricavare le coordinate a partire da un toponimo o un indirizzo. Bisogna però fare attenzione ad alcuni particolari:

- i servizi di geocoding **si basano su algoritmi di "address matching"**, sistemi intelligenti che cercano di capire come associare quanto richiesto ad una entità geografica (es. riuscire a capire che "Piazza Dante" o "P.za D. Alighieri" corrispondono allo stesso oggetto). Quando l'algoritmo non è in grado di restituire una informazione precisa, allora restituisce **una lista di alternative** con associato un valore di bontà. I servizi online però offrono questo risultato solo quando l'utente ne fa esplicita richiesta e, pertanto, si fermano a restituire il primo valore. Guardando questo, dal lato di una pubblica amministrazione è opportuno verificare che il risultato sia corretto ed è quindi richiesta la conoscenza reale del territorio.
- i servizi di geocoding si basano su **aggiornamenti periodici delle base dati**, quindi un risultato anche se corretto sul piano degli algoritmi, potrebbe restituire un valore errato. Pertanto, nuovamente, i risultati dovrebbero essere controllati⁶⁷
- infine, e questa è la questione più complessa, alcuni servizi di geocoding, avendo necessità di garantire sostenibilità, **hanno termini di riuso che non sempre permettono di riusare i dati senza vincoli aggiuntivi** a meno che non siano formalizzati accordi commerciali in merito⁶⁸.

In alternativa al geocoding, una buona pratica consolidata è l'integrazione di una mappa nei modelli di inserimento dei dati, per permettere di indicare coordinate geografiche con pochi click. Nella maggior parte dei casi, l'operazione consiste nel posizionamento di un punto sopra una mappa, ma non sono escluse funzioni come la creazione di geometrie più complesse.

Più in generale, è possibile recuperare le coordinate dei punti sul territorio o mediante le informazioni geografiche messe a disposizione dalla Provincia tramite il SIAT, Sistema Informativo Ambiente e Territorio e/o dai Comuni e/o mediante strumenti quali OpenStreetMap, per i quali sono anche disponibili delle API per interrogazioni sistematiche⁶⁹

66 http://wiki.openstreetmap.org/wiki/Blogger_map_widget

67 http://de.straba.us/2015/07/14/fan_di_vasco_con_gmaps_arrivi_tardi_al_concerto/

68 <http://de.straba.us/2013/08/20/4-motivi-per-una-pa-per-non-usare-google-maps/>

69 http://wiki.openstreetmap.org/wiki/Category:OSM_API

4.5 Indicazioni operative per i dati di tipo aperto nei capitolati di gara

L'art. 52 del CAD introduce, al comma 3, una disposizione che intende regolamentare la produzione di dati di tipo aperto in tutte quelle gare per la fornitura di prodotti e/o servizi che comportino la raccolta e la gestione di dati pubblici. Per poter attuare tale disposizione sono qui di seguito definite **alcune clausole che possono essere inserite nella documentazione tecnica di gara.**

Si raccomanda in primo luogo di includere **nello schema di contratto** una clausola simile alla seguente volta a richiamare la suddetta disposizione: *“Ai sensi dell’articolo 52 del Codice dell’Amministrazione Digitale (CAD), nella fornitura di prodotti e/o servizi che comportino la raccolta e la gestione di dati pubblici, tali dati, i relativi metadati, gli schemi delle strutture di dati e delle relative banche dati, devono poter essere acceduti telematicamente, nel rispetto del principio di neutralità tecnologica, e predisposti per essere riutilizzati da parte di persone fisiche e giuridiche secondo le modalità tecniche specificate nella documentazione di gara allegata e comunque nel rispetto dei principi e delle raccomandazioni dell’agenda e delle linee guida nazionali per la valorizzazione del patrimonio informativo pubblico rilasciate dall’Agenzia per l’Italia Digitale”.*

Questo implica, per chi sarà assegnatario dell’appalto, di includere all’interno della fornitura le funzionalità ed il codice necessario per la pubblicazione dei dati in modo programmatico nel catalogo dati.trentino.it, ovvero in grado di pubblicare direttamente o su richiesta le informazioni (metadati) relative ai dataset prescelti e del relativo end-point⁷⁰ da cui i dati potranno essere scaricati o acceduti programmaticamente.

A livello di capitolato tecnico di gara, diversi casi sono possibili secondo le peculiarità dell’appalto e comunque delle specificità dei prodotti e/o servizi oggetto della gara. In questo senso, supportati anche dall’analisi della relazione interpretativa che accompagna la disposizione appena citata, emerge come il legislatore abbia voluto in ogni caso riferirsi a un ampio spettro di forniture che comportino la realizzazione non solo di prodotti e/o servizi che abbiano direttamente come oggetto la gestione di dati pubblici, **ma anche quelli che indirettamente producano dati pubblici.**

Si distinguono quindi due diversi casi:

Caso 1

L’oggetto della fornitura riguarda prodotti e/o servizi che comportano **una produzione di dati pubblici che possono essere di tipo aperto (indirettamente)** in tempi successivi; in questo caso, l’amministrazione o la stazione appaltante può **inserire all’interno del capitolato tecnico una clausola** che obblighi l’aggiudicatario della gara verso l’amministrazione a uno o più dei seguenti punti:

- fornire tali dati pubblici in forma disaggregata e, ove possibile, tabellare consentendole di acquisire piena titolarità dei dati; di esplicitare i dati richiesti dall’amministrazione stessa in una clausola prevedendo le eventuali restrizioni normative (per es., norme sulla protezione dei dati personali) e/o di carattere industriale applicabili ai dati;

⁷⁰ https://en.wikipedia.org/wiki/Endpoint_interface

- esplicitare le modalità di acquisizione dei dati da parte dell'amministrazione (per es. tempistiche e meccanismi di trasferimento dall'aggiudicatario all'amministrazione);
- fornire una descrizione puntuale dei dati pubblici prodotti ivi incluse le attività ed entità che hanno dato origine alla produzione dei dati.

L'amministrazione, una volta acquisiti tali dati pubblici, si assume l'onere di renderli pienamente disponibili sotto forma di dati di tipo aperto, secondo le raccomandazioni delle presenti linee guida e gli obiettivi individuati dall'Agenda Nazionale per la valorizzazione del patrimonio informativo pubblico, al fine di incentivare il massimo riutilizzo.

Caso 2

- L'oggetto della fornitura riguarda prodotti e/o servizi che prevedono **la pubblicazione diretta di dati pubblici di tipo aperto**. In questo caso, l'amministrazione può inserire all'interno del capitolato tecnico una serie di clausole che obbligano l'aggiudicatario a produrre direttamente, a partire dall'oggetto della gara, anche dati di tipo aperto il più possibile disaggregati, nel rispetto comunque di eventuali restrizioni normative. Per es. in questo caso, il capitolato può essere dettagliato per richiedere espressamente:
 - la produzione di dati in formato aperto minimo di livello tre, secondo il modello proposto dalle presenti linee guida e tenuto conto dei principi di fruibilità, qualità, interoperabilità;
 - la produzione di metadati che accompagnano i dati aperti prodotti, secondo il modello per i metadati proposto dalle presenti linee guida, includendo anche le indicazioni sulla specifica licenza aperta da applicare ai dati ed ogni altra attività e/o entità che descriva come tali dati sono stati originati (provenance);
 - la pubblicazione dei dati aperti secondo specifiche indicazioni dell'amministrazione e comunque tenendo conto delle raccomandazioni delle presenti linee guida; o l'aggiornamento dei dati aperti, secondo la frequenza di aggiornamento individuata dall'amministrazione;
 - l'esportazione in formato aperto di ogni altro eventuale dato di reportistica generato. In entrambi i casi, l'amministrazione o la stazione appaltante individuerà i parametri dimensionali riguardanti le attività di fornitura di dati pubblici e produzione di dati aperti, rispettivamente, al fine della determinazione della valutazione economica del prodotto e/o servizio oggetto della gara.

4.6 Modello organizzativo

Per garantire la **sostenibilità e la messa a sistema del processo di cambiamento organizzativo** in corso, si definisce, in linea anche con quanto individuato dalle linee guida AgID, il seguente modello organizzativo, focalizzato alla governance della piattaforma dati.trentino.it.

Le **tematiche operative degli open data** vengono realizzate attraverso un team dedicato, a supporto della piattaforma dati.trentino.it (PaT, IT, FBK e altri), e l'individuazione nei dipartimenti della Provincia di figure che hanno il governo della materia sui propri domini verticali di azione.

4.6.1 Team Open Data

Promuove l'uso e la diffusione degli open data. Sviluppa, gestisce e implementa la piattaforma dati.trentino.it al fine di agevolare e creare **economie di scala a livello dell'intero sistema territoriale** (servizio pubblico – ricerca – impresa – comunità) nella gestione delle conoscenze in merito **alle dinamiche di cambiamento normativo, organizzativo e tecnico** connesse al processo di apertura dei dati della pubblica amministrazione.

Gestisce il processo di apertura dei dati dal punto di vista tecnologico e governa lo sviluppo della piattaforma e il suo allineamento alla piattaforma nazionale ed europea.

Si avvale inoltre della **collaborazione dell'Università di Trento e degli istituti di ricerca presenti sul territorio** per garantire l'evoluzione del processo di cambiamento organizzativo in corso allo stato dell'arte della materia e dei fattori che lo influenzano.

Cura all'interno della Provincia autonoma di Trento **i rapporti con il responsabile per la trasparenza e il responsabile per l'anticorruzione** e garantisce i processi di tutela dei dati personali in collaborazione con la struttura competente in materia di privacy.

Si occupa di garantire la formazione e l'aggiornamento diffuso, in particolare degli operatori che contribuiscono al processo di rilascio dei dati per conto dei vari titolari dei medesimi.

Collabora con il nodo **nazionale ODI - FBK** sui temi della qualità dei dati.

Supporta il **coordinamento fra tutti gli enti** che partecipano alla piattaforma e mette a disposizione le sue reti relazionali per tenere allineato l'intero sistema territoriale allo sviluppo della materia a livello nazionale e comunitario.

Cura i rapporti con le **società di sistema provinciali, il sistema degli enti locali, gli enti e le fondazioni di ricerca e i soggetti privati.**

Il gruppo è **coordinato dalla Provincia autonoma di Trento coadiuvata dalla società Informatica Trentina e fa riferimento – sotto il coordinamento della Direzione generale della Provincia – alle competenze del Servizio Supporto alla Direzione generale e ICT.**

È inoltre compito del Team Open Data redigere ogni anno, in modo allineato con la programmazione dei sistemi informativi e con l'attuazione della Agenda Digitale del Trentino, un'agenda (piano) annuale, per armonizzare l'apertura dei dataset del Trentino alle priorità individuate dalla pianificazione nazionale ed europea.

Il Team si coordina con gli attori del territorio interessati per garantire: **almeno un evento annuale**, come momento di dialogo sullo stato di avanzamento del processo di apertura dei dati; **attività di formazione** in collaborazione con università, enti di ricerca ed eventuali altre amministrazioni; le **attività laboratoriali già avviate con l'Università di Trento**, individuando attività su cui sviluppare possibili tirocini.

Il Team cura infine **le attività di comunicazione e si occupa del monitoraggio dell'intero processo**, anche attraverso lo sviluppo di un sistema di metriche in linea con il dibattito allo stato dell'arte sul tema.

4.6.2 Referente Open Data nei dipartimenti

Nel caso della Provincia, il referente è individuato dal dirigente generale del Dipartimento di norma fra il personale con maggior competenza in ambito di **comunicazione web, sistemi informativi e trattamento dati**; collabora con il dirigente generale **nella pianificazione degli sviluppi in ambito open data** e ne segue, entro il proprio specifico campo di competenza, le evoluzioni a livello nazionale ed europeo in tema di standard tecnici, metadatazione, evoluzione normativa, potendo avvalersi del supporto del Team di progetto Open Data. Collabora con il Team Open Data per definire l'agenda provinciale (piano dati) annuale.

4.6.3 Titolari dei dati

Il titolare del dato è la pubblica amministrazione o l'organismo di diritto pubblico che ha originariamente formato per uso proprio o commissionato ad altro soggetto pubblico o privato il documento che rappresenta il dato (d.lgs. 36/2003). Nel caso della Provincia, i dirigenti delle singole strutture/dipartimenti esercitano la titolarità per conto della Provincia, nel caso degli altri enti la Provincia si interfaccia con la struttura indicata dall'ente medesimo.

Spetta al titolare **valutare**, supportato dal Team Open Data, **la reale titolarità sui dati che intende rendere disponibili, la legittimità del processo, la pianificazione, la possibilità di sviluppare processi virtuosi di riutilizzo dei dati in sinergia con il sistema della ricerca e delle imprese locali.** Il

Team di progetto mette a disposizione strumenti per aiutare questi processi di valutazione ed è a disposizione per approfondimenti e delucidazioni specifiche.

4.6.4 Referenti tecnici

Il referente tecnico è individuato dai dirigenti delle singole strutture fra i propri collaboratori - meglio se già coinvolto in attività legate alla gestione del sito web, privacy, trasparenza, informatica, statistica – e rappresenta la figura che materialmente si occupa di operare, insieme al Team Open Data, sulla piattaforma dati.Trentino.it. Gestisce i dati, verifica il flusso di produzione e di comunicazione ad altri di quel dataset, crea i metadati, li aggiorna, si interfaccia con il referente di Informatica Trentina del Team per individuare soluzioni sempre più efficienti e automatizzate di pubblicazione dei dati. Utilizza i tools messi a disposizione dal Team per seguire processi corretti di pubblicazione.

4.6.5 Raccordo con le strutture provinciali, con il sistema trentino della pubblica amministrazione digitale e con la Community Open Data

I cambiamenti tecnologici, normativi e organizzativi connessi ad una nuova visione dell'amministrazione digitale con al centro il cittadino e quindi con i dati e le informazioni come elemento **strutturale di un dialogo nuovo, non più parcellizzato nelle visioni interne alle pubbliche amministrazioni**, ma a servizio di risposte coerenti e trasversali alle diverse strutture organizzative richiedono alla Provincia di affrontare in modo coordinato molte questioni trasversali alle diverse competenze delle strutture provinciali.

In questo contesto complesso, il processo di valorizzazione del patrimonio informativo pubblico ha conseguenze su molte attività della pubblica amministrazione e **può portare a cambiare pervasivamente molti processi quotidiani degli uffici pubblici e della pubblica amministrazione.** Questi processi hanno conseguenze sui temi **della privacy, dell'accesso, della trasparenza, dei rapporti con i cittadini e delle nuove forme di government**, e richiedono anche una diversa attenzione **alla comunicazione web.**

Per le ragioni indicate, è imprescindibile l'attivazione di un **Gruppo di coordinamento trasversale** che sviluppi le necessarie sinergie e le azioni di coordinamento tra tutte le strutture competenti sui vari tematismi richiamati e correlati allo sviluppo degli open data (profili legislativi, di privacy, di trasparenza, di semplificazione amministrativa, di statistica, di comunicazione web, di utilizzo delle ICT, ...), allo scopo di affrontare le varie tematiche potendo usufruire delle competenze specifiche di ogni ambito disciplinare per poi mettere le proprie conoscenze a servizio dell'intero sistema locale.

Per garantire un allineamento del sistema trentino attorno alla piattaforma dati.trentino.it la struttura competente in materia di open data stabilisce - in accordo con gli altri enti del territorio - le modalità di cooperazione sulla piattaforma dati.trentino.it.

Per facilitare la creazione di una comunità di pratiche attorno alla piattaforma sono **resi disponibili, a tutti gli enti che vi aderiscono, gli strumenti base che aiutino a definire processi, modalità di collaborazione, spazi di condivisione delle conoscenze**, una pianificazione di base di eventi, momenti formativi, i riferimenti dei referenti di ambito disponibili a partecipare a gruppi di lavoro nazionali su temi di interesse dei singoli enti a beneficio dell'intero sistema territoriale.

Infine, è importante evidenziare come il progresso degli open data si sia avvalso in questi anni e continui ad avvalersi di una comunità di aziende, professionisti, ricercatori, studenti che sono interessati al riutilizzo per vari finalità dei dati e per questo hanno garantito costante disponibilità a collaborare e mettere a disposizione della piattaforma le proprie conoscenze.

Per dare corso ad azioni in sintonia con gli sviluppi dell'ecosistema dell'innovazione trentina il Team di progetto Open Data può fare riferimento anche alla disponibilità del gruppo **“TOD - Trentino Open Data”** nell'indicare possibili criticità o anche utili linee possibili di sviluppo. Il Gruppo **“TOD - Trentino Open Data”** è **l'espressione locale della comunità informale open data italiana**. Il gruppo si raccoglie virtualmente attorno ad un profilo social⁷¹. Tale gruppo ha svolto anche una importante funzione di riferimento nella definizione delle presenti linee guida, che sono state riviste dalla community locale per valutarne l'eventuale impatto e raccogliere utili suggerimenti in modo partecipativo.

Il progetto Open Data ha un indirizzo e-mail di riferimento⁷², un profilo twitter⁷³ e anima il gruppo facebook Trentino Open Data.

In modo parallelo, la piattaforma dati.trentino.it si appoggia **alla pagina del progetto sul portale istituzionale della Provincia**⁷⁴, sulla quale vengono monitorati gli sviluppi con il contributo dei vari attori del sistema.

71 Pagina Facebook del gruppo <https://www.facebook.com/groups/todgroup/>

72 info@datitrentino.it

73 <https://twitter.com/datitrentinoit>

74 <http://www.innovazione.provincia.tn.it/opendata>

5. Modelli e standard tecnici di riferimento

Questo capitolo descrive **alcuni dettagli tecnici relativi agli standard** per operare sul catalogo dati.trentino.it in modo uniforme e **allineato al dibattito nazionale ed europeo**.

In particolare, sono elencati i formati che i dati possono assumere per poter essere considerati aperti; l'elenco dei metadati obbligatori previsti dallo standard internazionale e un aggiornamento sullo stato dell'arte del dibattito circa la definizione della specifica EU volta alla standardizzazione della metadattazione dei dataset nei cataloghi open data di tutta Europa (DCAT-AP); eventuali vocabolari/ontologie⁷⁵ di riferimento; e alcune indicazioni di come in Trentino si è affrontato e si stanno affrontando questi temi e in particolare il tema dei "linked open data".

5.1 Standard tecnici e modello "5 stars" per produrre e pubblicare dati aperti

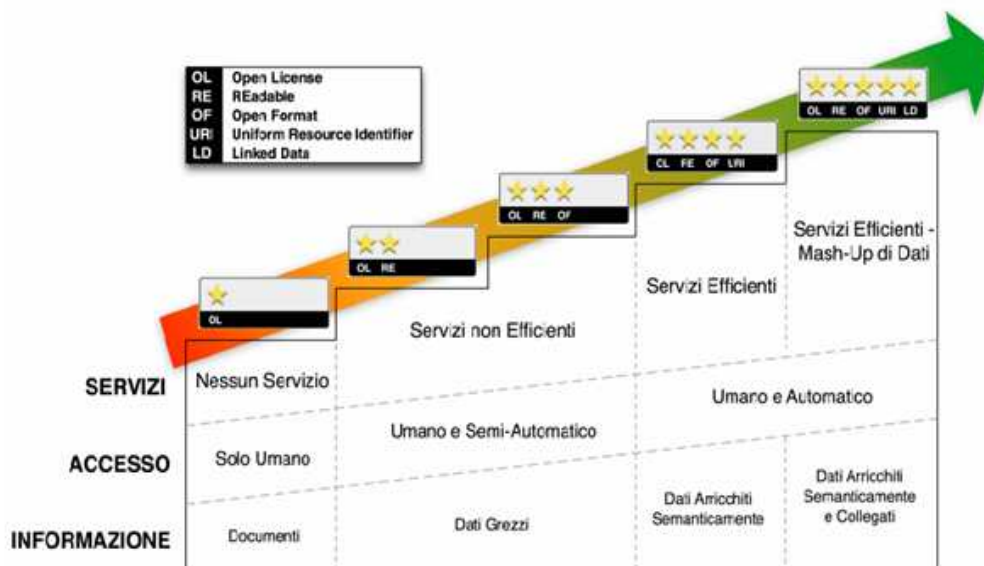
Il modello "5 stars" per la produzione e il rilascio di dati di tipo aperto riportato nella **Figura 4**⁷⁶ è il riferimento base della comunità open data internazionale per definire quando e attraverso quali caratteristiche di "arricchimento" i dati possono essere chiamati aperti al fine di sviluppare conoscenza e innovazione grazie ad un **loro ampio riutilizzo**.

Osservare il modello nel dettaglio aiuta a collocare le specifiche tecniche che vengono proposte in questo capitolo.

75 Il termine "ontologia" (formale) è entrato in uso nel campo dell'intelligenza artificiale e della rappresentazione della conoscenza, per descrivere il modo in cui diversi schemi vengono combinati in una struttura dati contenente tutte le entità rilevanti e le loro relazioni in un dominio.

76 Questo il portale del "five stars Open Data" <http://5stardata.info/> della World Wide Web Foundation

Figura 4 - Il modello "5 stars" per la produzione e il rilascio di dati di tipo aperto



Ogni colonna rappresenta un livello di qualità dei dati, nel dettaglio:

- prima colonna/una stella:** i dati, così come sono pubblicati nel web, sono solo **forniti di licenza aperta**, si noti che oggi tutte le informazioni del servizio pubblico pubblicate sui portali, salvo altra licenza esplicita, sono da considerarsi aperte, riutilizzabili o come si dice "open by default". A questo livello i dati possono essere riutilizzati da chiunque anche a scopo commerciale ma **richiedono molto lavoro e di conseguenza non facilitano il riutilizzo** e la crescita dell'economia. Come evidenziano le variabili descritte in orizzontale nel modello, a questo livello **la lettura è solo di tipo umano e quindi, solo in seconda battuta, si può utilizzare qualche software per poterne elaborare i contenuti e agevolare il riutilizzo**. Di fatto tutti i dati a questo livello sono da considerarsi come "documenti" che allo stato attuale dell'arte hanno a disposizione pochi servizi automatizzati ed estremamente costosi, per estrarne i contenuti e poterli mettere in circolo.
- seconda colonna/due stelle:** a questo livello troviamo, oltre alla licenza per il riutilizzo, anche un **formato leggibile facilmente da un programma software**. Il riutilizzo è facilitato dal computer, ma a questo livello, **se si usano formati proprietari**, il numero di programmi in grado di leggere correttamente il documento sarà **limitato e vi saranno dei vincoli** connessi alle licenze d'uso e dipendenti dalle scelte delle terze parti proprietarie dei formati. Per questi motivi il campo di applicazione del riutilizzo potrebbe essere ridotto.

- **terza colonna/tre stelle:** supera il limite del livello precedente quando si utilizza un formato non proprietario, cioè **con delle specifiche tecniche di pubblico dominio**, che lo descrivono e archiviano in modo libero da restrizioni legali per il suo utilizzo. (per. es. CSV, geoJSON, ...). In generale, sia il livello due che il livello tre si raggiungono quando si hanno **dati “grezzi”**, dati cioè che di solito sono prodotti per scopi specifici negli uffici e vengono successivamente proposti nei cataloghi open data o nei siti senza modifiche. Questi dati richiedono **un impegno minimo da parte degli uffici** perché non sono lavorati con l’obiettivo di aumentarne il loro valore d’uso. I servizi applicabili a questi dati possono però risultare poco efficienti e spesso necessitano di un forte intervento umano anche nella lettura automatizzata. C’è da dire **che molti dei dati che oggi sono stati aperti** nel nostro Paese⁷⁷ - compresi quelli nel catalogo dati.trentino.it - **raggiungono questo livello:** ciò perché nel processo di cambiamento organizzativo e di valorizzazione dei dati prodotti nelle pubbliche amministrazioni tale **livello è raggiungibile a costo quasi nullo**. Per i livelli successivi è invece richiesta **una maggior consapevolezza del valore insito nei dati** così che si possa **superare una prima fase esplorativa e mettere a sistema il processo di produzione e pubblicazione**.
- **quarta colonna/quattro stelle:** si realizza quando **i dati sono identificati da URI**⁷⁸ usando gli standard RDF (Resource Description Framework) per rappresentarli e SPARQL (Structured Protocol RDF Query Language) per interrogarli. In questo modo la ricerca dei dati nella rete e il riutilizzo è reso più “automatico” perché **facendo riferimenti a standard globali permette di collegare dati/conoscenze “locali” a dati/conoscenze “globali” presenti nel Web**. L’utilizzo di programmi **che sfruttano accessi di tipo web service** per recuperare i dati di interesse rendono molto più automatico e affidabile il processo attraverso il quale chi riutilizza il dato tiene alimentata la sua fonte di informazioni necessaria a dare servizi efficienti alla sua specifica nicchia di mercato.
- **quinta colonna/cinque stelle:** la presenza di **collegamenti (link)** fra dati diversi ma **legati da uno standard identificativo comune** renderebbe il campo di azione e di riutilizzo del dato praticamente infinito. Questo perché la possibilità di realizzare *servizi con mashup* (combinazione, fusione) *di dati* - applicazioni web che permettono di includere dinamicamente informazioni o contenuti **provenienti da più fonti, tipico dei linked open data - riuscirebbe ad arricchire potenzialmente in modo infinito i dataset disponibili al riutilizzo**. Questo, però, a condizione **che si trovassero accordi sugli standard comuni da usare**. A questi ultimi due livelli di apertura dei dati i programmi sono anche in grado di inglobare e riconoscere l'ontologia di riferimento e pertanto di elaborare i dati **quasi senza ulteriori interventi umani**.

E’ bene osservare che, nel caso in cui i dati fossero **arricchiti semanticamente**, l'intervento umano necessario a togliere l’ambiguità si potrebbe ridurre notevolmente. Le ambiguità sono sempre presenti nel linguaggio e quindi anche nei dati che ne sono una delle possibili formalizzazioni: l’arricchimento semantico cerca di ovviare al problema descrivendo i dati mediante vocabolari condivisi e ontologie definite, per esempio, in un gruppo di lavoro di uno specifico settore. Una buona riuscita del lavoro descritto poco sopra, ovvero una concettualizzazione condivisa almeno a

77 <http://www.agid.gov.it/agenda-digitale/open-data/basi-dati-pa>

78 http://it.wikipedia.org/wiki/Resource_Description_Framework <http://it.wikipedia.org/wiki/SPARQL>

livello di sovrastrutture, permetterebbe l'esportazione e lo scambio fluido di sorgenti di dati diverse tra differenti gruppi di lavoro. e aggiungendo opportuni metadati per contestualizzarli. Specialmente quando i dati **fossero anche resi linked**, i loro significati e ambiti di applicazione potrebbero risultare più evidenti, consentendo l'interoperabilità semantica tra applicazioni eterogenee, con un esponenziale aumento di valore anche economico dei dati prodotti dal servizio pubblico.

Il **condizionale utilizzato nei paragrafi precedenti è d'obbligo**, non essendo i ricercatori del settore ancora giunti a individuare una soluzione univoca, né a definire se si debba trovare una soluzione o non sia invece meglio individuare più soluzioni collegabili fra loro. Si tratta di un campo di ricerca e sviluppo **sul quale si stanno investendo da molti anni consistenti energie** per arricchire di significato e senso, e quindi per valorizzare, il patrimonio informativo rappresentato dai dati presenti in rete. Sta di fatto comunque che ogni piccolo sforzo teso a ridurre l'ambiguità di anche solo un frammento del complesso vocabolario utilizzato per descrivere i dati della Pubblica Amministrazione ha un immediato ritorno in termini di riduzione dei costi legati ai problemi di interoperabilità tra i sistemi di dati delle PPAA, oltre che in termini di trasparenza e dialogo verso i cittadini.

Di seguito viene descritto **dove il catalogo dati.trentino.it si collochi in questo dibattito** e come, dovendo operare scelte e definire passi concreti ora, si siano prese decisioni **per permettere standard minimi e apertura a possibili sviluppi**.

5.2 Formati aperti

Ogni dominio di applicazione utilizza propri standard tecnici: i formati dei documenti sono diversi da quelli delle tabelle e diversi da quelli di grafici, della cartografia, delle immagini, dei video ecc. Di seguito, si **definiscono i formati aperti più comunemente utilizzati e suggeriti per la pubblicazione** sul catalogo dati.trentino.it.

E' sempre auspicabile l'utilizzo di una pluralità di formati, ove possibile, in modo da offrire più opzioni ai riutilizzatori. I formati proposti nelle tabelle che seguono sono stati selezionati facendo riferimento agli **standard internazionali, alle prescrizioni normative previste dal Codice dall'Amministrazione Digitale (CAD) e agli standard indicati dalla deliberazione della Giunta provinciale n. 195 del 17 febbraio 2012** relativi al rilascio di un sottoinsieme di dati del Sistema Informativo Ambiente e Territorio (SIAT) dalla Provincia Autonoma di Trento secondo il paradigma degli Open Government Data.

5.2.1 Formati di tipo aperto per documenti

Nel caso **di documenti in generale**, si raccomanda, ove esistano, standard XML documentali internazionali o nazionali. La tabella seguente fornisce un elenco dei formati aperti maggiormente diffusi per i dati documentali.

Per quanto riguarda i documenti, quali ad esempio delibere, atti amministrativi si osserva che nel caso siano da pubblicare dei dati solo in forma tabellare senza elementi aggiuntivi (per es., grafici, formule, tipi di dato, metadati) si può evitare di utilizzare formati come PDF o ODT ma è possibile utilizzare il formato CSV. Nei casi con dati in forma tabellare con elementi aggiuntivi, si raccomanda l'uso, ove possibile di ODS.

Tabella 4 Formati per dati documentali documentali

Nome (Acronimo) Descrizione	Estensione
<p>ODT (Open Document Text).</p> <p>E' uno standard aperto per documenti testuali basato su XML. Fa parte dello standard OASIS Open Document Format for Office Applications [18]. E' stato adottato come formato principale per i testi in alcune suite per l'automazione d'ufficio come OpenOffice.org e LibreOffice ed è supportato da altre come Microsoft Office, Google Docs e IBM Lotus.</p>	.odt
<p>PDF (Portable Document Format).</p> <p>E' un formato aperto creato da Adobe per la rappresentazione di documenti contenenti testo e immagini che sia indipendente dalla piattaforma di lettura (applicativo, sistema operativo e hardware). È stato successivamente standardizzato dall'ISO (ISO/IEC 32000-1:2008) con una serie di formati differenti, ognuno avente una propria prerogativa (e.g., PDF/UA per l'accessibilità, PDF/H per documenti sanitari, PDF/A per l'archiviazione, ecc.). Se ne scoraggia assolutamente l'uso per distribuire tabelle.</p>	.pdf

5.2.2 Formati di tipo aperto per dati "strutturati generici"

La **Tabella 5** mostra un primo elenco dei formati aperti maggiormente diffusi nell'ambito dell'amministrazione pubblica ai quali si intende fare riferimento per le attività di pubblicazione di dati strutturati. Tali formati sono adatti per la rappresentazione e lo scambio di dati, anche quindi metadati, più o meno strutturati.

Tabella 5 Formati aperti adatti alla pubblicazione dei dati aperti di "tipo strutturato" generale, ovvero non relativo a specifici domini applicativi o di settore

Nome (Acronimo) e descrizione	Tipo di Dato	Estensione
<p>Comma Separated Value (CSV)</p> <p>Formato di file testuale usato per rappresentare informazioni con struttura tabellare, le cui righe corrispondono a linee e i cui valori delle</p>	Dato tabellare	.csv

<p>singole colonne sono separati da una virgola (o punto e virgola). Esso è spesso usato per importare ed esportare il contenuto di tabelle di database relazionali e fogli elettronici. Le righe delle tabelle corrispondono a righe nel file di testo CSV e i valori delle celle sono divisi da un carattere separatore.</p>		
<p>Tab Separated Value (TSV)</p> <p>Formato di file testuale per l'interscambio di tabelle, le cui righe corrispondono a linee e i cui valori delle singole colonne sono separati da un carattere di tabulazione</p>	Dato tabellare	.tsv
<p>ODS (Open Document Spreadsheet).</p> <p>Formato di file XML basato su standard aperto per fogli di calcolo. Fa parte dello standard OASIS Open Document Format for Office Applications [18]. Come nel caso precedente, è stato adottato come formato principale per i fogli di calcolo in alcune suite per l'automazione d'ufficio come OpenOffice.org e LibreOffice ed è supportato da altre come Microsoft Office, Google Docs e IBM Lotus. Si incoraggia a farne uso per la distribuzione di tabelle.</p>	Dato tabellare	.ods
<p>Extensible Markup Language (XML)</p> <p>XML è un linguaggio generale di marcatura, ovvero basato su un meccanismo che consente di definire e controllare il significato degli elementi contenuti in un documento o in un testo attraverso delle etichette (markup), standardizzato dal W3C usato per l'annotazione di documenti e per la costruzione di altri linguaggi più specifici per l'annotazione di documenti. Esistono strumenti che possono essere utilizzati per definire la struttura dell'XML come DTD e/o XML schema.</p> <p>Il formato XML è utilizzato per rappresentare altri formati standard quali ad esempio RDF, ODS e molti altri formati specifici di settore</p>	Dato strutturato	.xml
<p>Resource Description Framework (RDF)</p> <p>Formato di file XML basato su uno standard aperto proposto da World Wide Web Consortium (W3C) per la codifica, lo scambio e il riutilizzo di dati linked e metadati strutturati e consente l'interoperabilità tra applicazioni che si scambiano informazioni sul Web</p>	Dato strutturato	.rdf
<p>Notation3 (N3) e Turtle</p> <p>L'N3 è un formato di file pensato per essere più compatto rispetto a quella ottenuta utilizzando la sintassi XML del RDF. Essa risulta più leggibile da parte degli umani e possiede delle caratteristiche che esulano dall'uso stretto di RDF (e.g., rappresentazione di formule logiche). In modo simile, il formato Turtle (anche noto come Terse RDF Triple Language) è una versione semplificata (un sottoinsieme di funzionalità) di N3 che include solo aspetti relativi a RDF</p>	Dato strutturato	.n3
<p>JSON (JavaScript Object Notation)</p> <p>E' un formato maggiormente leggibile dagli umani e che mantiene,</p>	Dato strutturato	.json

rispetto a formati simili come l'XML, una sintassi poco prolissa. Questo aspetto ne fa un formato flessibile e compatto. Esso nasce dalla rappresentazione di strutture dati semplici nel linguaggio di programmazione JavaScript, ma mantiene indipendenza rispetto ai linguaggi di programmazione. Ha avuto ampia diffusione per il fatto di essere flessibile e compatto		
JSON-LD E' un formato di file per la rappresentazione di dati linked che si basa sulla sintassi JSON ed è pensato per essere più compatto rispetto alla rappresentazione ottenuta con il formato RDF. Il formato JSON-LD è stato standardizzato recentemente dal W3C, con la speranza di essere accolto favorevolmente dall'imponente numero di programmatori che già utilizzano e conoscono JSON ma che potrebbero avere necessità di manipolare dati linked.	Dato strutturato	.jsonld

5.2.3 Formati di tipo aperto per dati geografici

Fra i vari tipi di dati su specifici domini, i dati geografici sono stati i primi ad essere rivisti in ottica di standardizzazione e riutilizzo in modalità aperta, così che oggi l'ambito applicativo più diffuso e consolidato in tema di standard sono proprio i dati geografici-ambientali. Nella **Tabella 6** si elencano i formati utilizzati nel geoportale SIAT PAT ⁷⁹ e quindi resi consultabili assieme a tutti gli altri dati del sistema trentino sul catalogo dati.trentino.it.

Tabella 6 Formati aperti per la pubblicazione dei dati aperti di tipo geografico

Nome (Acronimo) Descrizione	Tipo di Dato	Estensione
Geographic Markup Language (GML) <i>Formato XML utile allo scambio di dati territoriali di tipo vettoriale. Si tratta di una grammatica XML che rappresenta un formato di scambio aperto per i dati territoriali. Essa, definita originariamente da OGC e diventata lo Standard ISO 19136:2008, fornisce la codifica XML (schemi XSD) delle classi concettuali definite in diversi Standard ISO della serie 19100 e di classi aggiuntive appositamente definite: geometrie, oggetti topologici, unità di misura, tipi di base, riferimenti temporali, feature, sistemi di riferimento, copertura</i>	Dato geografico vettoriale	.gml
Keyhole Markup Language (KML)	Dato geografico	.kml

⁷⁹ <http://www.territorio.provincia.tn.it/>

<p><i>È un formato basato su XML per rappresentare dati geografici. Nato con Google, è diventato poi uno standard OGC. Le specifiche della versione 2.2 presentano una serie di entità XML attraverso cui archiviare le coordinate geografiche che rappresentano punti, linee e poligoni espressi in coordinate WGS84 e altre utili a definire gli stili attraverso cui i dati andranno visualizzati. Si incoraggia l'uso solo se il file viene distribuito attraverso adeguata strutturazione degli attributi delle geometrie</i></p>	vettoriale	
<p>ESRI Shapefile (SHP)</p> <p><i>È il formato standard de-facto (non è un formato aperto ma sono note le specifiche) per la rappresentazione dei dati dei sistemi informativi geografici (GIS). I dati sono di tipo vettoriale. Lo shapefile è stato creato dalla società privata ESRI che rende comunque pubbliche le sue specifiche. L'apertura delle specifiche ha consentito lo sviluppo di diversi strumenti in grado di gestire e creare tale formato. Seppur impropriamente ci si riferisca a uno shapefile, nella pratica si devono considerare almeno tre file: un .shp contenente le forme geometriche, un .dbf contenente il database degli attributi delle forme geometriche e un file .shx come indice delle forme geometriche. A questi tre si deve anche accompagnare un file .prj che contiene le impostazioni del sistema di riferimento. I files possono essere distribuiti all'interno di un file archivio compresso (es. ".zip"). In questo caso è necessario indicare il formato dei dati contenuti nell'archivio come SHP. Si incoraggia a distribuire i dati completi del file .prj</i></p>	Dato geografico vettoriale	.shp, .shx, .dbf, .prj
<p>GeoJSON</p> <p><i>E' un formato aperto per la rappresentazione e l'interscambio dei dati territoriali in forma vettoriale, basato su JSON (JavaScript Object Notation). Ogni dato è codificato come oggetto che può rappresentare una geometria, una feature o una collezione di feature. Ad ogni elemento geografico sono associati degli attributi attraverso un insieme di coppie nome/valore (membri)</i></p>	Dato Geografico vettoriale	.json
<p>Geopackage</p> <p><i>E' un formato aperto per la rappresentazione di dati geografici che può essere considerato un'alternativa al formato shapefile prima introdotto. Esso supporta SpatialLite ovvero un'estensione dello schema del database SQLite. Se ne incoraggia l'uso e la distribuzione grazie alla duttilità del formato che, in un unico file, riesce a distribuire le geometrie, gli attributi collegati e le modalità di rappresentazione.</i></p>	Dato Geografico Vettoriale	.gpkg
<p>Formati Immagini</p> <p><i>I dati raster possono cioè essere memorizzati attraverso tipologie di file che sfruttano algoritmi di compressione diversi, gravando in modo differente sul supporto di memorizzazione. I formati raster più comuni sono i seguenti: non compressi; con</i></p>	Dato geografico Raster	.raw, .bmp, .png, .jpeg, .tiff

<i>compressione lossless (cioè senza perdita di informazione); con compressione lossy (cioè con perdita di informazione, per cui il file decompresso avrà qualità inferiore dell'originale)</i>		
---	--	--

5.2.4 Formati di tipo proprietario “aperti de facto”

La **tabella 7** fornisce un elenco dei formati proprietari “de facto” più diffusi di cui è comunque suggerita sempre la pubblicazione anche di una versione del dato in un formato di tipo aperto. Si sottolinea che, nel caso di pubblicazione di dati in forma solo tabellare, senza elementi aggiuntivi (es. metadati, grafici, formule) è possibile evitare di utilizzare formati quali PDF o ODT e utilizzare il formato CSV. Negli altri casi di dati in forma tabellare, con elementi aggiuntivi quali in particolare metadati, si raccomanda l’uso del formato ODS. Nel caso di documenti in generale, si raccomanda, ove esistano, l’utilizzo di standard XML documentali internazionali o nazionali.

Tabella 7 Formati proprietari “de facto”

Tipologia	Tipo di Dato	Estensione
Microsoft Excel (fino alla versione 2003) <i>Formato proprietario Microsoft, leggibile da programmi Open Source, utile alla memorizzazione di fogli di calcolo</i>	Dato tabellare	.xls
Microsoft Excel versioni successive (XSLX) <i>Formato proprietario Microsoft, leggibile da programmi Open Source qualora non siano introdotte formattazioni complesse, utile alla memorizzazione di fogli di calcolo</i>	Dato tabellare	.xlsx
DXF (Drawing Interchange Format) <i>è un formato per i file di tipo CAD, sviluppato da Autodesk come soluzione per scambiare dati tra il programma AutoCAD e altri programmi.</i>	Dato geografico vettoriale	.dxf

Tali formati sono quindi utilizzabili solamente insieme ad un formato di tipo aperto. Si pone attenzione al fatto che **l'utilizzo di tali formati è da limitare ove possibile ai soli casi di necessità, per agevolare il riuso** del dato e per permettere una transizione completa più graduale nel tempo ai formati di tipo aperto.

5.3 Modello per la metadatazione

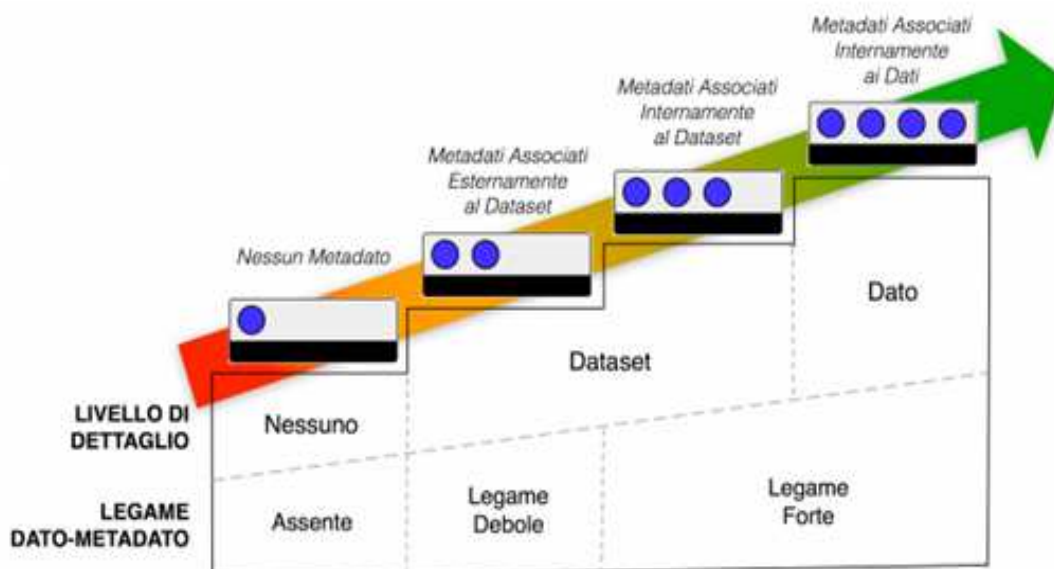
La metadatazione, ossia quella attività che **associa ai dataset e ai dati dentro un dataset tutte quelle informazioni necessarie per definirne le caratteristiche**, quali la titolarità, la scadenza o l'aggiornamento, l'ambito territoriale, eventuali Url di riferimento ecc., ricopre un ruolo essenziale per garantire che i dati esposti possano essere **ben interpretati e quindi riutilizzati da utenti esterni e da agenti automatici** in modo corretto così da garantire la qualità dei servizi creati su quei dati. Tali informazioni agevolano anche la ricerca dei dati all'interno dei cataloghi.

Esiste, anche in questo caso, un modello **per valutare la qualità dei metadati** funzionale allo sviluppo di servizi innovativi sulla base di quei dati e quindi che punta l'attenzione **sull'accessibilità dei metadati e al loro livello di granularità**.

5.3.1 Modello di riferimento

Il modello evidenziato nella **Figura 5** mette in evidenza **il legame tra dato-metadato** che è più forte più il metadato è collegato con il dato all'interno del processo di pubblicazione del dataset; il livello di **dettaglio dei metadati** che varia dall'assegnare metadati "generali" solo al dataset (per es, validità, licenza, autore, ecc.) all'assegnare metadati anche alle variabili interne ai dataset (unità di misura, formati, ecc.).

Figura 5 Modello di riferimento per la definizione dei metadati



Le combinazioni possibili di questi fattori individuano livelli diversi di **accessibilità da parte dei sistemi informativi**, inoltre ogni tipo formato che si utilizza permette livelli diversi di integrazione dei metadati. Le linee guida AgID, a questo proposito, specificano come **ogni diverso formato** siano associati in modo tecnologico gradi diversi di metadattazione:

- al **livello 1** non esiste metadattazione, o i dati sono autoevidenti o il loro riutilizzo è molto azzardato o richiede molto lavoro di decifrazione;
- al **livello 2** esiste una **descrizione testuale o un documento** esterno al dataset che spiega come è rappresentato il dato in linguaggio naturale (per es. il caso del formato CSV);
- al **livello 3 in un primo caso** l'uso di un formato quale l'ODS permette di raccogliere tutti i metadati del formato ODS attraverso il formato stesso, che ne permette anche la loro descrizione, considerando comunque la necessità di salvaguardare le indicazioni valide già indicate per il formato CSV, ovvero non utilizzare formule, non formattare in modo grafico o inserire note in linguaggio naturale dentro il file dati, limitandosi il più possibile a tabelle a "doppia entrata" e le altre buone pratiche riportate nel vademecum open data⁸⁰;
- **in un secondo caso, sempre al livello 3**, l'uso di un formato come l'XML con uno SCHEMA XML⁸¹ associato, permette di avere un maggiore legame tra dati e metadati, migliorando la possibilità di interpretazione da parte di sistemi automatici. In questo caso si può anche fare riferimento ad un dizionario di riferimento referenziato esternamente;
- al **livello 4** l'uso di un formato come nel punto precedente (XML + SCHEMA XML) ma con in più **l'uso di una o più ontologie all'interno del formato**⁸² capace di identificare in modo non ambiguo le entità descritte dalle variabili dei dataset, **permette il livello di interoperabilità più elevato.**

Essendo definiti i metadati come quei "dati/informazioni associate ai dati tali da informare gli utilizzatori del significato del dato stesso", ne deriva che **l'ampiezza dei metadati associabili** ad un dataset **è potenzialmente illimitata**, ed è per questo che è opportuno identificare **un insieme minimo ma condiviso, a livello nazionale ed europeo, di informazioni di base** che possano essere utilizzate ai fini della ricerca, dell'accesso e dell'utilizzo del dato pubblicato.

In tale senso, viene fornito qui di seguito **un set minimo di metadati considerati indispensabili per la pubblicazione dei dati aperti da parte della Provincia Autonoma di Trento** in accordo con lo stato dell'arte del dibattito in sede AgID e a livello europeo.

80 http://www.innovazione.provincia.tn.it/contenuti.php?t=opendata_tab&id=19

81 https://it.wikipedia.org/wiki/XML_Schema

82 <http://5stardata.info/gtd-5.html>

5.3.2 Verso uno standard unico per la metadattazione: DCAT-AP

Le indicazioni per la metadattazione definite in queste linee guida fanno riferimento alla specifica europea sulla metadattazione dei dati all'interno dei portali dati chiamata DCAT-AP⁸³.

L'Unione Europea raccomanda **lo standard DCAT-AP per la metadattazione dei dataset open data**. Il profilo DCAT-AP è una specifica basata sul vocabolario Catalogue Scheda (DCAT) per descrivere la struttura (non il contenuto) dei dataset del settore pubblico in Europa⁸⁴. L'elaborazione del DCAT-AP è stata una iniziativa comune della DG CONNECT, dell'Ufficio delle Pubblicazioni dell'Unione Europea e del Programma ISA. La specifica è stata elaborata da un gruppo di lavoro multidisciplinare con i rappresentanti di 16 Stati membri dell'Unione europea, alcune istituzioni europee e gli Stati Uniti. Nel mese di febbraio 2015, il programma ISA della Commissione europea ha lanciato un'attività di rivedere DCAT-AP con l'effetto di rilasciare una nuova versione (la 1.1) nell'ottobre 2015. Lo standard DCAT-AP utilizza vari vocabolari e ontologie di tipo generale.

Nel dettaglio, la raccomandazione DCAT-AP prevede tre classi: il **catalogo** - è un archivio che ospita i dataset; il **dataset** - è una raccolta di dati, pubblicati o curati da una singola sorgente e disponibili per l'accesso o il download di uno o più formati; la **distribuzione** - è una rappresentazione fisica del dataset in un particolare formato (es. file CSV). Ognuna delle 3 classi ha un livello di metadattazione, composto da un elenco di metadati e da dei vocabolari da utilizzare per la loro compilazione.

Alla data della redazione di questo documento, si prevede che tale specifica sarà recepita entro il 2016 da parte del portale europeo dei dati aperti (www.europeandataportal.com) e a cascata da tutti i portali open data nazionali tra cui il portale nazionale italiano dati.gov.it. In tale direzione il gruppo di lavoro open data di AgID sta lavorando insieme al gruppo di standardizzazione europeo in modo da istanziare correttamente DCAT-AP in Italia, permettendo il collegamento dei portali regionali al portale nazionale e conseguentemente al portale europeo.

5.3.3 Elenco metadati dati.trentino.it

La **Tabella 8** definisce il **set minimo obbligatorio** di metadati richiesto per la pubblicazione **dei dataset sul catalogo dati.trentino.it**. L'obbligatorietà è determinata dal fatto che la mancanza di tutti o alcuni di queste informazioni rende molto difficile ed incerta la possibilità di riutilizzo dei dati da parte di soggetti esterni.

83 https://joinup.ec.europa.eu/asset/dcat_application_profile/description

84 Il suo caso d'uso di base è quello di permettere una ricerca tra i vari portali open data europei in modo da permettere di trovare insieme i dati e rendere i dati del settore pubblico meglio ricercabili sia in modo transfrontaliero che intersettoriale, attraverso lo scambio di descrizioni di insiemi di dati tra portali dati.

Tabella 8 Metadati set minimo obbligatorio per i dataset

Metadato	Descrizione
Identificativo	Indica l'URI identificativa (non obbligatoriamente l'URI fisica) per il dataset.
Titolo dataset	Nome del dataset comprensibile da un utente generico. E' il nome che viene mostrato all'utente quando consulta il dataset (es. "Limiti amministrativi comunali del Trentino")
Titolare (rights holder)	Ente titolare del dataset, ovvero il soggetto o l'organizzazione che detiene e gestisce i diritti sul dataset, spesso coincide con il creatore
Tema/Categorie dataset	Categorie e temi associabili al dataset
Autore (author)	Indica il soggetto che ha prodotto il dataset. Spesso coincide con il editore/pubblicatore
Editore (pubblicatore/editore)	Ente che pubblica il dataset e che ne costituisce il referente (comprende un nome ed una email di contatto)
Punto di contatto	Nel caso non sia il titolare o l'editore, è il soggetto che agisce da punto di contatto per il dataset e per la sua distribuzione
Descrizione	Descrizione testuale del contenuto informativo del dataset in linguaggio naturale (testuale). Per esempio, nel caso di dato tabellare, gli attributi corrispondono alle colonne della tabella.
Data di rilascio	Indica la data di rilascio del dataset. Il formato della data segue le specifiche dello standard ISO 8601, es. <i>aaaa-mm-gg</i> o <i>aaaa-mm-ggThh:mm:ssZ</i>
Data di modifica	Indica la data di ultimo aggiornamento dei dati del dataset. Il formato della data segue le specifiche dello standard ISO 8601, es. <i>aaaa-mm-gg</i> o <i>aaaa-mm-ggThh:mm:ssZ</i>
Home Page	Pagina web (URI) in cui il dataset è messo a disposizione (fuori dal catalogo)
Frequenza di Aggiornamento	Indica la frequenza di aggiornamento dei dati del dataset. (tra i seguenti valori: giornaliera, settimanale, mensile, non prevista, continua). Vocabolario https://it.wikipedia.org/wiki/Resource_Description_Framework https://it.wikipedia.org/wiki/SPARQL
Zona Geografica	Indica la caratterizzazione spaziale dei dati. E' importante fare uso della notazione WKT(Well Known Text)con coordinate Latitudine e Longitudine espresse in WGS84. Se non possibile allora va indicato il toponimo. Vocabolari: http://publications.europa.eu/mdr/authority/continent/ , http://publications.europa.eu/mdr/authority/place/ , http://publications.europa.eu/mdr/authority/country/ , http://www.geonames.org/
Tag/Parole chiave	Lista di termini associati al dataset che lo descrivono, e rendono possibile la sua classificazione e la sua ricerca
Estensione temporale	Indica una caratterizzazione temporale dei dati. Data inizio e data di fine validità. Il formato della data segue le specifiche dello standard ISO 8601, es. <i>aaaa-mm-</i>

	<i>gg o aaaa-mm-ggThh:mm:ssZ</i>
Lingua	Indica la lingua con cui sono espressi i dati. Si consiglia di far riferimento al vocabolario controllato RFC 4646. Vocabolario: http://publications.europa.eu/mdr/authority/language/
Distribuzione	URI. Ad ogni dataset deve essere associata o collegata una rappresentazione fisica, alla quale è associata una licenza. Possono essere associate diverse licenze a diverse distribuzioni
Versione	Versione del dataset

La **tabella 9** definisce altri metadati che sono richiesti anche dal catalogo dati.trentino.it, in questo caso come **set obbligatorio** (o in via di modifica in questa direzione) **per la distribuzione**, ovvero per la rappresentazione fisica del dataset. Ogni dataset può avere più distribuzioni.

Tabella 9 Metadati: set obbligatorio per la distribuzione

Metadato	Descrizione
Titolo	Nome della distribuzione comprensibile da un utente generico.
Formato Distribuzione	Identifica il formato del file utilizzato dalla distribuzione del dataset. I formati possibili sono definiti nella tabella formati descritta precedentemente. Come vocabolario fare riferimento a http://publications.europa.eu/mdr/authority/file-type/
Codifica Caratteri Distribuzione	Identifica il codice dell'insieme di caratteri utilizzato dalla distribuzione del dataset. Si vedano le tabelle ISO 8859 (es. ISO 8859-15), gli standard Unicode (es. UTF-8) o quella Microsoft Windows (es. Windows-1250)
Dimensione in Byte Distribuzione	Numero indicante la dimensione della distribuzione del dataset espressa in byte. Rilevante se la distribuzione supera i 100 MB
Licenza	Vocabolario raccomandato http://publications.europa.eu/mdr/authority/file-type/ (da utilizzare anche per la proprietà ID).
URL EndPoint Accesso/Distribuzione	Nel caso la distribuzione del dataset abbia un endpoint di accesso, indica l'indirizzo del endpoint (es., REST endpoint) a cui possiamo accedere e/o sottoporre query sul dataset.
Data Modifica Distribuzione	Indica la modifica del file fisico che contiene i dati del dataset (non il dato). Il formato della data segue le specifiche dello standard ISO 8601, es. <i>aaaa-mm-gg o aaaa-mm-ggThh:mm:ssZ</i>

5.4 Altri standard, vocabolari e ontologie

Si è evidenziato come la definizione e l'uso di vocabolari strutturati e di ontologie sia un passaggio chiave per avere dei dati e dei metadati comprensibili da sistemi automatici. Una ontologia è una **rappresentazione formale, condivisa ed esplicita di un dominio di interesse, visto da un punto di vista più o meno generale**. Per esempio, nel dominio degli atti amministrativi un'ontologia potrebbe definire univocamente i vari tipi di atti amministrativi possibili o i rapporti che sussistono fra una delibera ed una serie di leggi, altre delibere, circolari, ecc.. **Qualora si riesca a trovare un accordo** per definire in modo univoco questo sistema di significati e relazioni, i programmi informatici possono poi usare quella ontologia per una varietà di scopi, tra cui il ragionamento tassonomico; la classificazione, e in generale la risoluzione di problemi. L'adozione di ontologie condivise **a livello internazionale/europeo sembra essere un passo fondamentale e probabilmente ineludibile per l'effettiva interoperabilità semantica delle amministrazioni europee**.

Si tratta di una direzione di sviluppo della ricerca ICT interessante ma anche molto complessa con ancora oggetto di ricerca e dibattito, ma con alcuni ambiti più esplorati di altri, ai quali è comunque utile fare riferimento. Esistono infatti oggi alcune ontologie di riferimento, utili e importanti da conoscere perché permettono di operare delle scelte nella gestione dei dataset, tali da valorizzare al meglio il potenziale di standardizzazione dei sistemi informativi.

La **tabella 10** ne presenta, a titolo esemplificativo, alcune di riferimento generale, evidenziando nella descrizione gli ambiti a cui si applicano e come in effetti alcune soluzioni che queste propongono permettono di "arricchire in modo standardizzato" di informazioni accessorie i dataset.

Tabella 10 esempio di alcune ontologie fra quelle più usate in diversi ambiti

Nome	Riferimento	Descrizione
XSD (XML Schema Definition)	http://www.w3.org/2001/XMLSchema#	Permette di esprimere in termini di vincoli sulla struttura e il contenuto dei documenti XML , al di sopra e al di là dei vincoli sintattici di base imposti dal formato XML stesso
RDFS (RDF Schema)	http://www.w3.org/2000/01/rdf-schema#	Fornisce un insieme di classi con le proprietà necessarie che abilita un modello estendibile di rappresentazione della conoscenza basato su RDF
SCHEMA.ORG	http://schema.org/	Insieme termini standard per descrivere la tipologia dei dati all'interno di pagine e documenti web (data markup) . Mira a permettere una migliore indicizzazione di pagine web da parte di servizi di ricerca come Google, Bing, etc.

DCAT (Data Catalogue Vocabulary)	http://www.w3.org/ns/dcat#	Vocabolario standard per descrivere cataloghi e facilitare l'interoperabilità tra i cataloghi di dati pubblicati sul Web
DCT (Dublin Core Schema)	http://purl.org/dc/terms/	Piccolo insieme di vocabolari che possono essere utilizzati per descrivere sia risorse Web (video, immagini, pagine web , ecc) che risorse fisiche, come libri o CD, e altri oggetti .
FOAF (Friend of a Friend Vocabulary)	http://xmlns.com/foaf/0.1/	Ontologia che descrive principalmente le relazioni fra le persone e con oggetti
SKOS (Simple Knowledge Organization System)	http://www.w3.org/2004/02/skos/core#	Raccomandazione del W3C progettata per la rappresentazione di vocabolari strutturati , quali thesauri, schemi di classificazione, tassonomie e per svariati altri tipi di vocabolario controllato strutturato.
ADMS (Asset Description Metadata Schema)	http://www.w3.org/ns/adms#	Vocabolario di metadati comune per descrivere gli standard di interoperabilità per il Web
DOLCE (Descriptive Ontology for Linguistic and Cognitive Engineering)	http://www.loa.istc.cnr.it/old/Papers/gai0122010.pdf	Ontologia fondazionale. DOLCE è stata sviluppata con lo scopo di catturare le categorie ontologiche che emergono nel linguaggio naturale e nel senso comune
BFO (Basic Formal Ontology)	http://ifomis.uni-saarland.de/bfo/	E' un ontologia di livello superiore che è stato progettata per il recupero delle informazione, l'analisi e l'integrazione in settori scientifici e altri domini applicativi
UFO (Unified Foundational Ontology)	http://oxygen.informatik.tu-cottbus.de/drupal7/ufo/	E' una ontologia che descrive i concetti molto generali che sono presenti in tutti i domini di conoscenza

5.5 Alcuni standard di riferimento su specifici ambiti settoriali

SDMX (Statistical Data and Metadata eXchange): si tratta di uno standard ISO per lo scambio di dati statistici basato su sintassi XML. Esso implementa un modello per la rappresentazione di dati multidimensionali e dunque descrive la struttura di un particolare “dataflow” attraverso un insieme di dimensioni (es. territorio o tempo), un insieme di attributi (e.g., unità di misura) e le associate code-list (o classificazioni). Sebbene SDMX sia nato come modello per lo scambio di dati, esso viene anche usato per la rappresentazione dei dati.

SPDX (Software Package Data Exchange): è un formato di file utilizzato per documentare informazioni sulle licenze software in base alle quali un certo pacchetto software è distribuito. SPDX cerca di standardizzare il modo in cui le organizzazioni pubblicano i loro metadati riguardanti le licenze e componenti software <http://spdx.org/rdf/terms#>

INSPIRE (Infrastructure for Spatial Information in Europe): nasce come direttiva della commissione europea per mettere a disposizione in modo armonico i dati geo-spaziali e i servizi associati tramite geo-portali e altri punti di accesso. Di fatto, definisce un'ontologia nata proprio con l'obiettivo di migliorare l'interoperabilità tra i sistemi geo-spaziali.

GeoDCAT-AP: è un'estensione di DCAT-AP per descrivere insiemi di informazioni geospaziali. Esso fornisce una sintassi RDF di binding per l'unione di elementi di metadati definiti nello standard ISO 19115:2003 e nella direttiva INSPIRE. Il suo caso d'uso di base è quello di rendere le informazioni spaziali migliori per la ricerca su portali dei dati in cross-settoriale e cross-regionale. **La specifica GeoDCAT-AP non rimpiazza le direttive INSPIRE.**

vCARD: è uno standard di formato di file per biglietti da visita elettronici definito dal W3C. I file in formato vCard sono spesso allegati ai messaggi di posta elettronica ma possono essere scambiati in altri modi, come ad esempio sul World Wide Web o messaggistica istantanea e possono contenere nome e l'indirizzo informazioni, numeri di telefono, indirizzi e-mail, URL, loghi e svariate altre informazioni. (<http://www.w3.org/2006/vcard/ns#>)

CIDOC-CRM: Standard ISO per la documentazione di dati provenienti dal dominio culturale. Trattasi di un'ontologia sviluppata dal 1995, e che viene continuamente curata da uno speciale gruppo di lavoro. Permette di documentare la vita e la gestione “dell'oggetto cultural”, ma la sua flessibilità ne permette l'uso in svariati contesti. È stata via via espansa grazie allo sviluppo di diverse estensioni per permettere la documentazione di oggetti digitali (CRMdig), osservazioni scientifiche (CRMsci), dati geografici (CRMgeo), archeologici (CRMarcheo), così come le argomentazioni di riferimento (CRMinf). Importante per il dominio bibliotecario è lo sviluppo insieme all'associazione internazionale delle biblioteche (IFLA) di FRBRoo, un'ontologia sviluppata a partire e come estensione di CIDOC-CRM che permette di descrivere i vari processi editoriali. (http://www.iso.org/iso/catalogue_detail?csnumber=57832)

5.6 Considerazioni sui Linked Data

Come definito anche nelle linee guida AgID, e accennato precedentemente, i dati di tipo linked, tipicamente resi disponibili in formato RDF, sono una modalità di **pubblicazione di dati arricchiti** da informazioni che li rendono più strutturati, più facilmente collegabili nel web ad altre informazioni. I linked data utilizzano tecnologie e standard web aperti quali il protocollo di comunicazione HTTP⁸⁵ e URI⁸⁶, si tratta di stringhe che identificano univocamente una risorsa generica dentro il web e ne estendono l'applicazione per fornire informazioni che possano essere lette e comprese da programmi software. Queste informazioni aumentano il valore dei dati in quanto **forniscono punti di riferimento importanti per la navigazione web e quindi per trovare informazioni utili e riutilizzabili collegabili a quel dataset** dentro il word wide web (WWW).

I dataset possono essere arricchiti in modalità linked perlomeno a 3 livelli: 1) dataset; 2) formati di distribuzione; 3) singola entità/oggetto contenuto nel dataset.

La letteratura parla di linked data quando i dati sono tali da rispettare le **seguenti 4 regole di base dei linked data**:

- **utilizzano degli indirizzi Web (URI)** come identificativi (nomi) delle “risorse” (alternativamente “oggetti” o “entità”);
- le URI sono tali che, attraverso la congiunzione con l'uso del protocollo HTTP, permettono di **individuare degli oggetti nel web** (dereferenziazione);
- garantiscono che quando viene cercata una URI, sia **sempre restituita una informazione utilizzabile**, usando gli standard facendo RDF e SPARQL;
- **includono anche i collegamenti ad altre URI**, così da permettere a chi cerca di scoprire nuovi collegamenti ad altre informazioni, ove possibile.

Ogni catalogo di dati aperti, adeguatamente metadato, fornisce già le URI del protocollo HTTP://, mentre è più complesso garantire l'uso di linguaggi quali SPARQL, usando il paradigma RDF⁸⁷ delle triple concatenate, in cui agli “oggetti” è assegnata una URI univoca su Web (principio 1 dei linked data). Conseguentemente tale URI può essere utilizzata per effettuare accessi diretti alle informazioni relative a quella entità (principio 2 dei linked data). I dati, in tale formato, sono chiamati infatti “linked” per la possibilità di riferenziarsi, o collegarsi, tra loro (principio 3 dei linked data). Nel riferenziarsi, si usano delle relazioni (“link”) che hanno un significato ben definito e che spiegano quindi il tipo di legame che intercorre tra le due entità coinvolte (principio 4 dei linked data).

85 https://it.wikipedia.org/wiki/Hypertext_Transfer_Protocol

86 https://it.wikipedia.org/wiki/Uniform_Resource_Identifier

87 https://it.wikipedia.org/wiki/Resource_Description_Framework

In sostanza, **il beneficio più evidente nell'utilizzo dei Linked Data è il poter sfruttare la semantica dei dati e le infrastrutture Web per abilitare la scoperta e l'inferenza di nuove informazioni**, con conseguenze positive e dirette sia per gli utenti che per gli sviluppatori.

In tale direzione, in Trentino, a partire dal 2012, sono stati pubblicati **a titolo sperimentale più di 135 dataset in formato Linked Data, in formato RDF**. In particolare si tratta di dati di tipo geografico appartenenti al Sistema Informativo Ambiente e Territorio della Provincia autonoma di Trento. Tali linked data sono il risultato di una prima attività svolta con i centri di ricerca del territorio, che ha permesso di evidenziare sia i requisiti tecnologici implicati nella produzione di dati in questa modalità e sia per valutare l'effettivo interesse sui dati in tale formato dai potenziali riutilizzatori (es. aziende). Un esempio **di riferimento per i linked data è quello offerto da DBPedia**, un progetto collaborativo che punta a trasformare in formato linked data tutte le informazioni contenute su Wikipedia e a renderle disponibili per agevolare il collegamento con altre banche dati. In particolare lo sviluppo del ramo italiano (DBPedia Italia) ha anche visto il coinvolgimento di alcuni attori del sistema delle imprese Trentino.

Sia nel caso del geocatalogo linked ispirato alle direttive Inspire, che il caso di Dbpedia, lo sforzo di attività per arricchire in modalità linked i dati si applica a un dominio ampio e molto trasversale, che può essere premiante, ma attualmente però, la domanda di linked data, al di fuori degli ambiti della ricerca, è risultata per nostra esperienza ancora molto embrionale, con limitate applicazioni, per quanto si riscontri negli ultimi tempi un'accelerazione di interesse grazie anche alcune aziende leader di mercato e di quanto sviluppatosi attorno al cluster della Semantic Valley di Trento.

Su questi temi il **sistema pubblico Trentino si pone come sperimentatore**⁸⁸, all'interno della cornice dei laboratori territoriali, con i centri di ricerca del territorio e con le aziende che intendono innovare in tale ambito, in collegamento con altre regioni italiane e in collegamento con AgID.

5.7 Standard nella struttura dei contenuti della comunicazione web del Sistema Trentino

I contenuti, i dati e le informazioni del servizio pubblico sono resi fruibili o attraverso i portali web, nel caso dei dati aperti, o tramite protocolli di scambio fra sistemi informativi più o meno vincolati da chiavi di accesso e abilitazioni.

⁸⁸ In tale ambito la Provincia autonoma di Trento è con progetto europeo FUSEPOOL P3#, insieme alla Regione Toscana e ad alcune aziende molto nell'ambito dei linked data, sta sperimentando direttamente la trasformazione di alcuni open data, pubblicati nel catalogo dati.trentino.it. La piattaforma FUSEPOOL P3, ancora sperimentale, integra lo stato dell'arte degli strumenti software del settore, come OpenRefine, OpenLink Virtuoso, Apache Stanbol, Squebi #e Pundit, e altri, e si pone l'obiettivo di realizzare in modo completo ed utilizzabile la piattaforma linked data LDP# standard, specificata da W3C. <http://tinyurl.com/pczpmwv>

E' indubbio che **i portali web siano oggi l'interfaccia prioritaria di contatto massivo fra il servizio pubblico e i cittadini**, le porte attraverso le quali permettere il riutilizzo dei dati del servizio pubblico. Tutti i dati presenti sui portali istituzionali del servizio pubblico sono "aperti by default" - salvo altra licenza - e quindi **già riutilizzabili anche a scopo commerciale**. Per questo motivo i portali non solo sono il luogo dove si trovano le informazioni ma anche il luogo dove queste **possono venire strutturate in modo tale da renderne più facile il riutilizzo** da parte di computer e piattaforme, o in altre parole, i sistemi di gestione dei contenuti dei portali, **i content management system (CMS)**, dal punto di vista open data, **sono uno spazio privilegiato per "valorizzare il patrimonio informativo pubblico che essi stessi veicolano"**.

Grazie ad una importante messa a sistema dell'esperienza maturata nell'ambito del Progetto ComunWeb del Consorzio dei Comuni Trentini, che ha offerto ai suoi associati una piattaforma di comunicazione web unica e soprattutto attenta alla strutturazione delle informazioni, il sistema provinciale **può oggi riutilizzare queste esperienze** e applicarle nella revisione del proprio stesso portale istituzionale.

Il sistema trentino nei prossimi mesi, in occasione della migrazione della piattaforma di gestione dei contenuti web (CMS) del portale istituzionale della PAT, adotterà la tecnologia della piattaforma adottata da ComunWeb e ad oggi da alcuni siti tematici come ad esempio Trentino Cultura, in modo tale da avere **una comune e semplice architettura dell'informazione in grado di abilitare l'interoperabilità informativa tra i siti web, il riutilizzo dei dati e la loro pubblicazione in modalità open data**.

Per architettura dell'Informazione⁸⁹ si intende la struttura organizzativa **logica e semantica delle informazioni, dei contenuti, dei processi e delle funzionalità di un sistema o ambiente informativo**. L'architettura dell'informazione costituisce la struttura portante di un qualsiasi insieme di contenuti e dati destinati alla fruizione e integrabile anche con le informazioni dei processi svolge un ruolo chiave nel definire il reale grado di fruibilità e di usabilità di un sistema informativo per l'utente finale. L'architettura sviluppata si basa su un insieme di **classi informative**⁹⁰, il cui primo insieme è stato definito attraverso l'esperienza di 150 comuni accumulata nel Progetto ComunWeb.

Adottando ed arricchendo le stesse strutture informative, si sta realizzando l'interoperabilità applicativa delle informazioni web di tutto il sistema degli enti locali del territorio provinciale, trasformando i portali web da vetrina di contenuti in depositi/contenitori/sistemi di informazioni strutturate, interrogabili con interfacce web (API), in formati di tipo aperto e direttamente presenti anche sulla piattaforma dati.Trentino.it senza ulteriori costi aggiuntivi.

Tale struttura informativa è messa a disposizione in modalità aperta e se ne incoraggia il riuso e l'adozione in modo tale da estenderne ulteriormente l'adozione, aumentando la qualità generale e la possibilità di riuso dei dati prodotti.

E' intuitivo capire come **con questo passaggio si potranno ottenere notevoli risultati** sul piano della standardizzazione, delle definizioni delle entità rilevate, dell'aggregazione delle informazioni in banche dati interoperabili alla fonte, in possibilità di arricchire i dati in modo coordinato tramite una

89 https://it.wikipedia.org/wiki/Architettura_dell%27informazione

90 <http://www.comunweb.it/openpa/classes>

medesima architettura al momento in cui le informazioni si generano, con importanti ricadute in termini di integrazione delle informazioni fra i due attori del servizio pubblico “comune” e “provincia” che grazie a queste sinergie potranno porre al centro della propria comunicazione istituzionale il cittadino dentro una rete di informazioni non ridondanti, coerenti, coordinate perché interoperabili e coerentemente strutturate nel momento in cui i dati sono generati nei portali.