

Agenda
digitale
dell'Umbria



Regione Umbria

Programma Trasversale #opendata
Modello Operativo Open Data (MOOD) Umbria
Versione 1.0

Indice

Open Data.....	4
Modello Organizzativo.....	7
Modello Operativo (i processi).....	9
1. Identificazione.....	10
1.1 Segnalazione interna.....	11
1.2 Segnalazione esterna.....	11
1.3 Identificazione del Referente Tecnico del Dataset.....	12
1.4 Attribuzione di un Referente Tecnico del Dataset.....	12
1.5 Compilazione della Scheda Informativa.....	12
1.6 Verifica finale della Scheda Informativa.....	13
1.7 Inserimento dei dati nella pipeline.....	13
2. Analisi e Pubblicazione.....	14
2.1 Analisi di riservatezza e titolarità dei dati.....	15
2.2 Archiviazione.....	16
2.3 Definizione degli obiettivi e opportunità di pubblicazione.....	17
2.4 Verifica Licenza.....	18
2.5 Analisi di completezza e qualità minima.....	18
2.6 Completamento del dataset.....	19
2.7 Analisi del formato.....	19
2.8 Dematerializzazione.....	20
2.9 Analisi della struttura.....	21
2.10 Strutturazione del dataset.....	21
2.11 Analisi dei dati personali.....	22
2.12 Tecniche di trattamento dei dati personali.....	24
2.13 Conversione in formato aperto.....	25
2.14 Definizione dei metadati.....	26
2.15 Applicazione della licenza d’uso.....	26
2.16 Inserimento nel Dataset catalogue.....	27
2.17 Validazione Dataset.....	28
2.18 Pubblicazione.....	28
2.19 Feedback.....	28
3. Monitoraggio.....	29
3.1 Valutazione interna.....	30
3.2 Valutazione esterna.....	31
4. Mantenimento.....	33
4.1 Analisi di aderenza agli obiettivi.....	34
4.2 Diffusione.....	34
4.3 Aggiornamento.....	36
4.4 Rimozione.....	37
4.5 Feedback.....	37
Allegati.....	38
Glossario.....	39

Il presente documento si avvale del riuso dei documenti:

- Linee Guida Nazionali per la Valorizzazione del Patrimonio Informativo Pubblico (anno 2014), AgID – rilasciato in licenza CC-BY-SA
- Vademecum – Open Data, Formez PA
- Modello e Standard per la Gestione degli Open Data, Comune di Guidonia Montecelio - rilasciato in licenza CC-BY-SA
- ODMC. Open Data Management Cycle. Modello Aperto per la Gestione degli Open Data, Present S.p.A. (www.odmc.org) - rilasciato in licenza CC-BY-SA
- Ten Open Data Guidelines - Transparency International Georgia - rilasciato in licenza CC-BY

Il documento *Modello Operativo Open Data (MOOD) Umbria* è rilasciato sotto licenza CC-BY-SA

Il portale di riferimento *Open Data Umbria* è disponibile all'indirizzo dati.umbria.it

Per ulteriori informazioni scrivere a opendata@regione.umbria.it

Open Data

Distribuire i dati pubblici in un formato aperto e libero da restrizioni sia dal punto di vista dell'accesso che dell'integrazione e del riutilizzo, rappresenta il presupposto di base affinché possa svilupparsi un vero e proprio processo di collaborazione tra le istituzioni e la comunità dei cittadini sulle scelte di governo, anche attraverso la rielaborazione in forma nuova e diversa dei dati messi a disposizione.

Mediante strategie di apertura dei dati pubblici, i cittadini non sono più soltanto consumatori passivi di informazioni messe a disposizione dalle amministrazioni. Hanno invece l'opportunità di riutilizzare e integrare i dati messi loro a disposizione, fino a sviluppare servizi e applicazioni a vantaggio dell'intera comunità di utenti, che vanno ad affiancarsi a quelli creati centralmente dalle istituzioni.

In riferimento al documento *Ten Open Data Guidelines - Transparency International Georgia*, i dati devono essere:

1. **Completi:** I dati devono comprendere tutte le componenti (metadati) che consentano di esportarli, utilizzarli online e offline, integrarli e aggregarli con altre risorse e diffonderli in rete.
2. **Primari:** Le risorse digitali devono essere strutturate in modo tale che i dati siano presentati in maniera sufficientemente granulare, così che possano essere utilizzate dagli utenti per integrarle e aggregarle con altri dati e contenuti in formato digitale;
3. **Tempestivi:** Gli utenti devono essere messi in condizione di accedere e utilizzare i dati presenti in rete in modo rapido e immediato, massimizzando il valore e l'utilità derivanti da accesso e uso di queste risorse;
4. **Accessibili:** I dati devono essere resi disponibili al maggior numero possibile di utenti senza barriere all'utilizzo, quindi preferibilmente attraverso il solo protocollo Hypertext Transfer Protocol (HTTP) e senza il ricorso a piattaforme proprietarie. Devono essere inoltre resi disponibili senza alcuna sottoscrizione di contratto, pagamento, registrazione o richiesta.
5. **Leggibili da computer:** Per garantire agli utenti la piena libertà di accesso e soprattutto di utilizzo e integrazione dei contenuti digitali, è necessario che i dati siano machine-readable, ovvero processabili in automatico dal computer.
6. **In formati non proprietari:** I dati devono essere codificati in formati aperti e pubblici, sui quali non vi siano entità (aziende o organizzazioni) che ne abbiano il controllo esclusivo. Sono preferibili i formati con le codifiche più semplici e maggiormente supportati.
7. **Liberi da licenze che ne limitino l'uso.** I dati aperti devono essere caratterizzati da licenze che non ne limitino l'uso, la diffusione o la redistribuzione.
8. **Riutilizzabili:** Affinché i dati siano effettivamente aperti, gli utenti devono essere messi in condizione di riutilizzarli e integrarli, fino a creare nuove risorse, applicazioni e servizi di pubblica utilità.

9. **Ricercabili:** I dati devono essere facilmente identificabili in rete, grazie a cataloghi e archivi facilmente indicizzabili dai motori di ricerca.
10. **Permanenti:** Le peculiarità fino ad ora descritte devono caratterizzare i dati nel corso del loro intero ciclo di vita.

In modo particolare, per la Pubblica Amministrazione italiana la normativa prevede che il dato debba possedere anche delle caratteristiche di **qualità** che assicurino l'integrità, il costante aggiornamento, l'indicazione della provenienza, la riutilizzabilità, etc. (art. 6 D.L. 14 marzo 2013, n. 33 “*Riordino della disciplina riguardante gli obblighi di pubblicità, trasparenza e diffusione di informazioni da parte delle pubbliche amministrazioni*”).

Per distinguere i diversi formati utilizzabili nella codifica dei set di dati, è stato proposto (dal W3C) un modello di catalogazione che li classifica in base alle loro caratteristiche su una scala di valori da 1 (una stella) a 5 (cinque stelle):

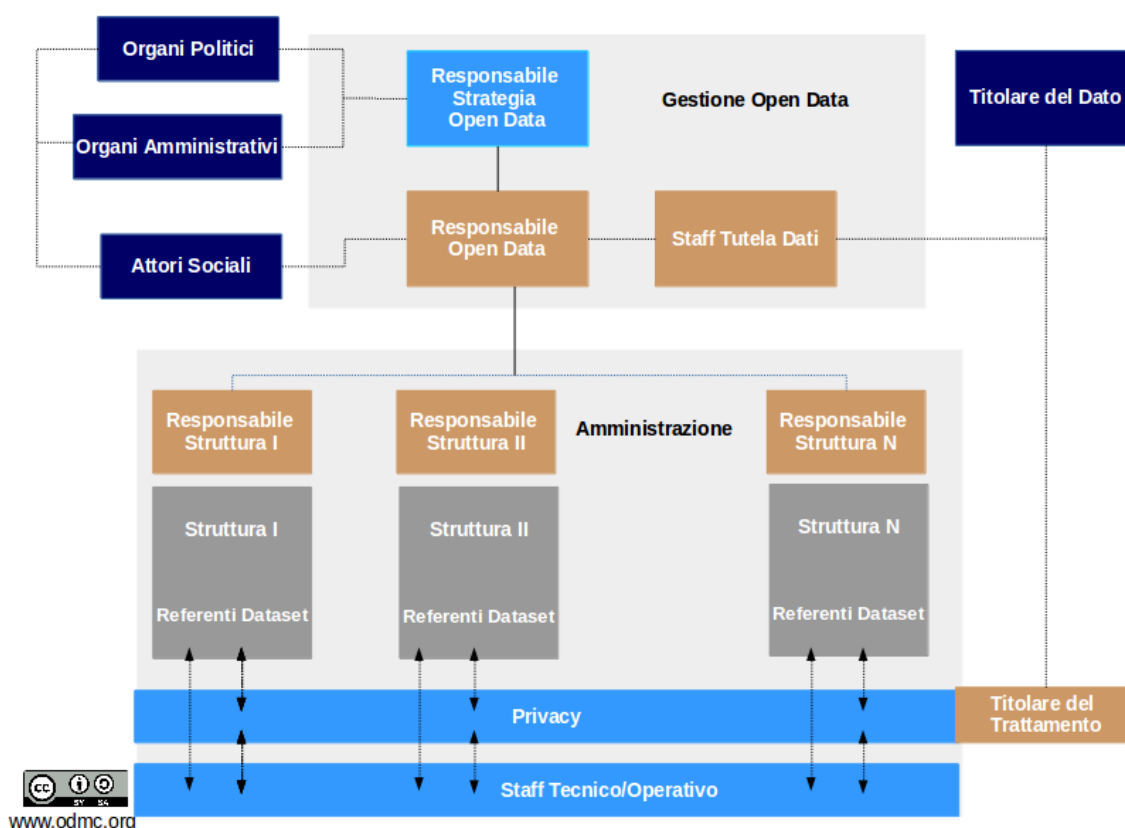
- (★) Una Stella. È il livello base, sono *dati statici* e non strutturati: ad esempio un'immagine in formato grezzo (formati come .gif, .jpg, .png), un documento in formato Microsoft Word, un file in formato Adobe PDF. Una sola stella indica la semplice disponibilità di una informazione e di un dato on line, in un formato qualsiasi, purché distribuito con licenza aperta. I dati distribuiti in questo formato sono leggibili e stampabili dagli utenti, possono essere conservati localmente su un PC e sono semplici da pubblicare. Tuttavia non sono un formato aperto in quanto non è possibile effettuare su di essi alcuna elaborazione.
- (★★) Due Stelle. Questo livello indica *dati strutturati* e codificati in un *formato proprietario* (ad esempio un documento in formato Microsoft Excel). Due stelle indicano la possibilità di effettuare elaborazioni sui dati, a patto di disporre del software necessario a gestire un file codificato con un formato proprietario. I dati caratterizzati dalle due stelle non sono un formato aperto in quanto per elaborarli è necessario un software proprietario, tuttavia di norma possono essere convertiti – essendo dati strutturati – in dati aperti.
- (★★★) Tre Stelle. Questo livello indica *dati strutturati* e codificati in un *formato aperto*. Ad esempio il formato .csv (Comma Separated Values) al posto del formato Microsoft Excel utilizzato nel caso precedente. Tre stelle indicano la possibilità di effettuare elaborazioni sui dati senza esser costretti ad utilizzare software proprietario. Quello caratterizzato dalle tre stelle è il formato più semplice di dati aperti.
- (★★★★) Quattro Stelle. Questo livello indica *dati strutturati* e codificati in un *formato aperto* che sono *dotati di un URI* che li rende indirizzabili sulla rete e quindi utilizzabili direttamente online, attraverso l'inclusione in una struttura basata sul modello RDF (Resource Description Framework). Quattro stelle indicano quindi il fatto che il singolo dato di un dataset, disponibile on line in un formato aperto (tipicamente XML/RDF) può essere richiamato attraverso un URL (Uniform Resource Locator) specifico. Ciò consente di puntare al dato o ad un insieme di dati da una applicazione o accedervi dall'interno di un programma che può poi elaborarlo in vari modi. Si pensi, ad esempio, a un dataset contenente gli indirizzi dei monumenti di una città opportunamente codificati: da qualsiasi software – finanche dal browser – è possibile collegarsi all'URL che indica il singolo monumento, potendolo ad esempio georeferenziare su una mappa.
- (★★★★★) Cinque Stelle. Questo livello indica quelli che vengono definiti *Linked Open Data* (LOD). Quei dati aperti, cioè, che – dal punto di vista del formato – oltre a rispondere alle caratteristiche indicate al punto precedente (classificazione a quattro stelle) presentano anche, nella struttura del dataset, collegamenti ad altri dataset. Grazie al modello di

descrizione dei dati RDF, è possibile collegare dinamicamente tra loro più dataset, incrociando così informazioni provenienti da fonti diverse, eventualmente gestite da diverse Amministrazioni. Si pensi ad esempio al caso del dataset contenente gli indirizzi dei monumenti di una città alla quale si è fatto riferimento in precedenza. Tale dataset, pubblicato dall'Amministrazione A, potrebbe essere collegato al dataset di un altro Ente – l'Amministrazione B – che dispone dell'archivio delle opere d'arte presenti all'interno di ogni monumento. In questo caso un sistema software potrebbe, dopo aver chiesto all'utente i suoi gusti e le sue preferenze in fatto di arte, strutturare un percorso georeferenziando i monumenti a partire dalle informazioni presenti nel dataset disponibile presso l'Amministrazione A, dopo averli preselezionati in base alle opere d'arte in essi contenute, identificate nel dataset dell'Amministrazione B.

★	Dati statici			
★ ★ ★ ★ ★	Dati strutturati			
★ ★ ★ ★ ★ ★ ★ ★ ★	Linked data			

Modello Organizzativo

Al fine di costruire un modello organizzativo efficiente per la produzione e gestione del dato pubblico di tipo aperto è auspicabile individuare, all'interno dell'amministrazione, delle figure strategiche per una corretta implementazione dello stesso. E' evidente che, dipendentemente dalle dimensioni delle amministrazioni, alcune delle figure professionali sotto indicate possono coincidere o possono essere ulteriormente distinte.



Dall'analisi dei documenti “Linee Guida Nazionali per la Valorizzazione del Patrimonio Informativo Pubblico (anno 2014)”, pubblicate dall'Agenzia per l'Italia Digitale, e “ODMC, Open Data Management Cycle, Modello Aperto per la Gestione degli Open Data” (recepito in parte nelle stesse linee guida AgID a seguito di consultazione pubblica) si evince la necessità di individuare figure chiare all'interno dell'ente pubblico, che mettano in moto il processo di pubblicazione e gestione degli Open Data. Di seguito si riportano tali figure:

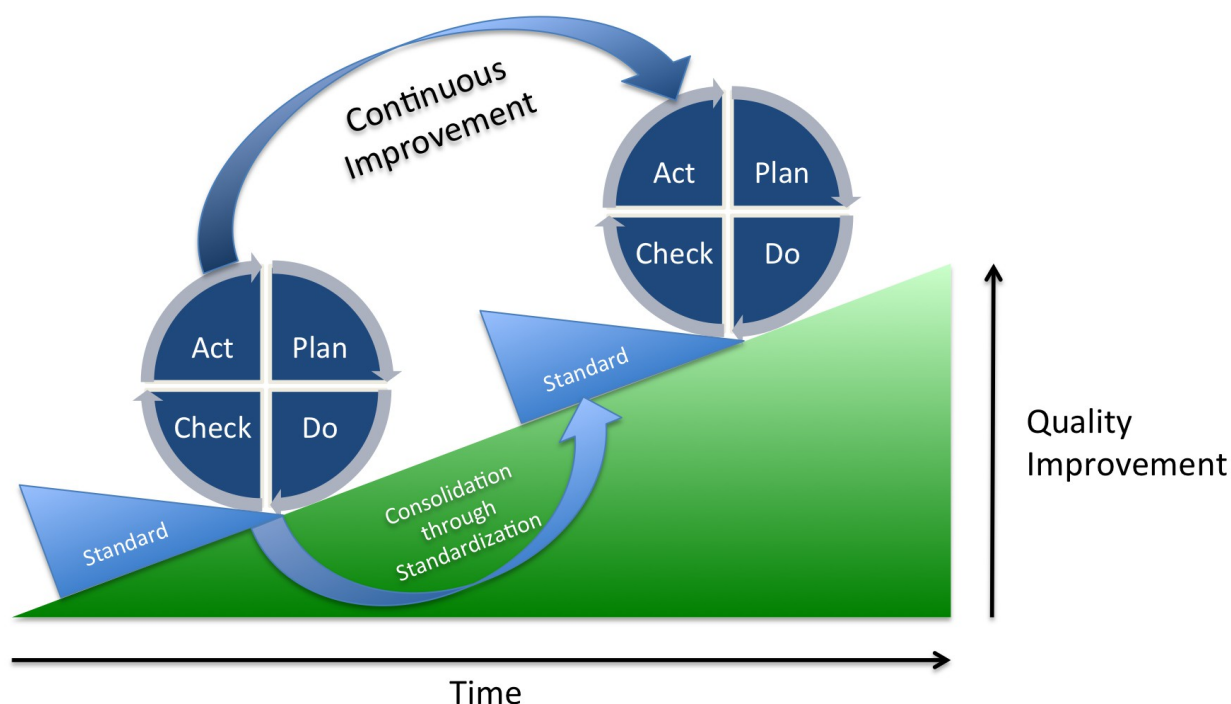
- **Responsabile Strategia Open Data:** il Responsabile della formulazione della Strategia Open Data, e successivamente della sua attuazione, dovrà essere una figura apicale della struttura dirigenziale dell'Amministrazione. Tale responsabile è, di norma, un Direttore (Direttore Generale, Segretario Generale). Opera in stretta collaborazione con i Dirigenti dei vari settori (Responsabili delle Strutture) e attua il collegamento tra gli Organi Politici e gli

Organi Amministrativi, con la la funzione duplice di attuare la strategia e le politiche degli Open Data e di garantire che la Strategia Open Data sia sempre allineata con quanto approvato a livello politico. IL /I responsabile/i della strategia attua gli indirizzi dell'Agenda Digitale e condivide le politiche per la pubblicazione degli opendata con la *Taskforce* (ex DGR 148/2012).

- **Responsabile Open Data:** Ha la responsabilità dell'intero ciclo di pubblicazione, di mantenimento e dell'aggiornamento degli Open Data.
- **Responsabile Struttura:** È il dirigente (o funzionario) incaricato di seguire gli Open Data presente in ogni Area/Settore dell'Ente. La sua responsabilità sarà quella di collaborare nelle decisioni sulla strategia (cosa pubblicare, quando e con quale priorità) e di implementare tali decisioni con riferimento al settore di appartenenza, fornendo al Responsabile Open Data i dataset e quant'altro necessario per avviare il processo di pubblicazione, curando gli aspetti gestionali interni al suo settore di appartenenza (autorizzazioni, firma, materiali, eventuali licenze, etc) e valutando l'opportunità della pubblicazione dei dataset tenendo conto degli aspetti legati alla riservatezza e alla tutela dei dati.
- **Staff Tutela Dati:** È una figura che fornisce supporto e pareri ai responsabili di struttura in merito a problemi relativi alla pubblicazione di dati (dati riservati, personali, coperti da segreto), al copyright ed alle licenze d'uso.
- **Staff Tecnico/Operativo:** Lo staff è costituito dal personale che collabora per le attività di routine e supporto ai processi Open Data, anche su problematiche di natura tecnico/informatica. All'interno dello staff è presente anche un "*Responsabile Monitoraggi e Attività di Coinvolgimento degli Utenti*" che, avendo competenza su parametri di valutazione e monitoraggio di dati aperti, nonché sull'utilizzo di strumenti Social e di marketing territoriale, cur l'attuazione delle iniziative di monitoraggio e di coinvolgimento dei cittadini, decise dal Responsabile della Strategia Open Data. Avrà quindi capacità di analisi dei dati e di reportistica e si farà promotore di iniziative per migliorare la collaborazione dei cittadini, concordandole con gli altri attori del processo Open Data.
- **Team comunicazione:** E' composto da figure con competenze di comunicazione istituzionale e non solo, in grado di curare il dialogo con i cittadini e la comunicazione verso l'esterno e verso l'interno. Ha la responsabilità di pianificare e coordinare le azioni di coinvolgimento verso l'esterno e di validare il risultato delle iniziative di data visualization relative agli Open Data, prima che queste vengano pubblicate. Cura, inoltre, la segnalazione verso l'esterno di tutte le novità relative agli Open Data dell'amministrazione (e.g., nuovi dataset aperti, aggiornamenti significativi, etc.).
- **Attori sociali:** Gli attori sociali sono cittadini, organizzazioni, associazioni, comunità di sviluppatori o altri portatori di interesse (stakeholder) esterni all'Ente, che possono contribuire con feedback e fruire liberamente dai dati aperti.

Modello Operativo (i processi)

Nella strutturazione dei processi connessi all'apertura dei dati della Pubblica Amministrazione, è utile fare riferimento al modello **PDCA** (*Plan*: Pianificare / *Do*: Eseguire / *Check*: Controllare / *Act*: Agire). Tale modello è alla base dell'organizzazione per processi dei sistemi di qualità orientati al miglioramento continuo. In sintesi, esso consente di strutturare processi complessi orientando la struttura stessa del processo così definito verso un modello che prevede momenti di analisi, verifica e cambiamento che ne consentono un'ottimizzazione continua.

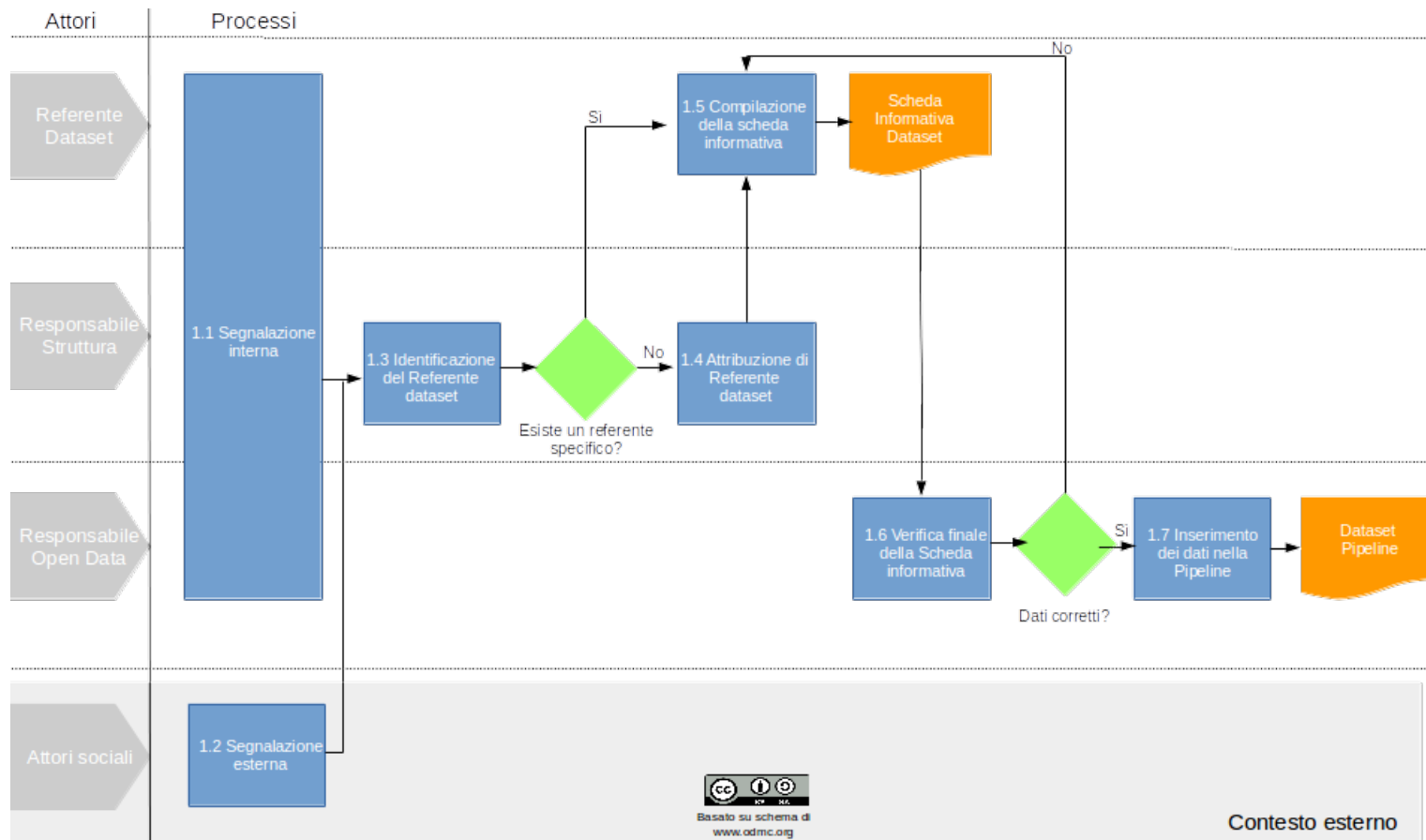


Applicando il modello in ambito Open Data, le quattro fasi del modello PDCA si possono tradurre come:

1. **Identificazione:** comprende i processi necessari all'identificazione e alla pianificazione dei potenziali dataset presenti all'interno dell'Ente;
2. **Analisi e Pubblicazione:** include i processi necessari all'esecuzione del programma di gestione dei dati, attraverso l'analisi dei dati raccolti e la definizione dei metodi di pubblicazione;
3. **Monitoraggio:** comprende i processi necessari per il controllo, lo studio e la raccolta dei risultati e dei feedback ricevuti dalle comunità interne ed esterne;
4. **Mantenimento:** include i processi necessari per rendere definitivi e/o migliorare la rappresentazione dei dati aperti.

1. Identificazione

In questo paragrafo vengono identificati i processi che, a partire dalla segnalazione (interna o esterna) di un dato, porta all'identificazione di uno o più dataset, delle relative figure di riferimento ed alla puntuale compilazione della “*Scheda Informativa*” (allegato A.1), che popola il catalogo denominato “*Dataset Pipeline*”. I processi rispettano il seguente schema:



1.1 Segnalazione interna

Tutte le risorse all'interno dell'Ente sono chiamate a segnalare in modo proattivo potenziali dati che possono diventare Open Data. In particolare, il Referente Tecnico del Dataset, ovvero la persona che lavora quotidianamente nella gestione del dato è l'attore principale e primo interessato all'apertura dei dati che gestisce. Molte volte chi segnala i dati su cui lavora quotidianamente non è riconosciuto come il Referente per il dataset. In tale caso, il processo *1.3 Identificazione del Referente Tecnico del Dataset* si occuperà di identificare il referente definitivo per la gestione del dataset o assegnarne uno se necessario.

La segnalazione può avvenire in modo strutturato (ad es.: con l'uso di un formulario o pagina intranet specifica per la segnalazione) o non-strutturato (attraverso conversazioni o segnalazioni informali tra colleghi). Può anche avvenire in modo attivo (la risorsa comunica proattivamente al Responsabile Open Data l'esistenza di un nuovo dataset) o passivo (il Responsabile Open Data si informa autonomamente sull'esistenza di nuovi dati, solitamente in concomitanza con l'avvio di nuovi progetti all'interno dell'Ente). Molto spesso il personale all'interno dell'Ente avrà difficoltà nel vedere il risultato del proprio lavoro come potenziali dataset. Chiedere loro quali informazioni vorrebbero rendere disponibili ai cittadini, o di immaginare quali applicazioni mobili potrebbero essere utili può essere un modo di aumentare la partecipazione e il flusso di dati da identificare.

Il Responsabile Open Data ha quindi il compito di valutare l'input (suggerimento di un nuovo dataset) e cercare in modo proattivo nuove sorgenti di dati in progetti esistenti, nuovi progetti o all'interno dei Piani Strategici in forma di archivi strutturati, elenchi o base dati prodotte da software in uso nell'Ente). La prima sorgente di dati è il sito web dell'Ente, da dove possono essere identificati molti dataset dai file PDF, fogli di calcolo, tabelle all'interno di documenti di testo o pagine HTML, modulistica, mappe, etc. Inoltre, molti "servizi online" possono contenere dataset di valore al loro interno.

1.2 Segnalazione esterna

I vari attori sociali (cittadini, associazioni, etc) sono un importante elemento all'interno di un modello di gestione dei dati aperti, in quanto veri destinatari dei dati pubblicati. Spesso hanno la necessità di accedere a dati su cui però sanno poco o niente in merito alla provenienza o gestione. Anche se per i cittadini sia possibile fare richiesta di accesso ad alcuni dati per i quali abbiano un interesse "diretto, concreto e attuale", avvalendosi dal diritto d'accesso introdotto dalla legge 241/90 o ai dati di pubblicazione obbligatoria tramite il "diritto di accesso civico" presente nel decreto legislativo n. 33/2013, entrambe modalità sono poco praticabili per gestire richieste di pubblicazione di nuovi dataset. Gli attori sociali possono ricercare e segnalare in modo attivo nuovi potenziali dataset, in modo strutturato (attraverso un apposito modulo "form" di contatto) o non strutturato (attraverso la casella di posta dedicata alle segnalazioni opendata@regione.umbria.it). L'Ente può anche sollecitare i vari attori sociali registrati all'interno del Portale Open Data (o registrati in una newsletter) in modo passivo, pubblicando o inviandogli periodicamente sondaggi

che hanno l'obiettivo di stabilire nuovi dataset di interesse collettivo.

Il Responsabile Open Data ha il compito di verificare e portare avanti le segnalazioni ricevute, consultando opportunamente il Responsabile della Struttura alla quale il dataset suggerito appartiene, il Referente Tecnico del Dataset se questo è già immediatamente identificabile e/o lo Staff Tutela Dati in modo preliminare rispetto a dubbi su dati eventualmente coperti da segreto di stato o altri potenziali impedimenti nell'ambito della privacy o diritto d'autore, anche se nella fase di Analisi e Pubblicazione ci siano processi specifici per controlli approfonditi. È una buona pratica gestire in modo continuativo ed interattivo le richieste di segnalazione esterna informando il richiedente dello stato di avanzamento della stessa o eventualmente con informazioni su come accedere ai dati se sono già stati pubblicati o sul perché non possono essere pubblicati.

1.3 Identificazione del Referente Tecnico del Dataset

Il Responsabile Open Data, a fronte di una segnalazione di identificazione di un nuovo dataset, deve identificare primariamente la risorsa che ha la maggiore conoscenza dei dati in questione. Tipicamente è la stessa risorsa che lavora quotidianamente nell'aggiornamento dei dati, o che gestisce la sorgente di provenienza del dato (progetto/servizio/sistema informativo).

Per compiere questa attività il Responsabile Open Data, se non riesce a riconoscere immediatamente la risorsa di competenza, deve avvalersi del supporto del Responsabile della Struttura di competenza di gestione del dato che indicherà i mezzi per arrivare al referente. Una volta identificata la risorsa, deve consultarla per confermare l'effettiva gestione dei dati ed ottenere dal responsabile della Struttura la nomina definitiva del Referente Tecnico del Dataset che dovrà essere registrata all'interno della Dataset Checklist.

Lo Staff Tutela Dati e lo Staff tecnico/operativo devono essere informati sull'esito del processo.

1.4 Attribuzione di un Referente Tecnico del Dataset

In assenza dell'identificazione di un referente per il Dataset segnalato, il Responsabile Open Data deve ricorrere al Responsabile della Struttura di appartenenza del dataset ed ottenere la nomina di un Referente Tecnico (che dovrà essere registrata nella Dataset Checklist), che abbia la conoscenza della natura dei dati e che possa fornire le informazioni necessarie al successivo censimento, estrazione ed aggiornamento continuo dei dati. Il Responsabile della Struttura è l'ultimo incaricato a trovare una risorsa che funga da Referente Tecnico del Dataset e, in caso tale identificazione non sia possibile, assume automaticamente il ruolo di Referente Tecnico del Dataset (preferibile perché appartiene alla stessa Area di competenza del dato), o in alcuni casi lo può delegare al Responsabile Open Data.

1.5 Compilazione della Scheda Informativa

Il Referente Tecnico del Dataset è responsabile per la compilazione della Scheda Informativa

(allegato A.1) preliminare, che censisce il dataset come parte del Patrimonio Informativo dell'Ente. La compilazione può avvenire in autonomia da parte del Referente Tecnico del Dataset, oppure nell'ambito di una sessione di intervista, guidata in questo caso dal Responsabile Open Data, incaricato dell'eventuale correzione dei dati incompleti o mancanti e dell'approvazione del deliverable prodotto in questo processo.

Lo Staff Tutela Dati può essere consultato per la compilazione della sezione sulla Privacy, anche se nella fase di Analisi e Pubblicazione ci saranno processi specifici per controlli approfonditi.

Il Responsabile della Struttura e lo Staff tecnico/operativo devono comunque essere notificati alla creazione o aggiornamento di ogni nuova Scheda Informativa.

1.6 Verifica finale della Scheda Informativa

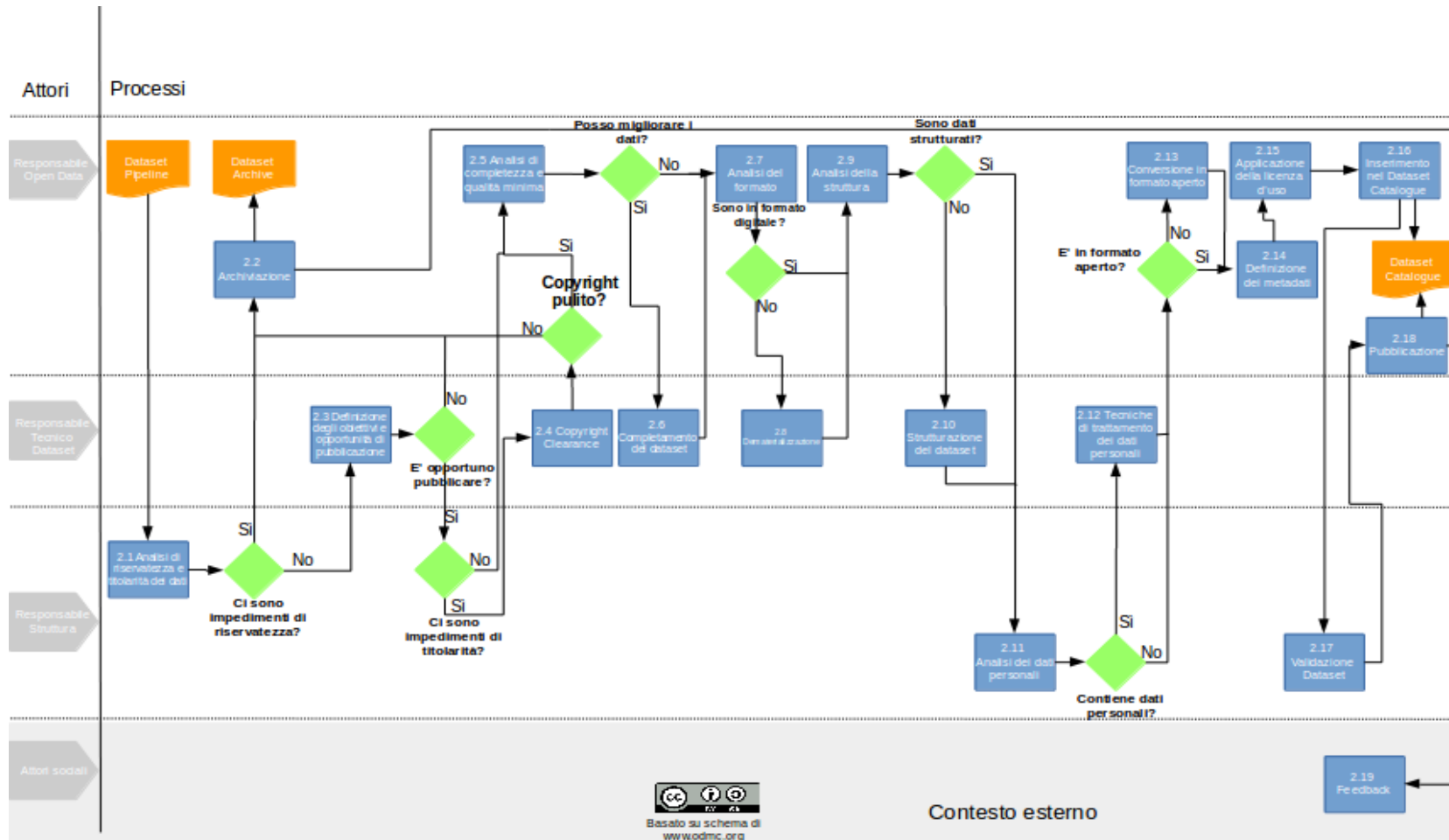
Una volta che la Scheda Informativa del dataset viene compilata (o aggiornata), il Responsabile Open Data deve assicurarsi che i dati siano validi, ovvero che non ci siano informazioni parziali, mancanti o di scarsa affidabilità. Deve eventualmente prendere nota di imprecisioni, integrazioni o correzioni necessarie, consultando lo Staff Tutela Dati e lo Staff tecnico/operativo per un peer-review. Il Responsabile Open Data è quindi incaricato di approvare la correttezza dei dati inseriti o aggiornati. La Dataset checklist dovrà contenere il nominativo della risorsa che ha validato la Scheda Informativa e la data di validazione.

1.7 Inserimento dei dati nella pipeline

La Scheda Informativa, compilata e validata in tutte le sue sezioni diventa un nuovo asset all'interno del Patrimonio Informativo gestito dall'Ente. Le informazioni contenute saranno utilizzate per la definizione dei metadati, ovvero informazioni che descrivono le caratteristiche principali dei dati all'interno del dataset (titolare del dato, frequenza di aggiornamento, copertura temporale, copertura geografica, etc) all'interno del Dataset pipeline. Il Responsabile Open Data, o in alternativa un membro dello Staff tecnico/operativo, deve essere il responsabile per l'inserimento delle informazioni della scheda all'interno del Dataset pipeline.

2. Analisi e Pubblicazione

In questo paragrafo vengono identificati i processi di lavorazione che portano ciascun dataset alla sua pubblicazione nella piattaforma dati.umbria.it, secondo il seguente schema:



2.1 Analisi di riservatezza e titolarità dei dati

Non tutti i dati prodotti all'interno dell'Ente potranno essere pubblicati liberamente. Ci sono infatti dati segreti (ad es.: soggetti a segreto di Stato, segreto statistico, etc) o con diritti di proprietà intellettuale di terzi che possono impedirne l'uso, il riutilizzo e la diffusione senza un'autorizzazione esplicita. Il Responsabile Struttura dovrà operare la classificazione del dataset compilando il campo "livello di riservatezza" nella checklist del dataset con uno dei seguenti valori:

- **Pubblico:** sono da classificare come di livello "pubblico" tutti i dataset che contengono dati afferenti ai documenti soggetti a pubblicazione obbligatoria per legge ai sensi del D.Lgs 33/2013, nonché tutti i dati già pubblicati online sul sito web dell'Ente. Sono inoltre da considerare pubblici tutti quei dati che, pur non essendo soggetti ad obbligo di pubblicazione, non ricadono nelle fattispecie indicate nei punti successivi;
- **Riservato:** sono da classificare come "riservati" tutti i dati coperti da segreto di Stato ai sensi della Legge n.801 del 24 ottobre 1977 e successive modificazioni, nonché le altre tipologie di dati escluse dal diritto di accesso e descritte all'articolo 16 della legge n. 15 del 11 febbraio 2005;
- **Vincolato:** sono "vincolati" tutti i dati in possesso dell'Ente sui cui terzi detengono diritti di proprietà intellettuale ai sensi della Legge n. 633 del 22 aprile 1941 e/o diritti di proprietà industriale ai sensi del D.Lgs n. 30 del 10 febbraio 2005, o che comunque sono stati originati da terze parti. Tali dati potranno essere aperti soltanto dopo la clearance dei diritti di copyright descritta nel processo *2.4 Verifica Licenza*.

Oltre alla classificazione del livello di riservatezza del dataset, il dirigente deve indicare nella checklist il nominativo e dati di contatto del 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; come definito dall'articolo 2, comma 1, lettera i, del D.Lgs n. 36 del 24 gennaio 2006 e successiva modificazione;
- terzi che detengano diritti di proprietà intellettuale ai sensi della Legge n. 633 del 22 aprile 1941 e/o diritti di proprietà industriale ai sensi del D.Lgs n. 30 del 10 febbraio 2005.

Lo spostamento dei dati da un sistema informativo ad un altro non modifica la titolarità e quindi rimane immutata la responsabilità sulla gestione dello stesso e sulla sua esattezza e veridicità, come previsto dall'art. 58 comma 1 del d.lgs. 82/2005. Ciò comporta che l'uso legittimo del dato in tutti gli altri casi (ossia dati formati da altri soggetti) avvenga ottenendo dal titolare apposita licenza che consenta la pubblicazione attraverso il processo *2.4 Verifica Licenza*.

Il Responsabile Struttura deve, nel rispetto della legge 633/1941, essere attento alla titolarità delle informazioni che si vogliono pubblicare al fine di evitare la violazione dei diritti di autore o altri

diritti di privativa di soggetti terzi. L'esistenza del diritto di autore e dei diritti connessi potrebbero impedire o limitare l'utilizzo dei dati per scopi diversi da quelli istituzionali.

I limiti alla conoscibilità dei dati sono quelli previsti dalle leggi e regolamenti in vigore con particolare riferimento alla tutela della riservatezza dei soggetti a cui i dati si riferiscono (d.lgs. 196/2003 "Codice in materia di protezione dei dati personali" e provvedimenti del Garante n. 17/2007 "Linee guida in materia di trattamento di dati personali per finalità di pubblicazione e diffusione di atti e documenti di enti locali" e n. 88/2011 "Linee guida in materia di trattamento dei dati personali contenuti anche in atti e documenti amministrativi, effettuati da soggetti pubblici per finalità di pubblicazione e diffusione sul web"). I dati anonimi sono comunque pubblicabili.

Il Responsabile Struttura, qualora lo ritenga necessario può avvalersi dello Staff Tutela Dati. Il responsabile tecnico del Dataset può essere consultato per il reperimento di queste informazioni. Il Responsabile Open Data e lo Staff tecnico/operativo devono essere informati una volta l'analisi sia conclusa.

2.2 Archiviazione

Nei casi in cui un dataset venga considerato non idoneo alla pubblicazione, il Responsabile Open Data è incaricato di compilare il documento Dataset checklist indicando la motivazione della richiesta di archiviazione:

- **Per riservatezza:** come indicato nel processo *2.1 Analisi di riservatezza e titolarità*, i dati classificati come "riservati" non sono pubblicabili come Open Data;
- **Per non titolarità:** il processo *2.4 Verifica Licenza* identifica, all'interno dei dati classificati come "vincolati" nel processo *2.1 Analisi di riservatezza e titolarità*, quali dati non possono essere aperti per impedimento contrattuale o mancati accordi per la clearance dei diritti d'autore;
- **Altro:** il processo *2.3 Definizione degli obiettivi e opportunità di pubblicazione*, attraverso la valutazione degli obiettivi strategici definiti dall'Ente e dal risultato dell'analisi del divario (gap analysis) potrebbe ritenere non opportuno continuare il percorso di apertura di un dataset (ad. es.: i costi di dematerializzazione, completamento e strutturazione di un dataset attualmente in formato cartaceo, l'utilità pubblica). In questi casi, sebbene la decisione di archiviazione sia più soggettiva che nei casi precedenti, l'analisi del divario e dei costi restano dati oggettivi che devono essere presenti e documentati nelle motivazioni per l'archiviazione del dataset.

Tutti gli altri attori devono essere informati delle motivazioni applicate al processo di archiviazione di un dataset. In particolare, se il dataset è stato analizzato in risposta a una richiesta da parte di uno degli attori sociali, è necessario fornire feedback diretto al richiedente in merito alla conclusione dell'analisi dei dati conforme le indicazioni del processo *2.17 Feedback*.

2.3 Definizione degli obiettivi e opportunità di pubblicazione

L'esecuzione di tutte le attività propedeutiche all'apertura dei dati previste dal modello dipendono da una valutazione complessiva dell'opportunità di apertura dei dati composta da:

- Analisi del divario (gap analysis) esistente tra la situazione attuale del dato, censita con la Scheda Informativa nella fase di Identificazione e lo stato finale di pubblicazione. Tale analisi ha come risultato la lista di interventi minimi necessari all'apertura dei dati (ad es.: dematerializzazione, Verifica Licenza, trattamento dei dati personali, completamento del dataset in merito al miglioramento della qualità, etc);
- Analisi del costo di apertura, considerando eventuali costi di gestione per eventuali progetti ad-hoc di adeguamento allo standard di pubblicazione, che in molti casi potrebbero gravare sull'area di appartenenza del dataset e non sulla struttura di gestione Open Data;
- Definizione degli obiettivi strategici, in relazione con il Piano Strategico dell'Ente in termini di priorità di apertura del dato, tempi di realizzazione, metriche che saranno usate nella fase di Monitoraggio per la valutazione della performance del dataset e risultati attesi che saranno valutati nella fase di Mantenimento. È necessario verificare se il dataset sia funzionale ai reali interessi della comunità di riferimento e se si trattano di dati necessari alla gestione delle attività (dati che servono all'Ente per il corretto funzionamento dei suoi processi), dati ottenuti dal risultato delle attività (dati che l'Ente produce nell'esercizio delle sue funzioni) o dati prodotti nella gestione delle attività-sottoprodotti di un processo (dati che l'Ente ottiene come sottoprodotto dei processi). Inoltre, deve essere definito il grado di apertura del dato secondo la classificazione a 5 stelle dei dati, proposta dall'ideatore del Web, Sir Tim Berners-Lee:
 - (1) ★ - dati disponibili sul web (in qualunque formato ad es. PDF, immagini) sotto una licenza aperta;
 - (2) ★★ - dati disponibili in un formato strutturato (ad es. MS Excel al posto di una scansione di una tabella);
 - (3) ★★★ - come (2) e utilizza formati non proprietari (ad es. CSV invece di MS Excel);
 - (4) ★★★★ - come (3) e utilizza standard aperti W3C (RDF, SPARQL) per identificare le risorse, in modo che si possano creare collegamenti ai dati;
 - (5) ★★★★★ - come (4) e collega i dati ad altri dati, al fine di creare un contesto.

Il Responsabile Open Data è incaricato di guidare la valutazione, documentando i risultati nella checklist del dataset e interagendo con il Responsabile della Struttura per la definizione degli obiettivi strategici e con gli altri attori per l'analisi del divario e le stime del costo di apertura. Infine, sulla base di quanto documentato, il Responsabile della Struttura di appartenenza del dataset potrà decidere se sarà opportuno o meno procedere con le attività di apertura dei dati.

2.4 Verifica Licenza

E' l'attività necessaria a contattare l'effettivo titolare del dato e predisporre un accordo di utilizzo o l'applicazione di una licenza aperta (v. processo 2.15 *Applicazione della licenza d'uso*) al dataset identificato, che permetta il suo riuso e la sua diffusione da parte dell'Ente e da chiunque abbia accesso ai dati.

Il Responsabile Struttura che tratta dati di cui non è titolare si attiverà per ottenere la "clearance" dei diritti d'autore per i dataset che abbiano superato la valutazione di opportunità nel processo 2.3 *Definizione degli obiettivi e opportunità di pubblicazione* e per i quali l'Ente non detiene la titolarità del dato. Il Responsabile Open Data potrà essere consultato in merito a requisiti e impedimenti specifici all'uso del dato in formato aperto.

Idealmente, si dovrebbe provare ad ottenere che il detentore dei diritti sui dati rinunci a tutti gli interessi sulla opera attribuendola al Pubblico Dominio con l'utilizzo dello strumento CC0, oppure CC-BY (come previsto dalla DGR 914 del 29/07/2013).

L'esito finale del processo (positivo o negativo) deve essere registrato nella checklist del dataset, con eventuali note sui termini dell'accordo di clearance dei diritti (ad es.: eventuale pagamento di royalties).

2.5 Analisi di completezza e qualità minima

Il Responsabile Open Data, assieme al Referente Tecnico del Dataset, sono responsabili per l'analisi *quantitativa* (completezza) e *qualitativa* (qualità) di dettaglio del dataset.

L'analisi quantitativa ha come obiettivo stabilire il grado di completezza del dataset in relazione alla realtà che rappresentano (ad es. se mancano elementi per rappresentare il 100% degli elementi attesi nel dataset completo). L'analisi qualitativa ha come obiettivo verificare che il dataset rispetti alcuni fattori minimi di conformità prima della pubblicazione (ad es. l'accuratezza sintattica e semantica dei dati, la consistenza interna ed esterna, etc).

Lo Staff tecnico/operativo può essere coinvolto per analisi puntuali dei singoli dati, con l'obiettivo di testare la loro consistenza, identificare e documentare i requisiti degli interventi per il miglioramento del dataset, e stimare con precisione tempi, costi e risultati attesi. Oltre ai problemi di qualità del singolo dato all'interno del dataset, tra i vari interventi di completamento che possono essere richiesti, è bene fare attenzione anche ai seguenti:

- Assenza di dati in determinati periodi di tempo (ad es. mancano i dati di un giorno o una settimana) e se è possibile ricrearli;
- La presenza di dati dove sia richiesta una password di accesso (ad es: fogli di calcolo protetti da password);
- Dati persi (ad es. era presente su postazione desktop che è stata formattata) e la possibilità di recuperare una copia di backup.

Il risultato dell'analisi deve essere documentato nel dataset checklist, assieme alla lista di eventuali interventi migliorativi che devono essere eseguiti nel processo 2.6 *Completamento del dataset* e messi a conoscenza del Responsabile della Struttura.

2.6 Completamento del dataset

Il Responsabile della Struttura deve revisionare ed approvare la lista di interventi di miglioramento della completezza e qualità del dataset prodotta come output del processo 2.5 *Analisi di completezza e qualità minima*. Il Responsabile Open Data, assieme allo Staff tecnico/operativo sono responsabili per la gestione del progetto di completamento del dataset che, tipicamente, consiste in attività di data cleaning o “pulizia dei dati”, ovvero il processo di individuazione, correzione o rimozione di record corrotti o inesatti. Dopo la pulizia, l'insieme di dati sarà coerente con altri dati simili nel sistema. Le incongruenze rilevate o rimosse possono essere state originariamente causate da errori di immissione degli utenti, dalla corruzione nella trasmissione o nella memorizzazione, o da differenti modi nella definizione dello stesso elemento (ad es. sinonimi).

L'esecuzione del progetto di completamento del dataset viene gestita di solito esternamente alla struttura di gestione Open Data in quanto realizzata idealmente nella sorgente stessa dei dati con l'obiettivo di rendere permanenti i miglioramenti di qualità dei dati stessi. Inoltre, la gestione del budget e la modalità di affidamento dell'attività (esecuzione interna o esterna all'Ente) restano sotto responsabilità dell'area di appartenenza del dataset.

2.7 Analisi del formato

Il Responsabile Open Data, assieme al Referente Tecnico del Dataset, sono responsabili per l'analisi del formato (o dei formati) del dato nella sua origine. Per origine (o sorgente) del dataset si intende dove il dato viene creato e gestito nello suo stato “originale” (ad es.: un foglio MS Excel, un sistema di ERP, un applicativo web, un archivio cartaceo...). Analizzare e documentare nella checklist del dataset il formato (o eventuali formati alternativi di accesso al dato originale) permette di classificare e comparare il grado di apertura attuale del dataset rispetto al grado finale di apertura desiderato e definito nel processo 2.3 *Definizione degli obiettivi e opportunità di pubblicazione* (da 1 a 5 stelle). A seconda del dato analizzato, gli attori responsabili del processo devono classificare la situazione del dato (se in formato digitale) all'origine tra le seguenti categorie e livelli:

- **Dati grezzi:** disponibili in formati che non consentono l'estrapolazione immediata degli stessi come ad esempio immagini nei diversi formati grafici (.gif, .jpg, .png, .bmp, etc) o documenti in formato proprietario Adobe PDF o Microsoft Word. Questi dati, senza trasformazioni, possono essere aperti con livello ★.
- **Dati strutturati:** dati disponibili in formati – aperti o proprietari - che ne consentono l'elaborazione in forma strutturata (es. righe e colonne). Ad esempio, fogli di calcolo in formato proprietario create con Microsoft Excel (in formato .xls, .xlsx) possono essere

aperti con livello ★★. Se salvati in formati non proprietari (come .csv, .ods) raggiungono il livello ★★★.

- **Database:** di norma di questi dati è possibile scegliere il formato di esportazione più idoneo (es. XML, JSON o altri formati aperti). L'esportazione può produrre file classificabili con livello ★★★★★ o ★★★★★★ (Linked Open Data) se relazionati ad altri dataset.

Come descritto nel processo 2.3 *Definizione degli obiettivi e opportunità di pubblicazione*, nel caso in cui l'Ente sia in possesso di dati grezzi (cioè non aperti e classificabili come ★), la trasformazione di questi in dati aperti dovrà essere valutata dal Responsabile Open Data, data la sua onerosità, sia in relazione al valore dei dati, sia in termini di risorse che di tempi.

Nel caso di dati a ★★ (strutturati, ma in formato proprietario) sarà opportuno convertirli in formati aperti e raggiungere quindi un livello ★★★.

I database che potranno generare dati a ★★★★★ o ★★★★★★ avranno bisogno di competenze tecniche specifiche dello Staff tecnico/operativo (o esterno), in quanto richiedono la costruzione di dataset basati su standard aperti definiti dalla W3C (come RDF, SPARQL) per identificare le risorse, in modo che si possano creare collegamenti ai dati.

Infine, non è raro trovare molti dei potenziali dataset in formato cartaceo. In tale caso, è necessario documentare accuratamente lo stato di conservazione, le caratteristiche, la quantità e il formato primario dei dati (tabelle, testo, immagini), informazioni necessarie ad avviare un progetto di dematerializzazione. Il risultato dell'analisi deve essere anche documentato nel dataset checklist.

2.8 Dematerializzazione

La dematerializzazione dei documenti è il processo mediante il quale gli atti transazionali (modulistica elettronica, incassi, pagamenti, assunzione o assolvimento di obbligazioni, etc) tra due o più soggetti si realizzano con il supporto informatico per l'acquisizione degli elementi costitutivi, l'elaborazione, l'archiviazione, il trasporto e la conservazione, con pieno valore tra le parti e verso i terzi. Il processo di dematerializzazione di un dataset non comprende quindi soltanto l'attività di scansione dei documenti cartacei e la conversione in file digitali, bensì la reingegnerizzazione della procedura che origina il flusso di dati. Tale progetto va quindi gestito in modo esterno alla struttura di gestione degli Open Data, perché impatta in modo diretto la modalità e gli strumenti di lavoro dell'area di appartenenza. Il Responsabile della Struttura è il primo responsabile per l'avvio e definizione dei requisiti generali della nuova procedura, che deve sostituire la precedente procedura cartacea. Il Referente Tecnico del Dataset è incaricato della definizione dei requisiti specifici della dematerializzazione, e di come il nuovo flusso di dati sia meglio integrato nel ciclo di gestione degli Open Data attraverso costante interfacciamento con il Responsabile Open Data e lo Staff tecnico/operativo.

2.9 Analisi della struttura

Il Responsabile Open Data, assieme al Referente Tecnico del Dataset, sono responsabili per l'analisi della struttura del dataset, la verifica di adeguamento al livello di apertura dei dati definito nel processo *2.3 Definizione degli obiettivi e opportunità di pubblicazione* e la documentazione di eventuali interventi di re-strutturazione dei dati tramite tecniche di ETL (Extract, Transformation, Load) per selezionare i dati di interesse (ad es.: estrarre tabelle dati da file PDF o MS Word), normalizzare i dati (ad es. eliminando duplicati), derivare nuovi dati calcolati (ad es.: aggiungere una nuova colonna con la somma dei valori estratti oppure ricavare le coordinate geografiche a partire da un indirizzo e numero civico). Le attività di ETL di questo processo non vanno confuse con le attività specifiche di trattamento dei dati personali eventualmente identificate nel processo *2.11 Analisi dei dati personali*, anche se potranno poi costituire un progetto unico per l'implementazione delle regole di trasformazione.

Lo Staff tecnico/operativo può essere consultato per fornire indicazioni e requisiti specifici per l'estrazione e trasformazione di dati in formati non strutturati (immagini, file PDF, MS Word) in dati strutturati e allineati agli obiettivi e requisiti di qualità già analizzati nei processi *2.3 Definizione degli obiettivi e opportunità di pubblicazione* e *2.5 Analisi di completezza e qualità minima*. Il Responsabile della Struttura va tenuto informato sull'esito dell'analisi.

2.10 Strutturazione del dataset

Un progetto per la ristrutturazione di un dataset deve considerare non solo gli obiettivi, i requisiti di qualità e gli interventi ETL necessari all'estrazione, trasformazione e caricamento dei dati già analizzati nei processi precedenti, ma soprattutto la creazione di una catena di valore per il riuso dei nuovi dati strutturati. Tale catena ha la sua maggiore espressione di valore rappresentata dal concetto di "data as a service", con l'utilizzo di URI specifiche per l'accesso alle risorse e, idealmente, un insieme di API (Application Programming Interface) per l'accesso granulare all'informazione che permetta agli attori sociali (cittadini, organizzazioni, etc) di sviluppare applicazioni o servizi integrati ai dati aperti (c.d. mashups).

Il progetto andrebbe seguito dal Responsabile Open Data e dal Referente Tecnico del Dataset che, con il supporto dello Staff tecnico/operativo possano assicurare la massima valorizzazione nel riuso del dato come materia prima per nuove attività economiche, maggior vantaggio competitivo per il territorio perché mette in evidenza il loro patrimonio informativo, oltre ai già noti principi di trasparenza, efficienza e collaborazione tra pubblico e privato. Al Responsabile della Struttura, viene attribuito il ruolo di stabilire il grado di strutturazione da applicare (dalla semplice trasformazione fino alla messa a disposizione di API specifiche per l'accesso ai dati) tenendo in mente i vantaggi del modello "data as a service" e valutando, con il supporto dello Staff tecnico/operativo, i costi complessivi di implementazione.

2.11 Analisi dei dati personali

La pubblicazione dei dati dell'Ente in formato aperto e di libero utilizzo non significa automatica condivisione di tutte le informazioni o accesso indiscriminato alle stesse; se tale eventualità è sicuramente applicabile ai dati "pubblici" cioè a quei dati che non possono essere riconducibili ad un soggetto fisico, lo stesso non può valere per la parte dell'informazione pubblica che ha carattere "personale" (ad esempio: registri anagrafici, registri sull'assistenza sociali, etc).

Per tali informazioni rimangono validi i principi e le disposizioni contenuti dalle leggi e dai regolamenti vigenti in tema di garanzia della riservatezza dei soggetti a cui i dati si riferiscono.

Con il termine "analisi dei dati personali" si fa quindi riferimento all'attività di valutazione delle informazioni soggette all'applicazione del d.lgs. 196/2003 "Codice in materia di protezione dei dati personali" e delle relative modalità di trattamento.

Il Responsabile Struttura è responsabile dell'analisi del dataset con l'obiettivo di stabilire se, tra le informazioni che lo compongono, sono comprese le seguenti tipologie di dato:

- **Dati personali**, ossia "qualunque informazione relativa a persona fisica, identificata o identificabile, anche indirettamente, mediante riferimento a qualsiasi altra informazione, ivi compreso un numero di identificazione personale" (art. 4, comma 1, lett. b del d.lgs.196/2003).
- **Dati sensibili**, ossia "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" (art. 4, comma e, lett. b del d.lgs.196/2003).
- **Dati giudiziari**, ossia "dati personali idonei a rivelare provvedimenti di cui all'articolo 3, comma 1, lettere da a) a o) e da r) a u), del d.P.R. 14 novembre 2002, n. 313, in materia di casellario giudiziale, di anagrafe delle sanzioni amministrative dipendenti da reato e dei relativi carichi pendenti, o la qualità di imputato o di indagato ai sensi degli articoli 60 e 61 del codice di procedura penale (art. 4, comma 1, lett. b del d.lgs.196/2003).

Nel caso in cui il dataset contenga le suddette tipologie di informazioni, Il Responsabile Struttura deve accertarsi, con il supporto delle funzioni competenti, che l'Ente:

- a. abbia già osservato i principi generali prescritti per il trattamento dei dati nella loro origine con particolare riguardo a quelli applicabili ai soggetti pubblici;
- b. abbia posto in essere gli adempimenti per la corretta applicazione della normativa.

Un quadro unitario delle regole che i soggetti pubblici sono tenuti ad applicare in relazione alle ipotesi di pubblicazione di dati personali sui propri siti istituzionali per finalità di trasparenza, pubblicità dell'azione amministrativa, nonché di consultazione di atti su iniziativa di singoli soggetti è contenuto nella Deliberazione 088/2011 del Garante Privacy "Linee guida in materia di

trattamento di dati personali contenuti anche in atti e documenti amministrativi, effettuato da soggetti pubblici per finalità di pubblicazione e diffusione sul web”. Considerando dataset contenenti dati personali occorre tener presente che in generale un ente pubblico:

- non può procedere alla diffusione dei dati idonei a rivelare lo stato di salute (art. 22, comma 8 del d.lgs. 196/2003);
- nel caso di trattamenti di dati diversi da quelli sensibili e giudiziari, (art. 19 del d.lgs. 196/2003):
 - la comunicazione ad altri soggetti pubblici è ammessa quando è prevista da una norma di legge o di regolamento. In mancanza di tale norma la comunicazione è ammessa quando è comunque necessaria per lo svolgimento di funzioni istituzionali e può essere iniziata se sia stata data comunicazione al Garante ai sensi dell’art. 39 del d.lgs. 196/2003 e sia decorso il termine di 45 giorni dal ricevimento della comunicazione e non è stata adottata una diversa determinazione del Garante.
 - la comunicazione a privati o a enti pubblici economici e la diffusione sono ammesse unicamente quando sono previste da una norma di legge o di regolamento.

Una volta completate le verifiche preliminari circa l’osservanza dei principi e la messa in opera degli adempimenti prescritti dalla normativa privacy, il Responsabile Struttura deve inoltre verificare che non sussistano le prescrizioni legislative dettate in tema di divieto di comunicazione e diffusione dei dati da parte di un soggetto pubblico negli art. 19 e 25 del d.lgs. 196/2003.

La pubblicazione del Dataset deve avvenire nel rispetto dei “Limiti alla trasparenza” fissati all’art. 4, commi 1, 2, 3 e 4 del d.lgs. 33/2013. In funzione dell’analisi sui dati personali basata sul d.lgs. 196/2003 si possono presentare i seguenti casi:

1° caso: Pubblicazione di dati personali, sensibili e giudiziari senza nessun tipo di operazioni preliminari sui dati.

Se ricorrono i presupposti del citato art. 18 del d.lgs. 196/2003 (svolgimento di funzioni istituzionali) e non sussiste uno specifico divieto. Il comma 1 dell’art. 4 del d.lgs. 33/2013 dispone la possibilità di una diffusione di dati personali attraverso siti istituzionali anche con modalità che ne consentono la indicizzazione e la rintracciabilità tramite i motori di ricerca web ed il loro riutilizzo ai sensi dell’articolo 7 del d.lgs. 33/2013.

Se è ridotta al minimo l'utilizzazione di dati personali (v. art. 3 del d.lgs. 196/2003) è possibile la pubblicazione nei siti istituzionali di dati relativi a titolari di organi di indirizzo politico e di uffici o incarichi di diretta collaborazione, nonché a dirigenti titolari degli organi amministrativi. Il comma 4, nei casi in cui norme di legge o di regolamento lo prevedano, dispone la pubblicazione (PUBBLICAZIONE OBBLIGATORIA) di atti o documenti che contengono dati personali pertinenti nonché dati sensibili e/o giudiziari indispensabili rispetto alle specifiche finalità di trasparenza della pubblicazione.

2° caso: Pubblicazione di dati informazioni e documenti con operazioni preliminari di

anonimizzazione dei dati personali eventualmente presenti

Il comma 3 dispone la possibilità di pubblicazione nei siti istituzionali di dati, informazioni e documenti per i quali non c'è l'obbligo di pubblicazione ai sensi del d.lgs. 33/2013 o sulla base di specifica previsione di legge o regolamento (PUBBLICAZIONE FACOLTATIVA), fermi restando i limiti e le condizioni espressamente previsti da disposizioni di legge, procedendo alla anonimizzazione dei dati personali eventualmente presenti.

3° caso: Pubblicazione di dati informazioni e documenti con operazioni preliminari di inintelligibilità dei dati personali, sensibili e giudiziari

Il comma 4, nei casi in cui norme di legge o di regolamento lo prevedano, dispone la pubblicazione (PUBBLICAZIONE OBBLIGATORIA) di atti o documenti purché le pubbliche amministrazioni provvedono a rendere non intelligibili i dati personali non pertinenti o, se sensibili o giudiziari, non indispensabili rispetto alle specifiche finalità di trasparenza della pubblicazione.

Resta fermo il limite alla diffusione dei dati idonei a rivelare lo stato di salute e la vita sessuale (art. 4, comma 6 del d.lgs. 33/2013).

A conclusione del processo, il Responsabile Struttura aggiorna la Dataset checklist con i risultati dell'analisi sui dati personali.

2.12 Tecniche di trattamento dei dati personali

Con il termine “tecniche di trattamento” si fa riferimento all'attività di individuazione degli strumenti tecnici per la protezione della privacy, necessari all'utilizzo dei dati personali da parte dell'Ente finalizzati alla pubblicazione in formato aperto e di libero utilizzo.

Tale attività, strettamente dipendente dagli output del processo *2.11 Analisi dei dati personali*, ha come obiettivo la tutela della riservatezza dei soggetti a cui i dati fanno riferimento, eliminando o minimizzando i dati personali allo scopo di prevenire processi non necessari e illegittimi, senza che ciò pregiudichi le finalità di trasparenza e di comunicazione alla base dell'Open Data.

In funzione dei risultati dell'Analisi dei dati personali riportata sulla dataset checklist, il Responsabile Struttura fornirà le indicazioni a cui lo staff tecnico operativo dovrà attenersi.

La principale tecnica di trattamento che minimizza o elimina del tutto l'utilizzo di dati personali è l'anonimizzazione. Il “dato anonimo” è il dato che non può essere associato ad un interessato identificato o identificabile.

Nell'ambito della comunicazione e diffusione, l'identificazione dell'interessato è posta in relazione ai mezzi ragionevolmente utilizzabili per procedere all'identificazione (risorse economiche, di tempo, archivi, risorse hardware e software) nonché alla probabilità di identificare l'interessato rispetto agli effetti di una violazione dei diritti degli interessati (rischio di identificazione).

La normativa privacy fornisce esempi diversi di tecniche di anonimizzazione in relazione a specifiche categorie di dati ma che possono costituire un valido riferimento anche per

l'individuazione delle tecniche di anonimizzazione da adottarsi ad altre categorie di dati:

- a. La prima che potremmo definire “anonimizzazione con chiave di collegamento o mascheramento” è il trattamento prescritto nel caso in cui i dati personali detenuti per scopi statistici siano oggetto di comunicazione a soggetti privati che non facciano parte del Sistema statistico nazionale. Tali informazioni devono essere comunicati sotto forma di collezioni campionarie, dati individuali privi di riferimenti con l'interessato. Le informazioni devono cioè riferirsi a soggetti identificati da un codice. La chiave che crea il collegamento tra il codice e i soggetti stessi rimane in possesso del Titolare.
- b. La seconda che potremmo definire “offuscamento” o “oscuramento” è il trattamento prescritto nel caso in cui su richiesta dell'interessato, o disposta “d'ufficio” dall'Autorità giudiziaria, siano precluse l'indicazione delle generalità e degli altri dati identificativi dell'interessato in caso di riproduzione della sentenza o provvedimento in qualsiasi forma per finalità di informazione giuridica, su riviste giuridiche, supporti elettronici o mediante reti di comunicazione elettronica.
- c. La terza, che fornisce una tutela più ampia delle precedenti e che potremmo definire “anonimizzazione in senso stretto”, è il trattamento prescritto in caso di diffusione dei dati dei minori e delle parti nei procedimenti giudiziari in materia di rapporti di famiglia e di stato delle persone. La norma impone di omettere non solo le generalità e gli altri dati identificativi dei soggetti tutelati ma anche gli “altri dati anche relativi a terzi dai quali può desumersi anche indirettamente l'identità di tali soggetti”.

2.13 Conversione in formato aperto

Il Responsabile Open Data, con il supporto dello Staff tecnico/operativo e attraverso consultazioni con il Referente Tecnico del Dataset, deve identificare quali saranno i formati aperti sui quali il dataset verrà convertito. Esistono diversi tipi di formati e la scelta del formato (o dei formati) da applicare al dataset deve dare priorità alla portabilità dei dati basata su standard aperti. Alcuni dei formati aperti standard di rappresentazione più comuni sono JSON, XML, CSV e RDF (KML e GML per dati geografici). L'obiettivo deve essere quello di utilizzare schemi standard che permettano agli sviluppatori un più veloce accesso alle informazioni e la maggiore facilità di manipolazione, proprio perché tali formati sono leggibili da computer e interpretabili da librerie software di sviluppo molto diffuse sul web.

Per i dati soggetti a frequenti modifiche, è utile pensare di esporre i dati attraverso web services. Ci sono diversi modi per creare un web service, ma alcuni dei più utilizzati sono SOAP e REST. I servizi web creati per accedere più facilmente ai dati faranno parte dell'API (Application Programming Interface) che sarà resa disponibile dall'Ente ed usata dagli sviluppatori per accedere ai dati e creare applicazioni basate su di essi, utilizzando paradigmi e protocolli come REST, HTTP e SQL. È importante, quindi, che il Responsabile Open Data lavori assieme allo Staff tecnico/operativo (o consulenti specializzati) in modo tale da garantire la creazione di un'architettura robusta ma estensibile che permetta alla comunità di sviluppatori di lavorare in

modo più efficace.

2.14 Definizione dei metadati

Il Responsabile Open Data, con il supporto dello Staff tecnico/operativo e attraverso consultazioni con il Referente Tecnico del Dataset, deve identificare quali saranno i metadati definitivi che saranno pubblicati con il dataset. Questa operazione di avvale anche dello schema di “mappatura” tra i campi della “Scheda Informativa” ed i metadati del “Dataset Catalogue”.

Infine, poiché i dataset possono essere soggetti a cambiamenti nel tempo (correzioni, aggiornamenti, elaborazioni, etc), è bene che lo Staff tecnico/operativo garantisca l’allineamento e la validità tra i metadati e i dati modificati. Alcuni riferimenti temporali da applicare ai metadati stessi debbono includere la data creazione dei metadati, la data dell’ultimo aggiornamento dei metadati, la lingua in cui i metadati sono scritti (ita, eng).

2.15 Applicazione della licenza d’uso

Una volta accertata l’assenza di vincoli di proprietà intellettuale che impediscano l’apertura dei dati (processo 2.1 *Analisi di riservatezza e titolarità*), è necessario che il Responsabile Open Data si occupi di identificare una licenza che ne definisca le condizioni e le modalità di riutilizzo da parte di chiunque, anche per scopi commerciali, conformemente all’articolo 8 del D.Lgs. 36/2006:

“1. Gli schemi di licenze standard per il riutilizzo sono predisposti dal titolare del dato, ove possibile in formato elettronico e resi disponibili sul proprio sito istituzionale. Gli schemi possono essere compilati elettronicamente e contengono eventuali limitazioni o condizioni all'utilizzo dei documenti, in considerazione delle loro peculiari caratteristiche, nonché l'indicazione dei mezzi di impugnazione, secondo criteri individuati dal titolare medesimo con proprio provvedimento.

2. Le condizioni e le limitazioni poste dal titolare del dato negli schemi di licenze standard sono individuate per categorie di documenti secondo criteri di proporzionalità e nel rispetto della disciplina sulla protezione dei dati personali e non possono costituire ostacolo alla concorrenza.”

dove per “licenze standard per il riutilizzo” lo stesso documento definisce “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”. Ci sono molte licenze che si prestano al riuso dei dati aperti ma, prima della scelta del tipo specifico di licenza da applicare, è bene che il Responsabile Open Data tenga in conto che tale scelta comporterà, in sostanza, la classificazione dei dati all’interno di uno di questi gruppi:

- **Dati di Pubblico Dominio (CC0):** sgrava l’Ente da ogni responsabilità su eventuali rielaborazioni sbagliate realizzate da terzi ed assomiglia i dati a tutte le altre tipologie di dati che i cittadini comunemente già percepiscono come dati di pubblico dominio (bilanci, dati statistici, calendari eventi, etc);
- **Dati con vincolo di attribuzione (CC-BY):** riconosce all’Ente lo sforzo impiegato nella

creazione (o “paternità”) dei dati, e richiede a chi lo riutilizza di citare l’Ente nell’opera derivata;

- **Dati con vincolo di condivisione allo stesso modo (CC-BY-SA):** richiede che il risultato del riuso venga condiviso con la stessa licenza, con impiego comune per tutti quei dati che sono stati raccolti in modalità “crowdsourced”, ovvero con l’aiuto dei cittadini in maniera volontaria (segnalazioni di problemi, idee per lo sviluppo, etc)

In ottemperanza alla DGR n.914 del 29/07/2013 la scelta della licenza di norma deve ricadere sulla “CC-BY”, che consente a chiunque di riprodurre, distribuire, comunicare al pubblico, esporre in pubblico, rappresentare, eseguire, modificare o usare l’opera, anche per **fini commerciali**.

Fatti salvi i limiti di legge inderogabili, le licenze da adottare vanno individuate dal Responsabile Open Data in modo tale che l’Ente licenziante non si assuma alcuna responsabilità per qualsiasi costo, onere, spesa o danno di qualsiasi natura che dovessero derivare dall’esercizio delle facoltà concesse in licenza, compreso l’utilizzo o il mancato utilizzo del documento da parte del licenziatario o da parte di terzi.

A ciascun dataset verrà abbinata la licenza con cui viene pubblicato (e registrata nel dataset checklist) e questa, idealmente, dovrà successivamente essere accessibile in lingua italiana. Il Responsabile Open Data deve prediligere una licenza leggibile e riconoscibile non solo agli utenti nazionali, ma anche agli utenti internazionali e che garantisca la massima possibilità di riutilizzo dei dati, a meno di specificità che richiedano l’adozione di licenze meno aperte (ad es.: uso dei dati a pagamento).

2.16 Inserimento nel Dataset catalogue

Il Responsabile Open data, con l’immediato supporto dello Staff tecnico/operativo deve organizzare l’attività di messa in produzione del dataset, nel rispetto di tutte le indicazioni previste nella dataset checklist. Può consultare il Responsabile della Struttura nel caso in cui si decida di eseguire tutti gli interventi dei processi esterni come parte di un unico progetto di apertura dei dati (ad esempio un dataset che preveda anche la dematerializzazione, la conversione del formato e modifiche nella struttura). In alternativa, il lavoro può essere eseguito direttamente dallo Staff tecnico/operativo se tutte le competenze informatiche siano presenti.

Lo Staff tecnico/operativo deve assicurarsi che i metadati (nel modo più completo possibile) siano presentati in maniera chiara all’interno di una pagina per il download del dataset con chiara visibilità dei formati disponibili ed eventuali modalità di accesso alternative (ad esempio tramite API). Oltre alla conversione del formato, il Responsabile Open Data deve valutare anche l’opportunità di offrire esperienze iterative per la navigazione dei dati del dataset direttamente dal “Portale Open Data”, come mappe iterative (basate per esempio su file KML e shape di un dataset), visualizzazioni grafiche di diverso tipo (istogrammi, grafici a barre, a linee, a dispersione, a torta, a bolle), visualizzazione tabulare dei dati, nonché accesso a API che permettano alla comunità di

sviluppatori di interrogare il dataset in tempo reale.

In ogni caso il Responsabile Open Data, prima della pubblicazione definitiva di un dataset deve assicurarsi che esso rispetti al massimo possibile le caratteristiche indicate in premessa.

Il Responsabile Open Data, con l'immediato supporto dello Staff tecnico/operativo, provvede quindi all'inserimento del dataset nel "Dataset Catalogue" della piattaforma dati.umbria.it, assegnandogli provvisoriamente l'attributo "*Privato*", quindi trasmette al Responsabile Struttura richiesta di validazione fornendo il link attraverso il quale possa visualizzare e valutare direttamente in piattaforma tutte le informazioni del dataset e le relative risorse collegate.

2.17 Validazione Dataset

Il Responsabile Struttura valuta opportunamente le informazioni del dataset e le relative risorse collegate. In caso di esito positivo, comunica tempestivamente al Responsabile Open Data la validazione del dataset, oppure, in caso di esito negativo, segnala in dettaglio al Responsabile Open Data i problemi riscontrati e suggerisce possibili soluzioni.

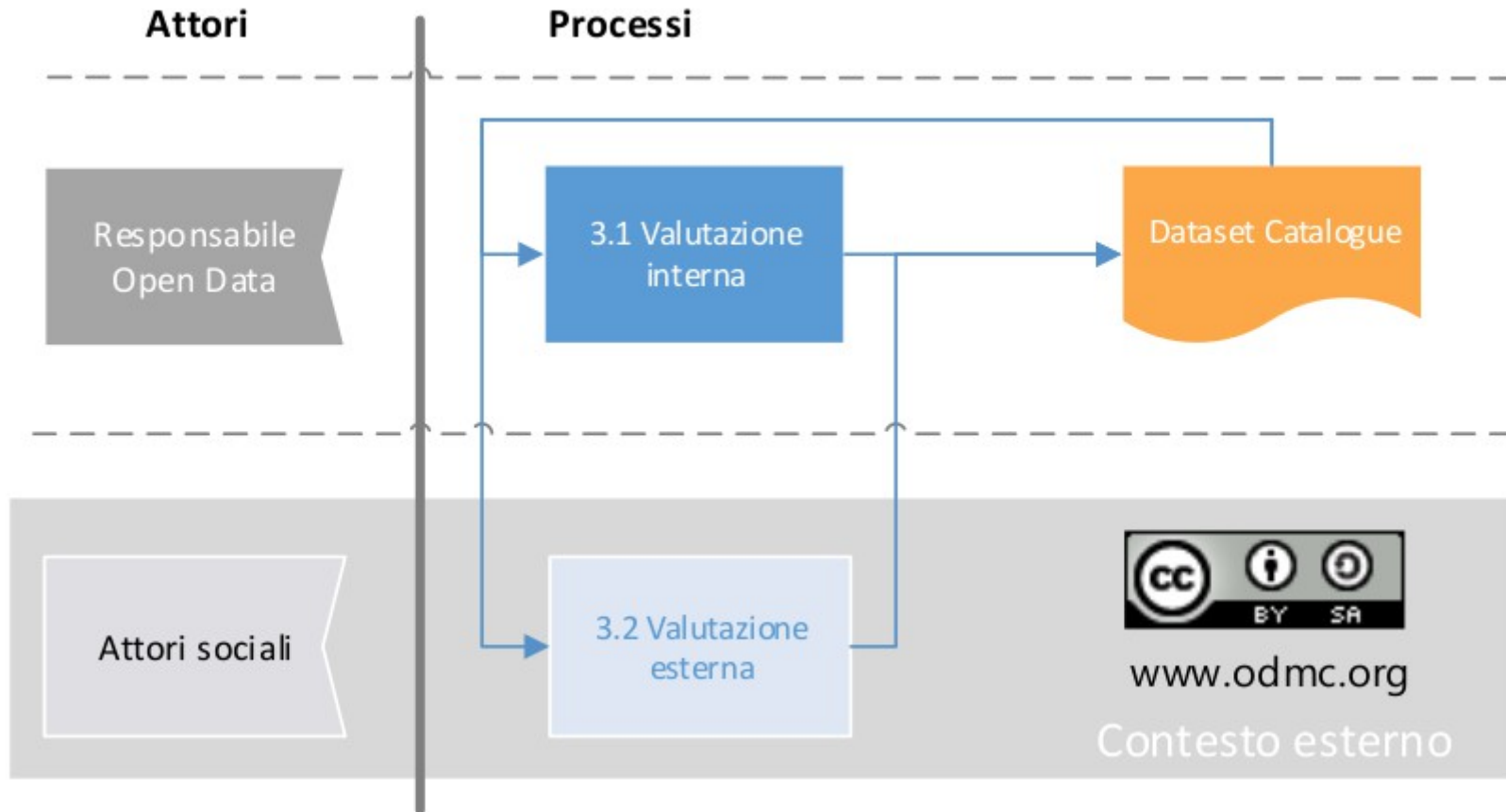
2.18 Pubblicazione

Il Responsabile Open Data, con l'immediato supporto dello Staff tecnico/operativo, ricevuta la validazione del dataset da parte del Responsabile Struttura, provvede nella "Dataset Catalogue" della piattaforma dati.umbria.it ad assegnare l'attributo "*Pubblico*" al dataset.

2.19 Feedback

Il Responsabile Open Data, con l'immediato supporto dello Staff tecnico/operativo deve dare tempestivo feedback agli utenti sulla pubblicazione del nuovo dataset che può avvenire sotto forma di pubblicazione di un annuncio di rilascio di un nuovo dataset in una sezione specifica del Portale Open Data (pagina "notizie" o "annunci"). In modo particolare, se la pubblicazione è avvenuta in seguito ad una o più richieste da parte degli utenti, è necessario rispondere in modo puntuale alla richiesta, informando dell'avvenuta pubblicazione. Mantenere una buona relazione di comunicazione con gli utenti che interagiscono con l'Ente tramite il Portale Open Data è uno dei punti chiavi nella costruzione di una comunità Open Data attiva e partecipativa. Infine, una nota interna di pubblicazione può essere emessa al Responsabile della Struttura ed al Referente Tecnico del Dataset, nonché ad altri soggetti potenzialmente interessati all'interno dell'Ente.

3. Monitoraggio



3.1 Valutazione interna

Una volta pubblicato sul Portale Open Data, il Responsabile Open data deve incaricare lo Staff tecnico/operative del monitoraggio della performance e del gradimento del dataset. Per la maggior parte delle metriche di misurazione la raccolta dei dati può essere automatizzata utilizzando le funzionalità del sistema di gestione del Portale Open Data o prodotti specifici di monitoraggio del traffico e del comportamento degli utenti del Portale (ad esempio con Google Analytics). Attraverso l'indirizzo web del dataset pubblicato sul Portale, e i log di accesso solitamente gestiti dai sistemi di web analytics appena citati, lo Staff tecnico/operativo potrà monitorare diversi parametri di performance del dataset quali, ma non solo:

- Disponibilità dell'accesso, ovvero se l'indirizzo è stato raggiunto in maniera corretta (oppure se sono presenti errori "404 Pagina non trovata");
- Il numero di visite (e visite uniche) alla pagina del dataset e la variazione rispetto a periodi precedenti;
- Conteggio dei download e la sua distribuzione nel tempo, in grado di fornire un utile indicatore dell'interesse per un determinato set di dati;
- Le sorgenti principali di traffico, per identificare i siti che portano la maggior parte dei visitatori al dataset pubblicato;
- Percentuale e numero assoluto di "conversioni" per obiettivi predefiniti (ad esempio numero di persone che hanno scaricato il dataset in un formato specifico - attraverso una URI permanente - rispetto il totale di accessi alla pagina del dataset).

Il Responsabile Open Data e lo Staff tecnico/operativo possono usare, inoltre, altre metriche non relazionate direttamente al traffico della pagina del dataset sul Portale Open Data per valutare il gradimento del dataset da parte degli utenti:

- Numero di applicazioni ("app") che usano il dataset: è difficile prevedere come verranno impiegati i dataset rilasciati, o quante applicazioni saranno sviluppate. Per questo motivo, si consiglia di lasciare a disposizione degli utenti un formulario di segnalazione di applicazioni basate sul dataset, ed eventualmente una sezione "vetrina" per le applicazioni create su ciascun dataset, di modo a incentivare gli sviluppatori a segnalare le proprie applicazioni;
- Numero di download di una app che si appoggia sul dataset: come per l'esistenza dell'app, il numero di downloads, se non segnalato in qualche modo dall'autore dell'applicazione può essere soltanto stimato in base alle valutazioni e dati presenti nella pagina dell'applicazione nei vari app store (Apple store, Google play store);
- Numero di "like" sull'app o sul dataset esposto: sistemi di valutazione (1-5 stelle, "mi piace/non mi piace") presenti sulla pagina del dataset e/o sulla pagina di download delle applicazioni che lo usano possono dare ulteriori segnali sul gradimento del dataset e del suo utilizzo da parte della comunità di riferimento;

- Segnalazioni ricevute: osservazioni, proposte e rilevazioni di eventuali inesattezze o errori sui dataset attraverso la pagina del dataset possono contenere ulteriori indicazioni sul gradimento/performance dei dati rilasciati;
- Risultato di sondaggi ad-hoc: è possibile pubblicare un sondaggio aperto sul Portale Open Data, oppure utilizzare dati di una “newsletter” per mandare un sondaggio ad-hoc a utenti registrati o a categorie potenzialmente interessate ai dati dell’Ente (ad es.: ordini professionali, confcommercio, albergatori, etc) e raccogliere informazioni più approfondite sul gradimento e utilizzo del dataset pubblicato.

I dati aperti creeranno valore solo quando qualcuno li userà, specialmente con profitto. È quindi importante tener conto dell’interesse commerciale che riscuotono i diversi dataset e valutare la qualità delle applicazioni che ne derivano. Al Responsabile Open Data tocca comprendere quali dati sono realmente in grado di creare valore pubblico o di generare profitto per le imprese, valutando in modo complessivo quanti più possibili dati statistici su di essi. Tale comprensione può essere utilmente mediata attraverso processi di co-progettazione che coinvolgano alcuni potenziali portatori di interesse, attraverso rapporti diretti e continuativi con gli sviluppatori, le imprese e le comunità open data, studenti, ricercatori e cittadini.

Per l’Ente questo deve significare la costruzione di un dialogo continuo con tutti gli attori sociali coinvolti, dove la capacità di ascolto e la volontà d’interazione sono due fattori indispensabili per il successo, ben più rilevanti degli aspetti tecnologici.

3.2 Valutazione esterna

Gli attori sociali (associazioni e comunità open data, sviluppatori, imprese, studenti, ricercatori, cittadini, altri enti pubblici) sono partecipanti attivi del processo di valutazione sulla performance e gradimento sul dataset pubblicato. Le loro osservazioni, proposte o rilevazioni di eventuali inesattezze o errori sui dataset sono fondamentali per la conclusione del ciclo di miglioramento continuo proposto. Se previste dal Portale Open Data, di seguito alcune funzionalità utili a permettere l’espressione di gradimento degli attori sociali sul dataset pubblicato:

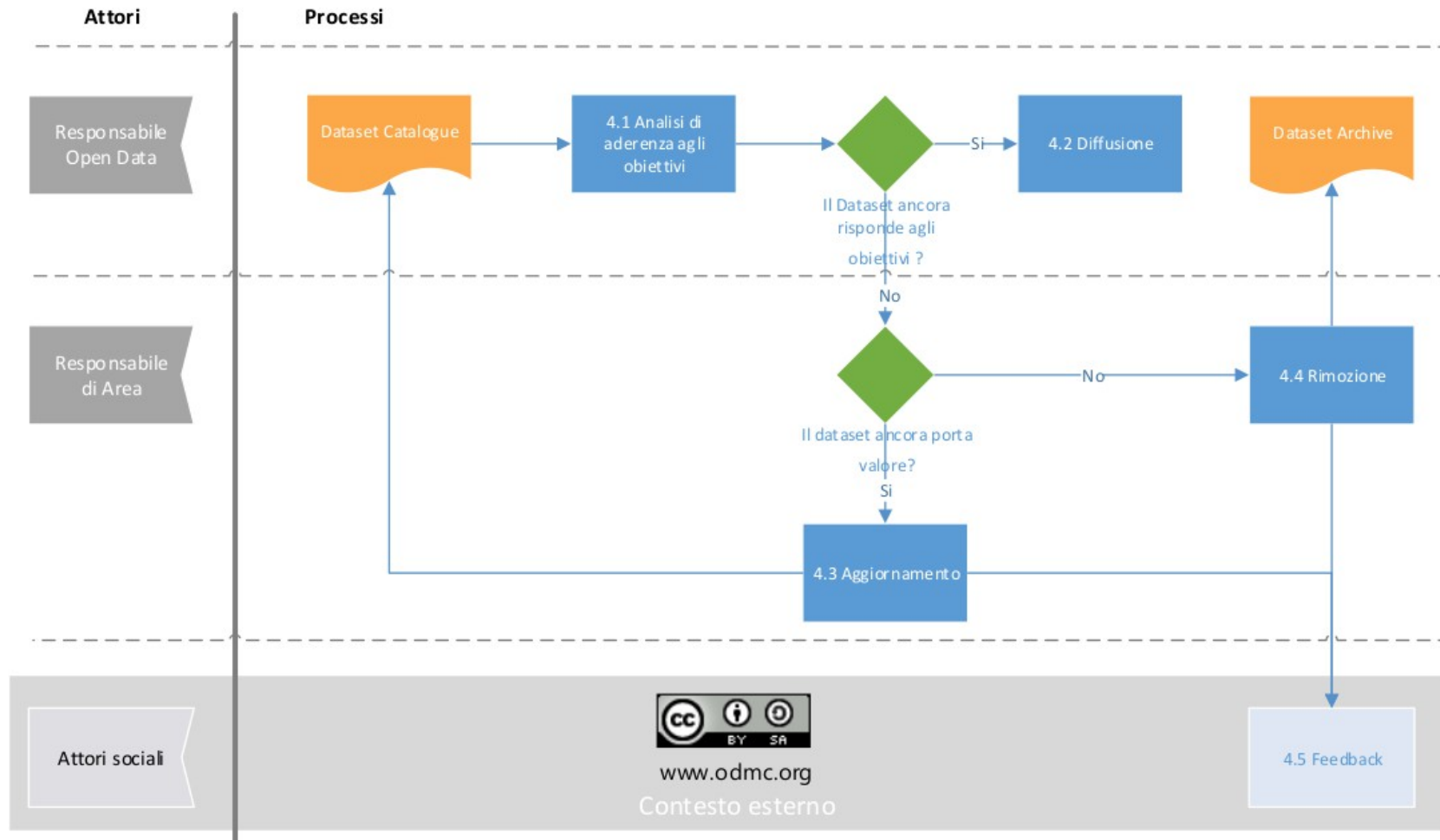
- Sistema di rating: l’abilitazione di un sistema di rating (ad es.: 1-5 stelle “mi piace/non mi piace”) sulla stessa pagina di download del dataset permette agli attori sociali di esprimere velocemente e in modo anonimo una preferenza generica sulla presentazione dei dati. Si può abilitare un sistema di valutazione più strutturato per valutare i dati in dimensioni più definite come “presentazione della pagina e interattività”, “qualità dei dati”, “accessibilità” tenendo come obiettivo però la semplicità d’uso di tale sistema (non presentare troppe dimensioni e fornire in modo chiaro una breve spiegazione sul significato di ciascuna dimensione da valutare);
- Formulario di contatto: un formulario di contatto dedicato (o una semplice casella di posta) permette agli attori sociali di fornire una valutazione più completa (per esempio classificate come “correzione”, “suggerimento”, “problemi di utilizzo”) in merito al

dataset. È importante fornire un campo a testo libero e, quando possibile, chiedere un recapito per eventualmente contattare l'utente e fornire una risposta quanto prima;

- Formulario di segnalazione app: permettere agli attori sociali (in modo particolare la comunità di sviluppo) di segnalare applicazioni realizzate sulla base del dataset pubblicato, ed eventualmente riproporre queste applicazioni in un'apposita sezione del Portale Open Data aiuterà a stabilire un legame più forte di collaborazione tra l'Ente e la comunità di riferimento.

È importante che il Responsabile Open Data abbia definito, all'interno della struttura di gestione dei dati aperti, quali saranno le modalità di dialogo con gli attori sociali. In modo particolare è importante assicurarsi che le segnalazioni provenienti dai formulari e che contengano il recapito dell'utente vengano sempre seguite e mai lasciate senza risposta, prima con una email di conferma del ricevimento della segnalazione e successivamente con uno o più messaggi che diano seguito al dialogo fino alla sua conclusione. Tale accorgimento aumenta in modo significativo la fiducia degli attori sociali coinvolti nell'impegno preso dall'Ente per la gestione dei dati aperti e l'ascolto verso i suoi veri destinatari.

4. Mantenimento



4.1 Analisi di aderenza agli obiettivi

Il processo 2.3 *Definizione degli obiettivi e opportunità di pubblicazione*, ha definito, all'interno del dataset checklist uno o più obiettivi strategici, allineati al Piano di Strategia Open Data definito dall'Ente, in termini di priorità di apertura del dato, risultati attesi e grado di apertura (1 a 5 stelle). Il Responsabile Open Data è incaricato di verificare l'aderenza del dataset agli obiettivi preposti, e compilare i risultati dell'analisi in una sezione di controllo del dataset check, assieme alla data del controllo e le risorse consultate (Responsabile della Struttura, Referente Tecnico del Dataset, Staff tecnico/operativo). L'analisi servirà a definire quali saranno i prossimi passi per la gestione del dataset, la sua diffusione, aggiornamento o eventuale dismissione.

4.2 Diffusione

Se il dataset analizzato nel processo 4.1 *Analisi di aderenza agli obiettivi* è ancora aderente agli obiettivi stabiliti, il Responsabile Open Data deve preoccuparsi di diffondere il dataset utilizzando il maggior numero di canali possibili. È essenziale tenere presente che gran parte dei dati non raggiungeranno gli utenti finali direttamente, ma tramite persone che prenderanno i dati e li trasformeranno per la presentazione. Ad esempio, la maggior parte dei destinatari di un dataset non vorrà (o non avrà bisogno) di navigare all'interno di una grande tabella con coordinate GPS, ma saranno decisamente più interessate a navigare i dati visualizzandoli all'interno di una mappa.

Lo Staff tecnico/operativo è il responsabile per attuare le azioni di comunicazione e diffusione del dataset tra i vari canali alternativi, con la supervisione del Responsabile Open Data:

- **Hackathon o incontri locali:** è probabile che ci siano già gruppi locali organizzati che lavorano per l'apertura dei dati pubblici e con i quali l'Ente può dialogare ed ottenere supporto per organizzare un incontro sul tema, migliorare i dati pubblicati o promuovere la creazione di nuove applicazioni basate sul dataset. Questi incontri sono destinati principalmente al coinvolgimento della comunità di sviluppatori, all'interno delle quali si trovano persone con forte motivazioni e competenze tecniche, pronte a mostrare le loro capacità e impiegare il loro talento nella costruzione di applicazioni e servizi per il bene pubblico. Incentivare queste attività di "civic hacking" è fondamentale per la crescita della popolarità del dataset;
- **Concorsi:** sulla scia dell'iniziativa americana "Apps for Democracy" e simile esperienza italiana "Apps4Italy", l'Ente può promuovere concorsi per incentivare la creazione di applicazioni/servizi innovativi basati sui dataset pubblicati;
- **Portali Aggregatori di Dati:** ci sono siti che raccolgono i dataset messi a disposizione dalle varie pubbliche amministrazioni (ad esempio: Dati.gov.it). Segnalare il proprio dataset aumenta la visibilità e la possibilità di riutilizzo dei dati;
- **Social Network:** se l'Ente gestisce canali sociali come Facebook, Twitter, LinkedIn, YouTube e/o altri social network, è possibile veicolare l'annuncio di apertura (o

aggiornamento) dei dati anche tramite essi;

- **Press releases:** non tutti i dataset meritano un comunicato stampa formale da parte dell'Ente. Ma il Responsabile Open Data potrà eventualmente produrne uno per quei dataset che portano un sostanziale valore alla comunità di riferimento, che sono di grande interesse pubblico, oppure che abbiano suscitato notevole interesse dalla comunità di sviluppo e per il quale ci siano diverse applicazioni funzionanti;
- **Blog:** Se nel Portale Open Data è presente un blog, un post che evidenzia gli obiettivi strategici e propone alla comunità di riferimento alcune idee sul potenziale di utilizzo dei dati può essere pubblicato dal Responsabile Open Data o dallo Staff tecnico/operativo. In alternativa, contattare cittadini che mantengono blog tematici e avvisarli dei nuovi rilasci può essere l'inizio per stabilire canali di comunicazione più vicini alle persone e meno formali;

Se il dataset analizzato nel processo *4.1 Analisi di aderenza agli obiettivi* non risponde pienamente agli obiettivi iniziali (primariamente perché il contenuto non è più attuale, ma anche rispetto al piano strategico dell'ente), ma porta ancora valore (è di interesse pubblico, ha una buona performance di accesso/utilizzo, ci sono applicazioni che lo usano), il Responsabile della Struttura di appartenenza del dataset dovrebbe avviare il processo di aggiornamento dei dati, consultando il Referente Tecnico del Dataset per informazioni su cambiamenti nella struttura e contenuti e lo Staff Tutela Dati per verificare che non ci siano problemi nell'aggiornamento di eventuali dati personali trattati prima della pubblicazione.

Il Responsabile Open data e lo Staff tecnico/operativo devono orientarsi verso una soluzione di aggiornamento dei dati che permetta il versionamento del dataset, permettendo agli utenti (ma soprattutto alle eventuali applicazioni che utilizzano il dataset tramite API) di continuare ad accedere ai dati della versione precedente, specialmente se il dataset cambia nella sua struttura. È importante registrare nella checklist del dataset tutti i cambiamenti apportati alla nuova versione del dataset, anche nei metadati (che devono essere anch'essi aggiornati).

Se il dataset analizzato nel processo *4.1 Analisi di aderenza agli obiettivi* non risponde più agli obiettivi iniziali e si considera che in qualche modo sia obsoleto (ad esempio perché è stato rilasciato un nuovo dataset che contiene gli stessi dati ed è più completo) e non porta più valore (o confonde gli utenti, come nel caso del dataset superato da un nuovo con struttura più completa e granulare), il Responsabile della Struttura di appartenenza del dataset dovrebbe avviare il processo di dismissione dei dati, all'interno del quale il Responsabile Open Data dovrà considerare tutte le implicazioni di tale azione (ad esempio che, come un normale servizio, la dismissione di un dataset potrebbe far smettere di funzionare applicazioni che usano le API di accesso ai dati). Un dataset va rimosso solo con giusta e motivata causa, registrata sul dataset checklist, e la dismissione di eventuali dataset con accesso tramite API va pianificata in anticipo, con rilascio di informazioni chiare su quali funzioni di accesso ai dati potrebbero smettere di funzionare (vedi processo *4.5 Feedback*).

Così come accade dopo la pubblicazione di un dataset (processo *2.17 Feedback*), il Responsabile

Open Data, con l'immediato supporto dello Staff tecnico/operativo deve dare tempestivo feedback agli utenti dopo l'aggiornamento di un dataset esistente (e/o cambiamenti nel modo di accedere ai dati tramite API) o dopo la dismissione del dataset. Tali annunci possono avvenire sotto forma di pubblicazione di una nota in una sezione specifica del Portale Open Data (pagina "notizie" o "annunci"). In modo particolare, se l'aggiornamento o rimozione è avvenuta in seguito ad una o più richieste da parte degli utenti, è necessario rispondere in modo puntuale alla richiesta, informando dell'avvenuta modifica. Mantenere una buona relazione di comunicazione con gli utenti che interagiscono con l'Ente tramite il Portale Open Data è uno dei punti chiavi nella costruzione di una comunità Open Data attiva e partecipativa. Infine, una nota interna sulle modifiche apportate può essere emessa al Responsabile della Struttura ed al Referente Tecnico del Dataset, nonché ad altri soggetti potenzialmente interessati all'interno dell'Ente.

Scheda Informativa dataset

- Partecipare attivamente all'interno delle **comunità Open Data**: le comunità Open data sul web comunicano spesso attraverso mailing lists. La principale in Italia è Spaghetti Open Data che raduna sviluppatori, funzionari pubblici, data journalists e semplici curiosi che si sono appassionati al tema dei dati aperti. Il Responsabile Open Data (ma possibilmente anche alcuni membri dello Staff tecnico/operativo) può interagire scrivendo personalmente alle comunità, aggiornandole gradualmente sul progresso di pubblicazione e mantenimento dei dataset, e aiutando i loro membri con le sue esperienze e l'esperienza dell'Ente nella gestione dei dati. La scelta di quante e quali comunità coinvolgere e partecipare può essere valutata in base a quale sia la relazione attesa con la comunità in futuro e che tipo di attività possono essere condotte insieme (concorsi, incontri, etc), alla presenza di leader e attivisti che guidano la comunità senza interessi politici, l'organizzazione di incontri offline che aumentano il senso di appartenenza. Può essere anche un modo pratico per mantenersi informati su cosa stanno facendo le altre pubbliche amministrazioni e quali dataset stanno aprendo.

4.3 Aggiornamento

Se il dataset analizzato nel processo *4.1 Analisi di aderenza agli obiettivi* non risponde pienamente agli obiettivi iniziali (primariamente perché il contenuto non è più attuale, ma anche rispetto al piano strategico dell'ente), ma porta ancora valore (è di interesse pubblico, ha una buona performance di accesso/utilizzo, ci sono applicazioni che lo usano), il Responsabile della Struttura di appartenenza del dataset dovrebbe avviare il processo di aggiornamento dei dati, consultando il Referente Tecnico del Dataset per informazioni su cambiamenti nella struttura e contenuti e lo Staff Tutela Dati per verificare che non ci siano problemi nell'aggiornamento di eventuali dati personali trattati prima della pubblicazione.

Il Responsabile Open data e lo Staff tecnico/operativo devono orientarsi verso una soluzione di aggiornamento dei dati che permetta il versionamento dei dataset, permettendo agli utenti (ma

soprattutto alle eventuali applicazioni che utilizzano il dataset tramite API) di continuare ad accedere ai dati della versione precedente, specialmente se il dataset cambia nella sua struttura. È importante registrare nella checklist del dataset tutti i cambiamenti apportati alla nuova versione del dataset, anche nei metadati (che devono essere anch'essi aggiornati).

4.4 Rimozione

Se il dataset analizzato nel processo *4.1 Analisi di aderenza agli obiettivi* non risponde più agli obiettivi iniziali e si considera che in qualche modo sia obsoleto (ad esempio perché è stato rilasciato un nuovo dataset che contiene gli stessi dati ed è più completo) e non porta più valore (o confonde gli utenti, come nel caso del dataset superato da un nuovo con struttura più completa e granulare), il Responsabile della Struttura di appartenenza del dataset dovrebbe avviare il processo di dismissione dei dati, all'interno del quale il Responsabile Open Data dovrà considerare tutte le implicazioni di tale azione (ad esempio che, come un normale servizio, la dismissione di un dataset potrebbe far smettere di funzionare applicazioni che usano le API di accesso ai dati). Un dataset va rimosso solo con giusta e motivata causa, registrata sul dataset checklist, e la dismissione di eventuali dataset con accesso tramite API va pianificata in anticipo, con rilascio di informazioni chiare su quali funzioni di accesso ai dati potrebbero smettere di funzionare (vedi processo *4.5 Feedback*).

4.5 Feedback

Così come accade dopo la pubblicazione di un dataset (processo *2.17 Feedback*), il Responsabile Open Data, con l'immediato supporto dello Staff tecnico/operativo deve dare tempestivo feedback agli utenti dopo l'aggiornamento di un dataset esistente (e/o cambiamenti nel modo di accedere ai dati tramite API) o dopo la dismissione del dataset. Tali annunci possono avvenire sotto forma di pubblicazione di una nota in una sezione specifica del Portale Open Data (pagina "notizie" o "annunci"). In modo particolare, se l'aggiornamento o rimozione è avvenuta in seguito ad una o più richieste da parte degli utenti, è necessario rispondere in modo puntuale alla richiesta, informando dell'avvenuta modifica. Mantenere una buona relazione di comunicazione con gli utenti che interagiscono con l'Ente tramite il Portale Open Data è uno dei punti chiavi nella costruzione di una comunità Open Data attiva e partecipativa. Infine, una nota interna sulle modifiche apportate può essere emessa al Responsabile della Struttura ed al Referente Tecnico del Dataset, nonché ad altri soggetti potenzialmente interessati all'interno dell'Ente.

Allegati

Scheda Informativa dataset

La *Scheda Informativa dataset* (disponibile come **Allegato A.1**) è il documento di supporto alla fase di identificazione e censimento di potenziali nuovi dataset. Tale documento consente di procedere all'analisi dei dati per garantire che non ci siano potenziali impedimenti nel rendere i dati pubblicamente disponibili.

Questionario di rilevazione delle Banche dati

Il *Questionario di rilevazione delle Banche dati* (disponibile come **Allegato A.2**) è utilizzato per il censimento delle basi dati a titolarità regionale che sono utilizzate dall'ente dalle quali è possibile desumere una molteplicità di informazioni finalizzate alla pubblicazione di dataset tematici open data.

Dataset checklist

La Dataset Checklist (disponibile come **Allegato A.3**) attesta e registra, per ciascun dataset, l'esito delle verifiche richieste dal modello di processi e contiene i riferimenti per ottenere le informazioni censite nella "Scheda Informativa dataset" in fase di Identificazione. Questo deliverable fornisce un modo uniforme di classificare caratteristiche importanti dei dati, la loro performance e qualità generale. Fornisce anche la base per delineare un percorso chiaro e legittimato per il miglioramento o la dismissione del dataset.

Nominativi e riferimenti Team Open Data Regione Umbria

I *Nominativi e riferimenti Team Open Data Regione Umbria* (disponibile come **Allegato A.4**) contiene la lista dei nominativi ed i riferimenti e-mail delle figure che compongono il Team Open Data.

Glossario

APPS: termine utilizzato per indicare le applicazioni software sviluppate per operare su terminali Mobili.

DATASET: una collezione di dati, generalmente riguardanti una stessa organizzazione, che vengono erogati e gestiti congiuntamente.

LICENZA: in ambito informatico, il contratto, o altro strumento negoziale, con il quale si regolano i diritti morali e patrimoniali del software e dei dati nonché le condizioni d'uso e di distribuzione degli stessi.

SEGRETO DI STATO: un vincolo giuridico che determina l'esclusione dalla divulgazione “di atti, documenti, notizie, attività e ogni altra cosa la cui diffusione sia idonea a recare danno all’integrità della Repubblica, anche in relazione ad accordi internazionali, alla difesa delle istituzioni poste dalla Costituzione a suo fondamento, all’indipendenza della Stato rispetto ad altri Stati e alle relazioni con essi, alla preparazione e alla difesa militare dello Stato”, ponendo delle sanzioni nei confronti di chi violi l'obbligo di non divulgazione.

SEGRETO STATISTICO: i dati raccolti nell'ambito di rilevazioni statistiche comprese nel programma statistico nazionale da parte degli uffici di statistica non possono essere esternati se non in forma aggregata, in modo che non se ne possa trarre alcun riferimento relativamente a persone identificabili e possono essere utilizzati solo per scopi statistici.

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 (d.lgs. 36/2003).