

# The Open Data Choosing Framework

Processi, strumenti e metodi per la selezione dei dati giusti



A PIECE OF CONTENT OR DATA IS OPEN IF ANYONE IS FREE TO USE, REUSE, AND REDISTRIBUTE IT - SUBJECT ONLY, AT MOST, TO THE REQUIREMENT TO ATTRIBUTE AND SHARE-ALIKE.

> [opendefinition.org](http://opendefinition.org)

NUOVO ORO,  
GIACIMENTI,  
MINIERE E MINATORI

OBIETTIVI DEL  
PAPER

SCEGLIERE I DATI  
GIUSTI

UN FRAMEWORK PER  
LA SCELTA DEGLI  
OPEN DATA



## Nuovo oro, giacimenti, miniere e minatori

In un articolo apparso su The Times il 31 dicembre 2011, il creatore del web Berners Lee ha definito i dati la nuova materia prima del XXI secolo, arrivando a sostenere che i nostri dati oggi valgono oro.

*“(...) Nel Regno Unito quando i dati pubblici sono stati rilasciati, sono subito seguite le applicazioni, da quelle per trovare via cellulare un dentista del servizio sanitario nazionale alle aziende che utilizzano i dati disponibili sulla spesa per consigliare le autorità locali su come ottenere il miglior rapporto qualità-prezzo. Queste applicazioni aperte stanno creando nuove occasioni di business per i loro sviluppatori e grandi risorse per tutti noi. Prendiamo, per esempio, i programmi per cercare un autobus (vedi London Bus Stop Live o BusMate London) (...).*

*(...) Negli Stati Uniti, dove i dati sul clima sono già disponibili a tutti, il mercato del settore privato delle previsioni meteorologiche vale 1,5 miliardi di dollari l'anno (...).*

*Tratto da: There's gold to be mined from all our data. Tim Berners-Lee and Nigel Shadbolt  
(www.eprints.ecs.soton.ac.uk)*

Se il dato rappresenta il nuovo metallo prezioso del nostro tempo, le Pubbliche Amministrazioni che lo detengono ne costituiscono immensi giacimenti. Occorre, pertanto, cogliere il suggerimento di Lee e vestirsi dei panni degli antichi cercatori d'oro per andare alla scoperta di questi nuovi grandi filoni.

Per farlo è necessario avere delle mappe, per potersi orientare durante la ricerca, nonché una conoscenza e una profonda consapevolezza del capitale di informazioni che è generato dai processi della PA e che spesso è archiviato in attesa di essere scoperto, estratto, e valorizzato. Una volta individuati i giacimenti potenziali, è poi importante verificare se siano già state create delle miniere per raccogliere, estrarre e trasportare in superficie il nuovo oro, oppure se occorre intervenire per crearne di nuove o modernizzare quelle esistenti.

Le miniere, in questa metafora dell'Open Data altro non sono che le applicazioni che compongono i sistemi informativi della nostra PA e che, in ragione di numerosi fattori, possono essere sia moderne e in piena efficienza, sia obsolete e bisognose di interventi di manutenzione e rinnovamento.

I minatori, infine, possono essere assimilati ai processi operativi delle PA che, in funzione della qualità con cui sono eseguiti, garantiscono qualità e quantità del materiale da trasportare in superficie.

Cercare, estrarre e portare in superficie dati e informazioni pubbliche che possono essere acquisite, analizzate, incrociate e utilizzate liberamente per creare nuovi servizi per cittadini, imprese e istituzioni favorendo lo sviluppo economico, sociale e culturale di una comunità è il significato più profondo dell'Open Data.

Attraverso questa collana di paper Engineering intende dare il proprio contributo di idee e riflessioni per la diffusione e lo sviluppo dell'Open Data.

## Obiettivi del paper

▼  
**Il paper si occupa degli aspetti collegati al cosa pubblicare e non del come pubblicare.**

Obiettivi di questo paper sono quelli di fornire metodi e strumenti di management utilizzabili dai dipartimenti IT di Istituzioni ed Enti Pubblici al fine di:

- individuare e valutare il capitale di informazioni (i giacimenti) che può essere utile pubblicare in modalità Open Data;
- stabilire la capacità delle applicazioni componenti il sistema informativo dell'Ente (le miniere) di estrarre tali informazioni e renderle disponibili in modalità aperta.

La figura 1 riporta il percorso di implementazione dell'Open Data suggerito nell'ambito Open Data Manual curato dall'Open Knowledge Foundation e identifica le fasi del processo rispetto alle quali il contenuto del paper intende fornire un contributo in termini di approcci, metodologie e strumenti.

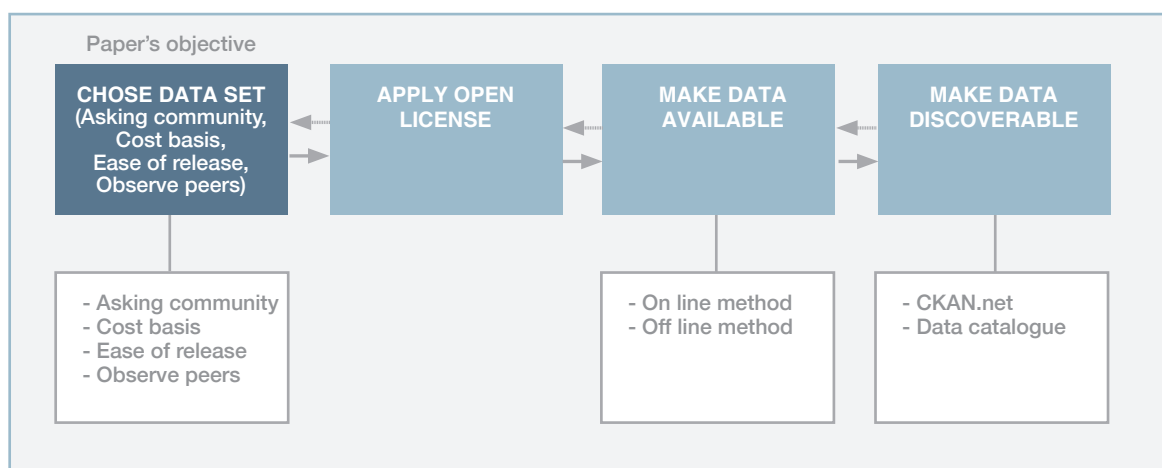


Figura 1 – I processi dell'Open Data Manual rispetto ai quali intende contribuire il paper.

Alcuni dei metodi e degli strumenti suggeriti nel paper possono essere utilizzati anche per favorire confronti tra diverse Amministrazioni (Observe peers) e iniziative di coinvolgimento di cittadini e imprese (Asking the community) nei programmi di rilascio degli Open Data.

Lo scopo principale del lavoro, tuttavia, è quello di fornire un framework per implementare all'interno della Pubblica Amministrazione processi di governance del percorso di selezione e pubblicazione dei dati, non quello di coinvolgere soggetti esterni nella scelta, nella valutazione e nel confronto delle "intenzioni" di pubblicazione o dei dataset eventualmente già pubblicati.

Non è invece obiettivo di questo paper discutere e suggerire modalità tecniche e operative attraverso le quali, una volta individuati i dati adatti, è possibile procedere alla loro pubblicazione e diffusione. Non si discuteranno, ad esempio, né aspetti relativi all'apertura legale dei dati attraverso la definizione di opportune licenze d'uso, né aspetti relativi alle modalità con le quali procedere all'apertura tecnica dei dati.

## Scegliere i dati giusti



**L'Open Data è una grande opportunità per sviluppare la relazione tra l'amministrazione e la cittadinanza, ma come ogni processo di comunicazione ha bisogno di essere gestito con attenzione.**

Pubblicare i dati in modalità Open è un tema di grande attualità. La stampa specializzata e generalista, il mondo dell'economia e delle istituzioni, la politica e i movimenti si occupano ormai quotidianamente di dati aperti, generando una spinta che, grazie anche alla cassa di risonanza del web, ha su molti amministratori pubblici un effetto esponenziale sulla percezione di urgenza e indifferibilità di pubblicare il più rapidamente possibile qualche dato.

In un momento come questo, in cui le spinte all'Open Data sono molto intense, la capacità di resistere alla tentazione di pubblicare quello che è possibile pubblicare in tempi rapidi, senza confondere il mezzo (pubblicare dati) con il fine (mettere a disposizione della comunità dati utili e utilizzabili) richiede molta forza e una certa disciplina.

La pubblicazione di un dato, infatti, rappresenta a tutti gli effetti un processo di comunicazione e – in quanto tale – un'operazione assai delicata per una Pubblica Amministrazione che richiederebbe:

- **Consapevolezza**, per quanto riguarda la qualità del dato disponibile e il rischio associato alla pubblicazione e all'utilizzo da parte di cittadini e aziende di dati poco affidabili;
- **Immaginazione**, per quanto riguarda la capacità di simulare possibili scenari di utilizzo dei dati e stabilire così quali dati avranno più impatti di altri sulla capacità di generare crescita economica, trasparenza amministrativa, cultura e progresso una volta messi liberamente a disposizione.

La selezione dei dati, in altre parole, è un problema complesso che coinvolge diverse variabili e che può essere utile affrontare in maniera strutturata e continuativa nel tempo, con l'obiettivo di aiutare le organizzazioni pubbliche a darsi una vera e propria agenda dell'Open Data, ossia una sequenza strategica di pubblicazione e aggiornamento dei loro dati basata su criteri misurabili di potenzialità e appropriatezza della pubblicazione.

## Un framework per la scelta degli open data

La Direzione Consulenza per la Pubblica Amministrazione di Engineering ha definito un framework per implementare programmi strutturati di ricerca e selezione dei dati da sottoporre a pubblicazione. Il framework è composto da quattro processi, un toolkit formato da cinque strumenti e un insieme di tecniche di management in grado di accompagnare lo svolgimento dei processi e l'impiego dei diversi strumenti.

I quattro processi del framework sono i seguenti:

- OPEN DATA VEINS & MINES MAPPING** È il processo attraverso il quale è possibile definire le coordinate da seguire per ricercare e individuare all'interno dell'organizzazione i luoghi in cui possono essere presenti i giacimenti di dati.
- OPEN DATA VEINS EXPLORATION** È il processo attraverso il quale è possibile verificare l'effettiva consistenza e ricchezza dei dati presenti nei giacimenti individuati e presi in considerazione.
- OPEN DATA MINES ASSESSMENT** È il processo attraverso il quale, una volta individuata la presenza di dati, è possibile stabilire se esistono miniere di applicazioni in grado di trattarli e, in caso affermativo, valutare il livello di complessità del processo di estrazione e trasporto "in superficie" (pubblicazione) dei dati.
- OPEN DATA MEASUREMENT** È il processo attraverso il quale è possibile a seguito dell'esplorazione e del successivo assessment associare a ciascuna miniera un valore numerico che permetta di stilare una lista in grado di stabilire una sequenza di estrazione e trasporto in superficie (pubblicazione) dei dati.

L'elenco dei processi e il loro ordine di esecuzione non è prescrittivo. Nell'applicazione del framework, la singola organizzazione dovrebbe individuare e selezionare i processi più adatti alla propria realtà e implementarli in maniera coerente ai propri obiettivi specifici e – se necessario – integrati con altri processi non previsti nell'ambito di questo paper.

Il toolkit che può essere impiegato a supporto dei processi appena descritti è composto dai seguenti strumenti:

- **Open Data Fields and Mines Map:** è un modello che permette di sistematizzare e generalizzare l'architettura applicative dell'organizzazione in maniera tale da rendere più facile e immediato individuare e confrontare i possibili Open Data;
- **Open Data Applications Potential Questionnaire:** è un questionario attraverso il quale è possibile ottenere informazioni circa i dati trattati da un applicazione oltre che giudizi e suggerimenti sui possibili utilizzi di tali dati da parte degli utenti delle applicazioni ma anche di soggetti esterni all'organizzazione come, ad esempio, i cittadini;

- **Open Data Applications Appropriateness Check list:** è un questionario attraverso il quale è possibile valutare il livello di prontezza dell'applicazione a supportare un programma di Open Data sia dal punto di vista tecnico, sia dal punto di vista organizzativo;
- **Open Data Application Scorecard:** è una scheda che, per singola applicazione, raggruppa, sintetizza e misura all'interno di aree omogenee i risultati delle analisi svolte sui livelli di potenziale e appropriatezza della pubblicazione;
- **Open Data Application Index:** è un indicatore ottenuto attribuendo dei pesi ai punteggi ottenuti dalle singole applicazioni e riportati all'interno delle singole scorecard che consolidati, in unico indice di sintesi, permettono di stilare una graduatoria relativamente alla propensione all'Open Data del portafoglio applicativo e di fornire elementi utili alla definizione di una strategia e una roadmap di pubblicazione dei dati.

Gli strumenti proposti sono funzionali allo svolgimento dei quattro processi del framework. Anche in questo caso è bene precisare che stabilire in maniera prescrittiva che tutti gli strumenti proposti a supporto di un certo processo debbano essere implementati per il suo corretto svolgimento non è obiettivo del framework. I singoli strumenti debbono essere selezionati in ciascun contesto in funzione delle esigenze e degli obiettivi specifici e – se necessario – integrati con altri strumenti non previsti nell'ambito di questo paper.

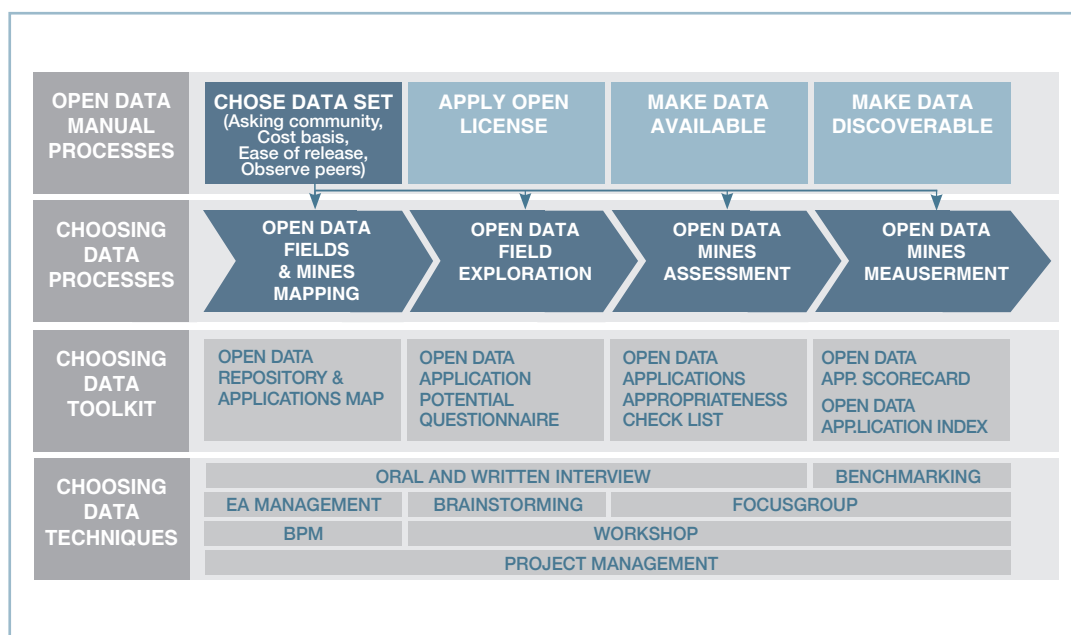


Figura 2 – Open Data selection Framework



**I Dati sono ovunque,  
ma la capacità di  
strutturarli in informazioni  
intelligenti e architetture  
finalizzate NO.  
In questo differenziale  
risiede il valore  
dell'ingegneria  
dell'informazione.**

Le tecniche di management che possono essere utilizzate nello sviluppo dei processi e nell'applicazione degli strumenti sono rappresentate da:

Oral & Written Interview / Workshop / Focusgroup / Brainstorming / Benchmarking / Business Process Management / Enterprise Architecture Management / Project Management .

Le tecniche proposte possono essere impiegate all'interno dei diversi processi e a supporto dei vari strumenti presentati in modalità che possono variare da contesto a contesto. L'elenco delle tecniche, inoltre, non è esaustivo, è possibile aggiungerne altre e combinarle in differenti modalità a seconda delle esigenze specifiche delle singole organizzazioni.