

CLASSIFICAZIONE DEL DOCUMENTO: CONSIP PUBLIC

ALLEGATO 5B

CAPITOLATO TECNICO

LOTTO 3

PROCEDURA RISTRETTA PER L’AFFIDAMENTO DEI SERVIZI DI *CLOUD COMPUTING*, DI SICUREZZA, DI REALIZZAZIONE DI PORTALI E SERVIZI ONLINE E DI COOPERAZIONE APPLICATIVA PER LE PUBBLICHE AMMINISTRAZIONI (ID SIGEF 1403).



| | |
|--|-----------|
| CLASSIFICAZIONE DEL DOCUMENTO: CONSIP PUBLIC | 1 |
| 1.1. Tipologia Servizi L3 - Cooperazione Applicativa | 5 |
| 1.1.1. Servizio L3.S1 - Porta di dominio | 6 |
| 1.1.1.1. Descrizione del Servizio | 6 |
| 1.1.1.1. Requisiti funzionali | 7 |
| 1.1.1.2. Requisiti tecnici e tecnologici | 7 |
| 1.1.1.3. Tipologia del Servizio | 10 |
| 1.1.1.4. Parametri di valutazione economica | 11 |
| 1.1.2. Servizio L3.S2 - Realizzazione interfacce web services | 11 |
| 1.1.2.1. Descrizione del Servizio | 11 |
| 1.1.2.2. Requisiti funzionali | 13 |
| 1.1.2.3. Requisiti tecnici e tecnologici | 13 |
| 1.1.2.4. Tipologia del Servizio | 15 |
| 1.1.2.5. Parametri di valutazione economica | 15 |
| 1.1.3. Servizio L3.S3 - Realizzazione client per la fruizione dei servizi | 16 |
| 1.1.3.1. Descrizione del Servizio | 16 |
| 1.1.3.2. Requisiti funzionali | 16 |
| 1.1.3.3. Requisiti tecnici e tecnologici | 17 |
| 1.1.3.4. Tipologia del Servizio | 17 |
| 1.1.3.5. Parametri di valutazione economica | 17 |
| 1.1.4. Servizio L3.S4 - Orchestrazione | 18 |
| 1.1.4.1. Descrizione del Servizio | 18 |



| | | |
|----------|--|----|
| 1.1.4.2. | Requisiti funzionali | 19 |
| 1.1.4.3. | Requisiti tecnici e tecnologici | 19 |
| 1.1.4.4. | Tipologia del Servizio | 20 |
| 1.1.4.5. | Parametri di valutazione economica | 20 |
| 1.2. | Tipologia Servizi L3 - Open Data | 22 |
| 1.2.1. | Servizio L3.S5 - Open Data | 23 |
| 1.2.1.1. | Descrizione del Servizio | 23 |
| 1.2.1.2. | Requisiti funzionali | 30 |
| 1.2.1.3. | Requisiti tecnici e tecnologici | 31 |
| 1.2.1.4. | Tipologia del Servizio | 31 |
| 1.2.1.5. | Parametri di valutazione economica | 31 |
| 1.3. | Tipologia Servizi L3 - Big Data | 35 |
| 1.3.1. | Servizio L3.S6 - Supporto alla memorizzazione dei Big Data | 36 |
| 1.3.1.1. | Descrizione del Servizio | 36 |
| 1.3.1.2. | Requisiti funzionali | 41 |
| 1.3.1.3. | Requisiti tecnici e tecnologici | 41 |
| 1.3.1.4. | Tipologia del Servizio | 42 |
| 1.3.1.5. | Parametri di valutazione economica | 42 |
| 1.3.2. | Servizio L3.S7 - Supporto all'analisi dei Big Data..... | 45 |
| 1.3.2.1. | Descrizione del Servizio | 45 |
| 1.3.2.2. | Requisiti funzionali | 50 |
| 1.3.2.3. | Requisiti tecnici e tecnologici | 50 |



| | | |
|----------|--|----|
| 1.3.2.4. | Tipologia del Servizio | 51 |
| 1.3.2.5. | Parametri di valutazione economica | 52 |



PREMESSA

In questo capitolo si fornisce la descrizione dei servizi previsti per il Lotto 3, raggruppati nelle seguenti tre macro-tipologie di servizio:

- Servizi di Cooperazione Applicativa
- Servizi di Open Data
- Servizi di Big Data

Sono previste due modalità di erogazione dei servizi:

- *On premise*, presso le strutture dell'Amministrazione richiedente, siano esse proprie oppure acquisite temporaneamente;
- *As a service*, servizi erogati mediante il Centro Servizi del Fornitore Aggiudicatario.

Si rimanda ai successivi paragrafi per l'indicazione delle specifiche modalità di erogazione previste per i singoli servizi compresi in ognuna delle tre macro-tipologie.

Di seguito vengono forniti i requisiti di tutti i servizi previsti nel Lotto. Salvo indicazioni diverse, tali requisiti sono da considerarsi minimi.

1.1. Tipologia Servizi L3 - Cooperazione Applicativa

La Cooperazione Applicativa realizza la modalità tecnica e organizzativa, proposta dal Codice dell'Amministrazione Digitale e dalle relative regole tecniche DPCM del 1 aprile 2008, per il colloquio tra i sistemi informativi delle Pubbliche Amministrazioni nell'ambito del Sistema Pubblico di Connettività. L'obiettivo della Cooperazione Applicativa è di rendere disponibile in modalità telematica, agli interlocutori istituzionali, le informazioni e i procedimenti amministrativi che competono ad una data Amministrazione.

I servizi di Cooperazione Applicativa previsti dal presente Capitolato hanno lo scopo di assistere le Amministrazioni nell'adozione del modello e nella realizzazione di scenari compiuti di Cooperazione Applicativa, anche nell'ottica di adempimenti previsti dal Codice dell'Amministrazione Digitale. In particolare, il Capitolato comprende i seguenti servizi, la cui applicabilità e il cui profilo verranno descritti nei successivi paragrafi:

- **L3.S1 - Porta di dominio**
- **L3.S2 - Realizzazione interfacce web services (“wrapping”)**
- **L3.S3 - Realizzazione client per la fruizione dei servizi**
- **L3.S4 - Orchestrazione**



1.1.1. Servizio L3.S1 - Porta di dominio

1.1.1.1. Descrizione del Servizio

La porta di dominio è un componente logico infrastrutturale che realizza il colloquio tra i domini delle Amministrazioni, collegati tra di loro tramite il Sistema Pubblico di Connettività; l'insieme di tutte le porte di dominio presenti e i servizi di registro SICA¹ (Servizi di Interoperabilità e di Cooperazione e Accesso) costituiscono l'infrastruttura di Cooperazione Applicativa.

La porta di dominio, come unico punto di accesso al dominio applicativo di una Pubblica Amministrazione, ha lo scopo di introdurre un grado di trasparenza rispetto alla distribuzione fisica dei sistemi presenti all'interno del dominio, e, implementando funzionalità centralizzate di *proxy/routing* applicativo rispetto agli stessi, di realizzare le necessarie funzionalità di sicurezza relative all'accesso ai servizi del dominio.

La porta di dominio ha il compito di veicolare le richieste verso i servizi applicativi e le relative risposte provenienti dagli stessi formulate secondo il protocollo SOAP (web services) che prevede un formato personalizzato di messaggio denominato "*busta e-gov*". Tale personalizzazione consiste nella presenza di un header customizzato, denominato "Intestazione"², predisposto a veicolare nei messaggi le informazioni relative ai soggetti coinvolti nello scambio, l'identificatore del servizio invocato e l'operazione richiesta.

La porta di dominio, intesa come *soap engine*, si configura quindi come l'agente associato all'header "Intestazione", e come tale ha il compito in fase di ricezione delle chiamate di servizio (i.e. funzionamento in modalità "porta applicativa") di sovraintendere alla gestione dell'autorizzazione all'erogazione dello stesso (*policy enforcement point*), in base agli accordi di servizio sottoscritti dalle Amministrazioni coinvolte nel rapporto erogatore-fruttore, nonché all'instradamento delle richieste verso il sistema di back-end in grado di soddisfarle.

Viceversa, in fase di invocazione di un servizio, su richiesta da parte di un applicativo client facente parte del dominio di una Amministrazione (i.e. funzionamento in modalità "porta delegata"), la porta di dominio dovrà essere in grado di individuare il dominio di erogazione del servizio richiesto e, preparata la busta di *e-gov* aggiornando correttamente i contenuti dell'header "Intestazione", di inoltrare il messaggio di richiesta verso la porta di dominio dell'Amministrazione erogatrice. In generale, a una Amministrazione è associata un'unica porta di dominio.

¹ <http://archivio.digitpa.gov.it/reti-della-pa/documenti-tecnico-operativi>

² Le specifiche sono presenti nei documenti "SPCoop-Busta-e-Gov_v1.1" e "SPCoop-Linee guida per la busta di e-gov" pubblicati nel sito istituzionale AGID (www.agid.gov.it)



Il servizio in oggetto “Porta di dominio” prevede la messa a disposizione di un sistema che realizzi le funzionalità della porta di dominio sopra elencate.

La porta di dominio deve:

- essere dedicata all’Amministrazione e resa disponibile alla stessa in cloud in modalità “*as-a-Service*”;
- essere qualificata per conto dell’Amministrazione prima di essere messa in esercizio.

Le attività principali previste per l’erogazione del servizio “Porta di dominio” sono:

- esercizio, mantenimento ed aggiornamento del sistema che realizza le funzionalità di porta di dominio;
- interfacciamento tra la porta di dominio e i servizi di back-end dell’Amministrazione, nel caso di funzionamento come *porta applicativa*, oppure, nel caso di funzionamento come *porta delegata*, degli applicativi dell’Amministrazione medesima operanti in modalità client.

1.1.1.1. Requisiti funzionali

Non applicabile per il servizio “Porta di dominio”.

1.1.1.2. Requisiti tecnici e tecnologici

Nell’ambito del servizio “Porta di dominio”, il Fornitore deve garantire almeno i seguenti requisiti tecnici e tecnologici:

- gestire l’instaurazione di sessioni https con le altre porte di dominio. Il colloquio deve avvenire solo con porte di dominio qualificate facendo uso dei certificati SSL rilasciati dall’Agenzia per l’Italia Digitale;
- gestire correttamente la busta di *e-gov*:
 - utilizzare la sintassi della busta di *e-gov* indicata nel documento di Specifiche della busta di *e-gov* e nelle Raccomandazioni previste dal documento “Linee guida all’uso della busta di *e-gov*” pubblicato dall’Agenzia per l’Italia Digitale;
 - la porta di dominio deve accettare tutte e sole le buste di *e-gov* conformate come specificato nei suddetti documenti e emettere solo buste conformi a come negli stessi specificato;
- in caso di funzionamento in modalità “porta applicativa”:
 - la porta di dominio deve ricevere le chiamate di servizio provenienti dalle porte di dominio delle Amministrazioni esterne e inoltrarle al servizio di back-end di competenza previa verifica della correttezza dell’invocazione



e dell'autorizzazione alla fruizione. Tali attività devono essere operate sulla base delle informazioni veicolate dall'header "Intestazione" della busta di e-gov di richiesta del servizio;

- in base alle informazioni presenti nei seguenti elementi dell'header "Intestazione":
 - a) "/Intestazione/IntestazioneMessaggio/Mittente",
 - b) "/Intestazione/IntestazioneMessaggio/Destinatario",
 - c) "/Intestazione/IntestazioneMessaggio/Servizio";
 - la porta di dominio deve verificare il "non ripudio" dell'invocazione del servizio, ovvero che la stessa provenga dall'Amministrazione (soggettoSPCoop) referenziata nell'elemento a);
 - in fase di autorizzazione all'invocazione del servizio, la porta deve verificare, sulla base del contenuto dell'elemento b), che le richieste pervenute siano relative a servizi resi disponibili dall'Amministrazione a essa afferente e, sulla base del contenuto dell'elemento c), che le richieste provengano da Amministrazioni che hanno sottoscritto l'accordo di servizio relativo al servizio invocato;
- la correttezza della corrispondenza del formato del body della busta di e-gov relativa alla richiesta con quanto previsto nell'accordo di servizio (WSDL) deve essere demandata al servizio di back-end;
- la porta di dominio deve invocare il servizio di back-end sulla base del relativo WSDL reso disponibile dall'Amministrazione;
- in caso di funzionamento in modalità "porta delegata":
 - la porta di dominio deve ricevere le chiamate di servizio provenienti dai sistemi informativi interni al dominio dell'Amministrazione e inoltrarle verso la porta di dominio dell'Amministrazione erogatrice il servizio richiesto;
 - la porta di dominio deve essere in grado di individuare la corretta porta e il corretto servizio da invocare e predisporre opportunamente la busta di e-gov;



- gestire la tracciatura dei messaggi di e-gov transitanti, utilizzando il formato specificato in Appendice 1 del documento “Sistema pubblico di cooperazione: Porta di Dominio”³, e mantenere una storicizzazione delle tracce emesse;
- rendere disponibile alle Amministrazioni una consolle di monitoraggio per:
 - verificare lo stato della porta di dominio,
 - consultare ed estrarre le tracciate dei messaggi;
 - monitorare le chiamate di servizio transitate attraverso la porta,
- gestire eventuali *header* relativi alla sicurezza presenti nella busta di e-gov specificati secondo lo standard “WS-Security”, finalizzati alla:
 - firma dei messaggi,
 - cifratura dei messaggi,
 - veicolazione token SAML relativi ad asserzioni di identità e di attributo. Gli eventuali token SAML contenuti nella busta e-gov e relativi ad asserzioni di identità o di attributo non gestiti direttamente dalla porta di dominio devono essere propagati integri verso il dominio dell’Amministrazione e senza comprometterne l’utilizzo da parte di altre entità preposte allo scopo;
- mantenere inalterate nell’attraversamento della porta eventuali messaggi nel formato “SOA with Attachment/MTOM” (gestione del “SOAP with attachment/MTOM”);
- prevedere un colloquio con i sistemi di back-end, sia in funzionamento “porta delegata” che in quello “porta applicativa”, in modalità web service su canali https;
- prevedere una modalità per l’interfacciamento/integrazione della porta di dominio con i sistemi informativi del dominio interno dell’Amministrazione, propedeutica alla erogazione (modalità “porta applicativa”) o all’invocazione (modalità “porta delegata”) dei servizi. Tale modalità dovrà essere implementata sulla base di specifiche in formato WSDL;
- gestire l’autorizzazione al servizio tenendo conto anche del profilo dell’utente che opera per conto dell’Amministrazione richiedente, ove questa modalità sia prevista nell’accordo di servizio.

³ Documento disponibile al seguente link web:

http://www.agid.gov.it/sites/default/files/documentazione/spcoop-portadominio_v1.1_0.pdf



1.1.1.3. Tipologia del Servizio

Il servizio “Porta di dominio” dovrà essere erogato in modalità “continuativa” di tipo “*as-a-Service*”. In tale modalità, la porta di dominio è erogata dal Fornitore presso il proprio Centro Servizi.



1.1.1.4. Parametri di valutazione economica

La Modalità di remunerazione del servizio “Porta di dominio” è “a canone (su base annuale)”. Ai fini della valutazione economica del servizio “Porta di dominio” è richiesta una quotazione [€/anno] per il canone annuale relativo all’erogazione della porta di dominio in modalità “*as-a-Service*”, comprensivo delle attività di esercizio, mantenimento e aggiornamento del sistema realizzante la porta nonché delle attività propedeutiche all’interfacciamento dei servizi esposti attraverso la porta.

1.1.2. Servizio L3.S2 - Realizzazione interfacce web services

1.1.2.1. Descrizione del Servizio

Il servizio in oggetto “Realizzazione interfacce web services” ha come obiettivo la messa a punto delle componenti software necessarie a far sì che un sistema informativo, già esistente presso l’Amministrazione, possa rendere accessibili banche dati o funzionalità, già presenti e disponibili, in modalità “web services”.

Il servizio è offerto per essere di ausilio alle Amministrazioni nell’attuazione delle prescrizioni del Codice per l’Amministrazione Digitale, in particolare per gli articoli 58 e 60 relativi alla fruibilità dei dati pubblici e alle banche dati di interesse nazionale.

Nel caso di accesso alle banche dati detenute dall’Amministrazione, lo scopo è di fornire delle funzioni di interrogazione invocabili attraverso web services.

Nel caso più generale della fruibilità dei dati pubblici, a seguito di stipula di convenzioni ai sensi dell’art. 58 comma 2 del Codice per l’Amministrazione Digitale, il servizio è finalizzato alla messa a punto delle componenti software necessarie a far sì che un sistema informativo, già esistente presso le Amministrazioni, possa rendere accessibili in modalità web services funzionalità già presenti e disponibili. Tali funzionalità devono realizzare passi di un procedimento amministrativo istituzionale di competenza dell’Amministrazione e oggetto delle citate convenzioni.

In entrambi i casi previsti, i web services realizzati dovranno costituire servizi di Cooperazione Applicativa da rendere fruibili attraverso la porta di dominio, con accordo di servizio pubblicato nel registro SICA.

Per realizzare gli obiettivi del servizio si prevede l’implementazione di un adattatore (denominato “*wrapper*”) in grado di gestire correttamente invocazioni secondo il protocollo SOAP su *http* e di interfacciarsi con il sistema informativo o con la base di dati interni all’Amministrazione, al fine di soddisfare le richieste pervenute. Il web services così predisposto dovrà essere pronto per l’interfacciamento con la porta di dominio perché possa essere fruito secondo le modalità previste per la Cooperazione Applicativa.



Il servizio di “Realizzazione interfacce web services” è comprensivo di 12 (dodici) mesi di garanzia per i web services realizzati.



1.1.2.2. Requisiti funzionali

Nell'ambito del servizio "Realizzazione interfacce web services", il Fornitore deve garantire almeno i seguenti requisiti funzionali:

- predisposizione dell'Accordo di Servizio Parte Comune da pubblicare nel registro SICA. La predisposizione dell'Accordo di Servizio Parte Comune deve essere fatta in accordo con quanto specificato nel documento "spcoop-accordoservizio_v1.1" emesso da DigitPA - ora Agenzia per l'Italia Digitale - reperibile presso il sito istituzionale della stessa Agenzia;
- supporto all'Amministrazione nella pubblicazione dell'Accordo di Servizio Parte Comune. La pubblicazione di tale Accordo presso il registro SICA avviene a cura dell'Amministrazione con l'assistenza del Fornitore secondo le modalità descritte nel Manuale utente del portale dei servizi SICA;
- gestione dell'esercizio, della manutenzione correttiva (MAC) ed evolutiva (MEV) del sistema realizzato con il servizio in oggetto;
- opzionalmente, su richiesta dell'Amministrazione, fornitura di una piattaforma application server virtualizzata erogata in modalità "as-a-Service", destinata ad ospitare tutte le componenti *wrapper* acquisite dall'Amministrazione attraverso questo servizio. Tale piattaforma dovrà essere opportunamente dimensionata e quindi avere una capacità elaborativa adeguata alle necessità e alla numerosità dei sistemi ospitati. Tale piattaforma dovrà essere mantenuta e aggiornata dal Fornitore.

1.1.2.3. Requisiti tecnici e tecnologici

Nell'ambito del servizio "Realizzazione interfacce web services", il Fornitore deve garantire almeno i seguenti requisiti tecnici e tecnologici:

- progettazione dell'interfaccia web services in linguaggio WSDL v1.1;
- il web services da predisporre deve mettere a disposizione le "operation" che realizzano:
 - i passi del *workflow* associato al procedimento amministrativo che l'Amministrazione vuole rendere fruibile in Cooperazione Applicativa,
 - le interrogazioni sulla base di dati che ospita le informazioni che l'Amministrazione vuole rendere fruibile in Cooperazione Applicativa.

In entrambi i casi parametri in ingresso e in uscita delle "operation" devono essere opportunamente scelti in modo da evitare operazioni ridondanti. Il WSDL risultante dalla fase di progettazione deve essere infine compatibile con le raccomandazioni WS-I;



- realizzazione del codice software in grado di gestire le chiamate SOAP dell'interfaccia web services. Sulla base del WSDL relativo all'interfaccia progettata, devono essere realizzati i componenti software del *wrapper* in grado di:
 - ricevere le chiamate SOAP,
 - attivare i connettori per l'invocazione dei sistemi di back-end che realizzano le funzionalità associate alla varie “*operation*” presenti nell'interfaccia,
 - restituire i risultati elaborati in modalità SOAP;
- in caso di fornitura di una piattaforma application server virtualizzata erogata in modalità “*as-a-Service*”, tale piattaforma deve essere in grado di poter elaborare in parallelo almeno 10 (dieci) richieste ogni secondo nel 95% dei casi, e almeno 8 (otto) richieste ogni secondo nel 100% dei casi;
- in caso di fornitura di una piattaforma application server virtualizzata erogata in modalità “*as-a-Service*”, tale piattaforma deve essere in grado di poter mantenere in coda almeno 10 (dieci) richieste ogni secondo nel 95% dei casi, e almeno 8 (otto) richieste ogni secondo nel 100% dei casi;
- realizzazione di un connettore verso il sistema informativo dell'Amministrazione, in grado di invocare nel linguaggio nativo del sistema le funzionalità, già presenti, da rendere fruibili o, nel caso di Basi di Dati, realizzare le interrogazioni previste. Gli scenari di integrazione previsti nel servizio sono i seguenti:
 - a) **Scenario 1:** sistemi informativi realizzati in linguaggi orientati agli oggetti (i.e. java, .net) o in linguaggi compilati (i.e. C, C++). Presupposto necessario per l'applicabilità del servizio in questo scenario è la possibilità di invocare le funzionalità richieste attraverso opportune funzioni o procedure già esistenti, nel caso di linguaggi strutturati, o metodi su classi già disponibili.
 - b) **Scenario 2:** sistemi informativi interfacciabili attraverso basi di dati relazionali o attraverso basi di dati non relazionali (i.e. basi di dati gerarchiche e/o reticolari). Presupposto necessario per l'applicabilità del servizio in questo scenario è la possibilità di realizzare, nel caso di interfacciamento attraverso basi di dati relazionali, le funzionalità di interesse attraverso delle query SQL e, nel caso di interfacciamento attraverso basi di dati non relazionali, le funzionalità di interesse attraverso delle interrogazioni operate attraverso specifici connettori;



1.1.2.4. Tipologia del Servizio

Il servizio “Realizzazione interfacce web services” dovrà essere erogato in modalità progettuale “*on-premise*”.

In caso di richiesta da parte dell’Amministrazione di una piattaforma virtualizzata, tale piattaforma sarà erogata dal Fornitore presso il proprio Centro Servizi in modalità “*as-a-Service*”.

1.1.2.5. Parametri di valutazione economica

Ai fini della valutazione economica del servizio “Realizzazione interfacce web services”, sono di seguito esplicitati i relativi parametri:

Realizzazione interfacce web services

La modalità di remunerazione delle attività di realizzazione delle interfacce web services è “a corpo” sulla base del numero di “*operation*” da realizzare per i web services.

Ai fini della valutazione economica è richiesta una quotazione [€/operation] per la realizzazione di una “*operation*”, indipendentemente dallo Scenario di integrazione previsto dal servizio (§1.1.2.3 “Requisiti tecnici e tecnologici”).

Il numero massimo di “*operation*” remunerate per singolo web services è pari a 6 (sei).

Manutenzione interfacce web services

La modalità di remunerazione per le attività di gestione dell’esercizio, della manutenzione correttiva (MAC) ed evolutiva (MEV) dei web services realizzati con il servizio in oggetto è a “canone annuale” in base al numero di *operation* realizzate su cui effettuare la manutenzione. Tale canone decorre dal termine della garanzia di 12 (dodici) mesi compresa nella realizzazione dei web services, con durata massima sino alla scadenza del Contratto Quadro.

Ai fini della valutazione economica è richiesta una quotazione [€/canone annuo operation] per il canone annuale di una singola *operation* realizzata nell’ambito del servizio.

Piattaforma application server virtualizzata “as-a-Service”

La modalità di remunerazione per la fornitura “*as-a-Service*” di una piattaforma application server virtualizzata, destinata ad ospitare tutte le componenti *wrapper* acquisite dall’Amministrazione attraverso questo servizio, è a “canone annuale”.

Ai fini della valutazione economica è richiesta una quotazione [€/anno] per il canone annuale della piattaforma virtualizzata.



1.1.3. Servizio L3.S3 - Realizzazione client per la fruizione dei servizi

1.1.3.1. Descrizione del Servizio

Il servizio “Realizzazione client per la fruizione dei servizi” ha come obiettivo la messa a punto delle componenti software necessarie alla fruizione dei servizi di Cooperazione Applicativa già pubblicati sul registro SICA e per le finalità previste del Codice per l’Amministrazione Digitale agli articoli 58 e 60, relativi alla fruibilità dei dati pubblici e alle banche dati di interesse nazionale.

Sono previste due tipologie di client in base alla modalità di utilizzo dei servizi di Cooperazione Applicativa:

- **Client per interazione “*Application to Application*”** - sviluppo di un software in modo da consentire che un sistema informativo, già esistente presso l’Amministrazione, possa invocare i servizi pubblicati sul registro SICA realizzando un colloquio di tipo *Application to Application*;
- **Client per interazione “*User to Application*”** - realizzazione di una consolle a uso dell’Amministrazione che consenta l’invocazione dei servizi e la presentazione dei risultati in modalità “*User to Application*”.

Il servizio di “Realizzazione client per la fruizione dei servizi” è comprensivo di 12 (dodici) mesi di garanzia per i *client* realizzati.

1.1.3.2. Requisiti funzionali

Si riportano di seguito i requisiti funzionali minimi che il Fornitore deve obbligatoriamente garantire per il servizio “Realizzazione client per la fruizione dei servizi”:

- stesura e pubblicazione nel registro SICA dell’Accordo di Servizio Parte Specifica;
- in caso di sviluppo del client per interazione “*Application to Application*”, dovrà essere prevista la realizzazione di un componente integrato con un preesistente sistema informativo di back-end in grado di invocare, per il tramite della porta di dominio, le *operation* previste dalla componente WSDL dell’accordo di servizio e relativa gestione delle risposte;
- in caso di sviluppo del client per interazione “*User to Application*”, dovrà essere prevista la realizzazione di una web application in grado di gestire le invocazioni delle *operation* previste dalla componente WSDL dell’accordo di servizio, acquisendo i dati di ingresso da parte dell’utente e fornendo a video una presentazione dei risultati;
- opzionalmente, su richiesta dell’Amministrazione, fornitura di una piattaforma application server virtualizzata erogata in modalità “*as-a-Service*”, destinata ad



ospitare tutte le componenti *client* acquisite dall'Amministrazione attraverso questo servizio. Tale piattaforma dovrà essere opportunamente dimensionata e quindi avere una capacità elaborativa adeguata alle necessità e alla numerosità dei sistemi ospitati. Tale piattaforma dovrà essere mantenuta e aggiornata dal Fornitore.

1.1.3.3. Requisiti tecnici e tecnologici

Nell'ambito del servizio "Realizzazione client per la fruizione dei servizi", il Fornitore deve garantire almeno i seguenti requisiti tecnici e tecnologici:

- in caso di fornitura di una piattaforma application server virtualizzata erogata in modalità "as-a-Service", tale piattaforma deve essere in grado di poter elaborare in parallelo almeno 10 (dieci) richieste ogni secondo nel 95% dei casi, e almeno 8 (otto) richieste ogni secondo nel 100% dei casi;
- in caso di fornitura di una piattaforma application server virtualizzata erogata in modalità "as-a-Service", tale piattaforma deve essere in grado di poter mantenere in coda almeno 10 (dieci) richieste ogni secondo nel 95% dei casi, e almeno 8 (otto) richieste ogni secondo nel 100% dei casi.

1.1.3.4. Tipologia del Servizio

Il servizio "Realizzazione client per la fruizione dei servizi" dovrà essere erogato in modalità progettuale "on-premise". Il servizio in oggetto si intende svolto mediante sistemi e infrastrutture del Centro Servizi messo a disposizione dal Fornitore; l'Amministrazione potrà comunque riservarsi la possibilità di disporre dei team di sviluppo presso le proprie sedi.

1.1.3.5. Parametri di valutazione economica

Il costo del servizio "Realizzazione client per la fruizione dei servizi" è determinato dalle seguenti voci di costo:

Realizzazione *client*

La Modalità di remunerazione delle attività di sviluppo del *client* si basa sulla seguente modalità di dimensionamento del software:

- Tecnica di misura Punti Funzione - IFPUG vers. 4.3. o successive per i requisiti funzionali del *client*.

Ai fini della valutazione economica del servizio "Realizzazione client per la fruizione dei servizi" è richiesta una quotazione [€/PF] per un singolo Punto Funzione.



Manutenzione *client*

La modalità di remunerazione per le attività di gestione dell'esercizio e della manutenzione correttiva (MAC) dei *client* realizzati con il servizio in oggetto è a "canone annuale". Tale canone decorre dal termine della garanzia di 12 (dodici) mesi compresa nella realizzazione dei *client*, con durata sino alla scadenza del Contratto Quadro.

La manutenzione evolutiva (MEV) di un *client* realizzato per la fruizione dei servizi dovrà essere assimilata a un nuovo sviluppo limitatamente alle nuove funzionalità aggiunte al sistema.

Ai fini della valutazione economica è richiesta una quotazione [€/PF] per il canone annuale di un singolo Punto Funzione sviluppato nell'ambito del servizio di realizzazione del *client*.

Piattaforma virtualizzata "as-a-Service"

La modalità di remunerazione per la fornitura "as-a-Service" di una piattaforma application server virtualizzata, destinata ad ospitare tutte le componenti *client* acquisite dall'Amministrazione attraverso questo servizio, è a "canone annuale".

Ai fini della valutazione economica è richiesta una quotazione [€/anno] per il canone annuale della piattaforma virtualizzata.

1.1.4. Servizio L3.S4 - Orchestrazione

1.1.4.1. Descrizione del Servizio

Il servizio prevede la messa a disposizione di un sistema che realizzi le funzionalità di "web services orchestration", ossia la composizione di servizi di Cooperazione Applicativa, già esistenti e pubblicati nel registro SICA, ai fini della realizzazione di un accordo di cooperazione (**orchestrazione esterna**) o la composizione di web services già presenti nel dominio dell'Amministrazione al fine di realizzare un servizio di cooperazione applicativa (**orchestrazione interna**).

Orchestrazione esterna

L'orchestrazione esterna è finalizzata alla composizione di più servizi di cooperazione applicativa, già esistenti e pubblicati nel registro SICA, erogati da diverse Amministrazioni.

Ai sensi del DPCM 1 Aprile 2008 - Regole tecniche di SPC -, le Amministrazioni erogatrici dei singoli servizi costituiranno un "dominio di cooperazione" e il servizio composto sarà definito da un "accordo di cooperazione" pubblicato nel registro SICA. Una delle Amministrazioni costituente il dominio di cooperazione si assumerà l'onere di orchestrare i suddetti servizi di cooperazione attraverso il servizio di orchestrazione oggetto della



fornitura e di rendere disponibile il *web services* costituente il servizio composto - risultato dell'orchestrazione - attraverso la propria porta di dominio.

Orchestrazione interna

L'orchestrazione interna è finalizzata all'integrazione di sistemi informativi interni al dominio dell'Amministrazione e dotati di interfacce *web services* (native o realizzate tramite il servizio "Realizzazione interfacce *web services*"), allo scopo di realizzare un servizio composto che realizzi un procedimento amministrativo avente visibilità esterna del quale i *web services* orchestrati costituiscano un sottoinsieme dei passi previsti. I *web services* risultato della composizione dovranno costituire un servizio di Cooperazione Applicativa reso fruibile attraverso la porta di dominio, con accordo di servizio pubblicato nel registro SICA.

1.1.4.2. Requisiti funzionali

Nell'ambito del servizio "Orchestrazione", il Fornitore deve garantire almeno i seguenti requisiti funzionali:

- stesura del BPEL di orchestrazione;
- in caso di erogazione di una piattaforma di orchestrazione in modalità "as-a-Service", dispiegamento dell'orchestrazione realizzata sulla piattaforma di orchestrazione;
- in caso di orchestrazione interna, predisposizione dell'accordo di servizio;
- in caso di orchestrazione esterna, predisposizione dell'accordo di cooperazione;
- supporto all'Amministrazione nella pubblicazione dell'accordo di servizio o dell'accordo di cooperazione di cui al punto precedente;
- gestione dell'esercizio, della manutenzione correttiva (MAC) ed evolutiva (MEV) dell'orchestrazione realizzata;
- fornitura all'Amministrazione di una piattaforma di orchestrazione erogata in modalità cloud, se non già disponibile o perché messa a disposizione dall'Amministrazione o già precedentemente acquisita attraverso sottoservizio di orchestrazione interna.

1.1.4.3. Requisiti tecnici e tecnologici

Nell'ambito del servizio "Orchestrazione", il Fornitore deve garantire almeno i seguenti requisiti tecnici e tecnologici:

- In caso di erogazione di una piattaforma di orchestrazione in modalità "as-a-Service", tale piattaforma dovrà supportare i seguenti standard:
 - OASIS WS-BPEL 2.0



- W3C WSDL 1.1
- In caso di erogazione di una piattaforma di orchestrazione in modalità “*as-a-Service*”, tale piattaforma dovrà essere opportunamente dimensionata e quindi avere una capacità elaborativa adeguata alle necessità e alla numerosità dei servizi orchestrati. Tale piattaforma dovrà essere mantenuta e aggiornata dal Fornitore;
- In caso di erogazione di una piattaforma di orchestrazione in modalità “*as-a-Service*”, tale piattaforma deve essere in grado di poter elaborare in parallelo almeno 10 (dieci) richieste ogni secondo nel 95% dei casi, e almeno 8 (otto) richieste ogni secondo nel 100% dei casi;
- In caso di erogazione di una piattaforma di orchestrazione in modalità “*as-a-Service*”, tale piattaforma deve essere in grado di poter mantenere in coda almeno 10 (dieci) richieste ogni secondo nel 95% dei casi, e almeno 8 (otto) richieste ogni secondo nel 100% dei casi.

1.1.4.4. Tipologia del Servizio

Il servizio sarà erogato in modalità progettuale “*on-premise*”. In caso di richiesta da parte dell’Amministrazione di una piattaforma di orchestrazione, tale piattaforma sarà erogata dal Fornitore presso il proprio Centro Servizi in modalità “*as-a-Service*”.

1.1.4.5. Parametri di valutazione economica

Il costo del servizio “Orchestrazione” è determinato dalle seguenti voci di costo:

Realizzazione orchestrazione

La Modalità di remunerazione delle attività di realizzazione dell’orchestrazione dei servizi è “a corpo” sulla base del numero di orchestrazioni interne o esterne. Ciascuna orchestrazione deve portare alla realizzazione di un servizio di cooperazione applicativa, sia esso relativo ad un accordo di servizio, nel caso di orchestrazione interna, o a un accordo di cooperazione, nel caso di orchestrazione esterna.

Ai fini della valutazione economica delle attività di orchestrazione è richiesta una quotazione [€/orchestrazione] per i seguenti elementi:

- Singola orchestrazione per una orchestrazione complessiva di meno di 10 (dieci) servizi;
- Singola orchestrazione per una orchestrazione complessiva di 10 (dieci) o più servizi.

Il servizio di “Orchestrazione” è comprensivo di 12 (dodici) mesi di garanzia per i *client* realizzati.



Manutenzione *orchestrazione*

La modalità di remunerazione per le attività di gestione dell'esercizio, della manutenzione correttiva (MAC) ed evolutiva (MEV) delle orchestrazioni realizzate con il servizio in oggetto è a "canone annuale" in base al numero di servizi orchestrati.

Tale canone decorre dal termine della garanzia di 12 (dodici) mesi compresa nella realizzazione delle orchestrazioni, con durata sino alla scadenza del Contratto Quadro.

Ai fini della valutazione economica è richiesta una quotazione [€/servizio orchestrato] per il canone annuale di manutenzione di un singolo servizio orchestrato nell'ambito del servizio.

Piattaforma virtualizzata "*as-a-Service*"

La modalità di remunerazione per la fornitura "*as-a-Service*" di una piattaforma application server virtualizzata, destinata ad ospitare tutte le orchestrazioni acquisite dall'Amministrazione attraverso questo servizio, è a "canone annuale".

Ai fini della valutazione economica è richiesta una quotazione [€/anno] per il canone annuale della piattaforma virtualizzata.



1.2. Tipologia Servizi L3 - Open Data

Nell'ambito dell'Agenda Digitale Europea il pilastro sul mercato unico digitale prevede all'azione 3 di rivedere la direttiva 2003/98/CE del Parlamento europeo e del Consiglio per introdurre norme minime necessarie a garantire la massima apertura e riutilizzo dell'informazione del settore pubblico nell'Unione Europea. Dando seguito a tale azione, la recente Direttiva apre quindi ai nuovi paradigmi dell'Open Data, ribadendo la possibilità di generare valore economico dai dati pubblici rilasciati dalle pubbliche amministrazioni seguendo principi di interoperabilità, qualità, riuso e condivisione.

A livello nazionale, con la definizione dell'Agenda Digitale, l'Italia ha introdotto alcune disposizioni, all'art. 52 del Codice dell'Amministrazione Digitale, relative alle modalità di gestione e accesso ai dati pubblici, prevedendo il principio dell' "Open Data by default" per i dati già pubblicati e politiche di valorizzazione del patrimonio pubblico per facilitare il rilascio di quelli già prodotti e da produrre secondo i criteri dell'Open Data (art. 68 del Codice dell'Amministrazioni Digitale).

In questo contesto, si definisce dato di tipo aperto, un dato pubblico, ovvero un dato della Pubblica Amministrazione conoscibile da chiunque e non soggetto a restrizioni temporali (e.g., diritto all'oblio), che:

- è disponibile secondo i termini di una licenza che ne permetta l'utilizzo da parte di chiunque, anche per finalità commerciali;
- è disponibile nella forma più disaggregata possibile;
- è accessibile mediante tecnologie ICT in formato di tipo aperto (secondo quanto indicato dalle Linee guida sulla valorizzazione del patrimonio informativo pubblico pubblicate dall'Agenzia per l'Italia Digitale);
- è provvisto dei relativi metadati;
- è disponibile gratuitamente attraverso le tecnologie ICT oppure è reso disponibile ai costi marginali sostenuti per la sua riproduzione e divulgazione.

Sono esclusi quindi tutti i dati a conoscibilità limitata, ad esempio i dati:

- coperti da segreto di stato;
- coperti da segreto statistico;
- coperti dal diritto d'autore;
- per i quali trova applicazione la normativa vigente sulla protezione dei dati personali (i.e., Dlgs n. 196/2003 e Linee guida in materia di trattamento di dati personali rilasciate dal Garante Privacy).



1.2.1. Servizio L3.S5 - Open Data

1.2.1.1. Descrizione del Servizio

Il servizio Open Data è pensato per essere di ausilio alle Amministrazioni per l'attuazione delle disposizioni degli art. 52 e 68 del Codice per l'Amministrazione Digitale e delle disposizioni contenute all'art. 7 del Dlgs. 33/2013 in materia di pubblicità e trasparenza della Pubblica Amministrazione ma, in quest'ultimo caso, limitatamente a quanto indicato dalle Linee guida del Garante e dalle Linee guida per la valorizzazione del patrimonio informativo pubblico (anno 2014)⁴ dell'Agenzia per l'Italia Digitale.

Il servizio consente alle Pubbliche Amministrazioni di:

- produrre dati di tipo aperto con i relativi metadati;
- produrre documenti di tipo aperto con i relativi metadati;
- pubblicare dati di tipo aperto e i relativi metadati in un portale

a partire da documentazione e basi di dati già presenti presso le Amministrazioni.

In particolare, il servizio consente di identificare, analizzare, bonificare, trasformare dati pubblici in un formato di tipo aperto non proprietario, metadattare i dati, identificare e associare ai dati una licenza aperta per il loro riutilizzo, anche per finalità commerciali, e pubblicare i dati in un portale.

Il prodotto di tale servizio sarà principalmente la produzione, la modellazione e la pubblicazione di dati che appartengono solo ai seguenti livelli del modello per i dati aperti proposto dall'Agenzia per l'Italia Digitale, e pubblicato nell'Agenda nazionale per la valorizzazione del patrimonio informativo pubblico: livello tre, livello quattro, livello cinque.

Il servizio consiste nella gestione di tutte le fasi necessarie alla produzione, modellazione e pubblicazione di dati di tipo aperto secondo le modalità previste dalle linee guida per la valorizzazione del patrimonio informativo pubblico rilasciate dall'Agenzia per l'Italia Digitale e dalle linee guida per "l'interoperabilità semantica attraverso i Linked Open Data" pubblicate nell'ambito del Sistema Pubblico di Connettività (SPC).

⁴ Il link web al documento è:
http://www.agid.gov.it/sites/default/files/linee_guida/patrimoniopubblicolg2014_v0.7finale.pdf



Il servizio in oggetto è di natura progettuale ed è composto dalle seguenti cinque Fasi di processo:

- **Fase 1** - Censimento, analisi e bonifica dei dati presenti presso l'Amministrazione;
- **Fase 2** - Produzione e metadatazione di dati a livello 3;
- **Fase 3** - Produzione e metadatazione di dati di livello 4 e/o produzione di Linked Open Data (LOD) o dati di livello 5;
- **Fase 4** - Pubblicazione dei dataset realizzati;
- **Fase 5** - Aggiornamento e conservazione dei dataset prodotti.

Si precisa che per dataset si intende una collezione di dati tipicamente omogenei per contenuto gestite da un soggetto pubblico nell'ambito delle proprie attività istituzionali al quale è associato una licenza aperta che eventualmente ne possa consentire anche il riutilizzo a fini commerciali.

L'individuazione dei dataset da produrre deve essere fornita dall'Amministrazione di concerto con il Fornitore; in ogni caso, aggiornamenti degli stessi dati contenuti nella collezione non costituiscono un nuovo dataset.

Si precisa che l'Amministrazione può acquisire l'intero percorso oppure può selezionare le fasi a seconda delle specifiche esigenze, anche dipendenti dalle tipologie di dati da rendere disponibili come Open Data, ma tenendo in considerazione i seguenti vincoli di acquisizione e di erogazione del servizio:

- Le Fasi 2, 3 e 4, per essere acquisite, devono necessariamente essere precedute dalla Fase 1;
- La Fase 4 deve essere preceduta dalle Fasi 2 o 3, nonché dalla Fase 1.

Fase 1. Censimento, analisi e bonifica dei dati

Nel caso in cui l'Amministrazione abbia necessità di produrre dati di tipo aperto a partire da documentazione e basi di dati già presenti presso l'Amministrazione, questa Fase prevede le seguenti attività in carico al Fornitore:

- a) affiancamento e supporto nel censimento dei dati pubblici esistenti al responsabile Open Data dell'Amministrazione, al responsabile trasparenza e ai responsabili delle basi di dati (o laddove non previsti tali responsabili, affiancamento e supporto ai Referenti indicati dall'Amministrazione);
- b) per ciascuna base di dati identificata al precedente punto a), affiancamento e supporto all'Amministrazione nello svolgere un'analisi delle fonti individuate per determinare i diritti ed eventuali vincoli di pubblicazione. Tale analisi è condotta utilizzando almeno la "checklist" inclusa nelle Linee guida sulla valorizzazione del patrimonio informativo pubblico (anno 2014);



- c) affiancamento e supporto all'Amministrazione nello svolgere un'analisi della qualità dei dati, una volta eseguiti i punti a) e b), in cui evidenziare, tra gli altri, eventuali problemi di inconsistenza, di ambiguità semantica, di dati da bonificare, di dati mancanti e/o incoerenti.

Il Fornitore dovrà indicare nell'offerta le metodologie di erogazione del servizio, con particolare riferimento alla data quality e alle relative dimensioni di analisi della qualità del dato che intende utilizzare per assicurare che i dati possano essere effettivamente utilizzabili. Per le dimensioni di analisi della qualità del dato dovranno essere indicate le misure e le soglie che consentano di discriminare la bontà di un dato rispetto alla dimensione in esame. Tali misure e soglie dovranno essere in ogni caso concordate con l'Amministrazione in caso di attivazione del servizio.

In particolare, tenuto anche conto dello standard di riferimento per il modello sulla qualità del dato ISO/IEC 25012, almeno le seguenti dimensioni devono essere soddisfatte per rendere il dato fruibile: completezza, accuratezza, coerenza, originalità, attualità (o tempestività dell'aggiornamento).

Come risultato delle azioni di cui ai punti a) b) e c), il Fornitore supporta l'Amministrazione nella redazione di un documento di sintesi nel quale indicare almeno le seguenti informazioni:

- la lista di tutte le basi di dati individuate nel censimento;
- per ciascuna base di dati individuata, le caratteristiche descrittive dei dati oggetto del processo di apertura;
- le caratteristiche peculiari dei dati oggetto del processo di apertura;
- il tasso temporale di aggiornamento dei dati oggetto del processo di apertura;
- eventuali limiti tecnici e giuridici identificati per la pubblicazione;
- l'assessment della qualità dei dati riscontrata.

Sulla base del documento, e sulla base del livello di produzione dei dati di tipo aperto (livello tre o quattro/cinque) che l'Amministrazione desidera acquisire, la Fase 1 consente di svolgere, previa approvazione dell'Amministrazione, le necessarie azioni correttive sui dati.

Il Fornitore potrà supportare l'Amministrazione nell'applicare le azioni correttive scegliendo tra una o più delle seguenti modalità, indicate a titolo esemplificativo, e comunque considerando il livello di produzione (Fase 2 o Fase 3) del dataset che l'Amministrazione desidera acquisire:

- affiancamento ai soggetti responsabili dei dati per una correzione congiunta;
- confronti incrociati con altri dati (matching);



- analisi e revisione dei processi di produzione dei dati che hanno causato la scarsa qualità.

Fase 2. Produzione e metadattazione di dati a livello 3

Questa fase segue la Fase 1 precedente. Le attività previste sono le seguenti:

- a) **produzione di dataset** leggibili da umani e da agenti automatici (*human e machine-readable*), strutturati e disponibili in un formato non proprietario o in un formato definito da specifiche aperte (i.e., dataset di livello tre del modello dei dati proposto dall'Agenzia per l'Italia Digitale).

Sono previste nel servizio le seguenti tre classi di formati di tipo aperto per la produzione dei dataset:

- **Classe A** - formato CSV;
- **Classe B** - formati XML e JSON;
- **Classe C** - formati di dati geospaziali shapefile, GeoJSON, GML, KML 2.2

Infine, nel caso di documenti, questa fase consente di produrre documenti aperti secondo le raccomandazioni e i formati individuati dalle linee guida per la valorizzazione del patrimonio informativo pubblico dell'Agenzia per l'Italia Digitale;

- b) **metadattazione dei dataset prodotti**. Sono richieste le attività di metadattazione dei dataset:

- **Metadattazione per formati di Classe A (formato CSV) e Classe B (formati XML e JSON)**. I dataset di questo formato devono essere arricchiti con almeno l'insieme di metadati obbligatori e obbligatori condizionatamente previsti dalle Linee guida nazionali per la valorizzazione del patrimonio informativo pubblico;
- **Metadattazione per formati di Classe C (formati di dati geospaziali shapefile, GeoJSON, GML, KML 2.2)**. I dataset geospaziali devono essere arricchiti con l'insieme dei metadati previsti nel contesto del Repertorio Nazionale dei Dati Territoriali (RNDT) come stabilito dalle regole tecniche del Repertorio e le relative guide tecniche; tali metadati saranno inseriti nel RNDT dall'Amministrazione con il supporto del Fornitore.

I dataset prodotti al precedente punto a), ove necessario per meglio documentare il contenuto dei dataset stessi, possono essere arricchiti con uno schema che descriva il dominio di riferimento e gli elementi del dominio;

I dataset prodotti al precedente punto a) possono essere accompagnati, qualora disponibili, dalle informazioni di provenienza, riguardanti entità, attività e



persone che hanno contribuito alla formazione del dataset, seguendo almeno le raccomandazioni W3C del PROV framework⁵;

- c) **selezione per ciascun dataset prodotto al punto a) di una licenza aperta** che ne consenta il riutilizzo anche per finalità commerciali seguendo le raccomandazioni incluse nelle Linee guida per la valorizzazione del patrimonio informativo pubblico.

Il Fornitore dovrà indicare obbligatoriamente nell'offerta l'intero processo di produzione e metadattazione dei dati e/o documenti, evidenziando gli strumenti utilizzati a tale scopo.

Fase 3. Produzione di dati di livello 4 e 5

Questa Fase si suddivide in due sotto-fasi:

- **Sotto-fase 3a** - Produzione di dati di livello 4;
- **Sotto-fase 3b** - Produzione di Linked Open Data (LOD) o dati di livello 5.

Si noti che la sotto-fase 3b può essere abilitata su richiesta dell'Amministrazione e segue la sotto-fase 3a.

La Sotto-fase 3a di produzione di dati di livello 4 avviene nel rispetto delle raccomandazioni incluse nelle linee guida su "interoperabilità semantica attraverso Linked Open Data"⁶ prodotte dalla Commissione di Coordinamento SPC e pubblicate dall'Agenzia per l'Italia Digitale. Essa consiste nella seguenti attività:

- partendo dall'analisi effettuata alla "Fase 1. Censimento, analisi e bonifica dei dati", definizione di un'ontologia attraverso l'uso di RDFS o OWL che specifichi i concetti ("classi"), le relazioni tra i concetti ("proprietà"), i vincoli di cardinalità ("restrizioni") oltre che i commenti e le annotazioni e ogni altro elemento per la rappresentazione del dominio di riferimento.

Nella definizione dell'ontologia, il Fornitore dovrà indicare dettagliatamente come e quali ontologie già esistenti e utilizzate a livello internazionale possono essere riutilizzate per la modellazione di parte o dell'intero dominio di riferimento. In tale attività, il Fornitore, ove possibile, deve riferirsi alla lista di ontologie proposte nelle Linee guida per la valorizzazione del patrimonio informativo pubblico L'Amministrazione ha comunque la piena facoltà di scelta sulla modellazione del dominio, e quindi sulle ontologie da sviluppare o riutilizzare;

⁵ <http://www.w3.org/TR/prov-o/>

⁶ http://www.agid.gov.it/sites/default/files/documentazione_trasparenza/cdc-spc-gdl6-interoperabilitasemopendata_v2.0_0.pdf



- produzione dei dati in RDF, serializzati in almeno uno tra N3, N-Triple, Turtle, JSON-LD, RDF/XML ovvero la produzione di proposizioni (“triple”) nella forma <oggetto> <predicato> <oggetto> dove ogni elemento è identificato in maniera univoca da un URI deferenziabile e persistente;
Il Fornitore dovrà indicare nell’offerta l’intero processo di trasformazione dei dati in RDF e gli strumenti utilizzati a tale scopo;
- arricchimento dei dati secondo quanto indicato dalle linee guida per “l’interoperabilità semantica attraverso Linked Open Data” e comunque, per la metadateazione, assicurando la presenza dei metadati obbligatori e obbligatori condizionatamente, in conformità con quanto indicato dalle Linee guida nazionali per la valorizzazione del patrimonio informativo pubblico.

La Sotto-fase 3b di produzione di Linked Open Data o dati di livello 5 consiste nelle seguenti attività:

- identificazione di altri dataset da collegare al dataset prodotto nella Sotto-fase 3a. I dataset identificati possono essere interni ovvero prodotti dall’Amministrazione stessa, e dataset esterni ovvero già presenti nel Web dei dati;
- creazione delle triple di collegamento esterno (*interlinking*), da aggiungere al dataset prodotto nella Sotto-fase 3a, anche mediante l’uso di strumenti di record linkage.

Tali Sotto-fasi 3a e 3b sono complementate dalla selezione, per ciascun dataset di livello 4 e/o 5, di una licenza aperta che ne consenta il riutilizzo anche per finalità commerciali seguendo le raccomandazioni incluse nelle Linee guida per la valorizzazione del patrimonio informativo pubblico. I metadati necessari per identificare chiaramente le condizioni di riutilizzo dei dati costituiscono parte integrante dell’arricchimento dei dati previsto per il servizio.

Non è prevista la mera esportazione automatica del dataset in sola e semplice sintassi RDF.

Fase 4. Pubblicazione dei dati

La pubblicazione dei dati consiste delle seguenti attività:

- a) su indicazione dell’Amministrazione, pubblicazione di dataset di tipo aperto in un portale/piattaforma esterna disponibile o come estensione di un CMS esistente;
- b) messa a disposizione di meccanismi per l’interrogazione dei dati nelle modalità espresse nel seguito.



L'attività di cui al punto a) consiste nelle seguenti attività specifiche:

- realizzazione di una sezione specifica o di singole pagine Web per l'Open Data da integrarsi all'interno di un portale indicato dall'Amministrazione (e.g. un portale Open Data esistente o il sito Web istituzionale); a tal fine, quest'ultima dovrà fornire al Fornitore gli elementi, anche grafici, che consentano di abilitare tale integrazione.

La sezione specifica o le singole pagine Web devono essere predisposte per pubblicare i dataset prodotti a seguito dell'acquisizione delle Fasi 2 e/o 3.

La realizzazione deve essere fatta rispettando requisiti di accessibilità e deve consentire agli utenti che vi accedono di effettuare almeno le seguenti attività, compatibilmente con il portale in cui tali pagine o sezione si integrano:

- scaricare i dataset;
- votare i dataset;
- commentare i dataset;
- ottenere una preview dei dataset;
- condividere/segnalare i dataset su social network;
- accedere alle applicazioni sviluppate a partire dai dataset;
- interrogare i dataset secondo le modalità indicate nel seguito.

Per ogni dataset, le pagine Web o la sezione dovranno necessariamente contenere le informazioni sui metadati, in conformità con quanto indicato dalle Linee guida sulla valorizzazione del patrimonio informativo pubblico, la licenza associata e i formati disponibili;

- predisposizione di quanto realizzato al punto precedente affinché si possa integrare con le funzioni, ove presenti, di:
 - ricerca - sia semplice che avanzata, sulla base di motori semantici;
 - navigazione dati;
 - visualizzazione dati;
 - federazione con il portale nazionale dati.gov.it (la predisposizione deve essere svolta su specifica richiesta dell'Amministrazione o di Consip/AgID).

L'attività di cui al punto b) consiste nelle seguenti attività specifiche, distinte in base al tipo di livello del dataset da interrogare:

- nel caso di dataset di tipo aperto di livello tre, l'attività di interrogazione dei dati, abilitata su richiesta dell'Amministrazione, prevede l'esposizione di interfacce API realizzate dal Fornitore per l'accesso diretto ai dati di tipo aperto;



- nel caso di dataset di tipo aperto di livello quattro e cinque (Linked Open Data), la funzionalità di interrogazione dei dati prevede l'esposizione di uno SPARQL *end-point*. Il Fornitore dovrà inoltre implementare opportuni meccanismi per una corretta risoluzione degli URI.

Il Fornitore dovrà dettagliatamente indicare nell'offerta tecnica tutte le fasi del processo di pubblicazione e gli strumenti utilizzati evidenziando la conformità delle attività alle raccomandazioni incluse nelle Linee guida nazionali per la valorizzazione del patrimonio informativo pubblico.

Fase 5. Aggiornamento e conservazione

Questa fase consiste delle seguenti attività:

- aggiornamento dei dataset disponibili sotto forma di open data sulla base del tasso di aggiornamento richiesto dall'Amministrazione. Tale attività deve essere garantita per tutta la durata del contratto esecutivo;
- conservazione delle serie storiche dei dataset prodotti avendo cura di predisporre un opportuno sistema di archiviazione nel quale mantenere almeno le informazioni sulla versione del dataset, sul momento di creazione del dataset, sul momento di pubblicazione e sull'intervallo temporale al quale il dataset si riferisce.

1.2.1.2. Requisiti funzionali

Nell'ambito del servizio "Open Data", il Fornitore deve garantire almeno i seguenti requisiti funzionali, oltre a quelli già indicati all'interno del paragrafo §1.2.1.1 "Descrizione del Servizio":

- l'analisi delle fonti dati prevista nella Fase 1 "Censimento, analisi e bonifica dei dati" deve essere condotta utilizzando almeno la "checklist" inclusa nelle Linee guida sulla valorizzazione del patrimonio informativo pubblico (anno 2014);
- l'analisi sulla qualità dei dati prevista nella Fase 1 "Censimento, analisi e bonifica dei dati" deve prevedere almeno la valutazione di problemi di inconsistenza nei dati, di significati ambigui, di dati "sporchi", di dati mancanti e/o incoerenti.



1.2.1.3. Requisiti tecnici e tecnologici

Nell'ambito del servizio "Open Data", il Fornitore deve garantire almeno i seguenti requisiti tecnici e tecnologici, oltre a quelli già indicati all'interno del paragrafo §1.2.1.1 "Descrizione del Servizio":

- le classi di formato di tipo aperto previste per la produzione di dataset di livello 3 sono le seguenti:
 - CSV, XML e JSON;
 - Formati di dati geospaziali shapefile, GeoJSON, GML, KML 2.2.
- i dataset prodotti nei formati CSV, XML e JSON tramite la Fase 2 devono essere arricchiti con almeno l'insieme di metadati obbligatori e obbligatori condizionatamente previsti dalle Linee guida nazionali per la valorizzazione del patrimonio informativo pubblico;
- i dataset prodotti nei formati geospaziali tramite la Fase 2 devono essere arricchiti con l'insieme dei metadati previsti nel contesto del Repertorio Nazionale dei Dati Territoriali (RNDDT) come stabilito dalle regole tecniche del Repertorio e le relative guide tecniche;
- devono essere seguite almeno le raccomandazioni W3C del PROV framework nel caso in cui i dataset prodotti nella Fase 2 possano essere accompagnati, qualora disponibili, dalle informazioni di provenienza, riguardanti entità, attività e persone che hanno contribuito alla formazione del dataset;
- la definizione di una ontologia nella Fase 3a deve essere svolta attraverso l'uso di RDFS o OWL;
- i dati prodotti in RDF devono essere serializzati in almeno uno tra N-3, N-Triple, Turtle, JSON-LD, RDF/XML.

1.2.1.4. Tipologia del Servizio

Il servizio dovrà essere erogato in modalità progettuale "on-premise". Il servizio prevede opzionalmente l'erogazione in modalità "as-a-Service", presso il Centro Servizi del Fornitore, di uno SPAQRL *end-point* per la pubblicazione di dati di livello 4 e/o 5.

1.2.1.5. Parametri di valutazione economica

Ai fini della valutazione economica, i relativi parametri sono di seguito esplicitati in base alle cinque Fasi di processo previste per il servizio.

Il costo OD del servizio di "Open Data" è determinato dalla seguente formula:

$$OD = C_{fase1} + \sum_{i=1}^{N_{dataset}} (C_{fase2_i} + C_{fase3_i} + C_{fase4_i}) + C_{fase5}$$



Di seguito l'evidenza della modalità di remunerazione per ogni voce di costo presentata nella formula:

C_{fase1}: Costo Fase 1. Analisi, censimento e bonifica dei dati

La Modalità di remunerazione della Fase 1 è a “corpo [gg/p]” in base ai giorni persona concordati ed erogati dal Fornitore.

Ai fini della valutazione economica, il Fornitore dovrà indicare la quotazione espressa in [€/giorno] per ognuna delle figure professionali, la cui descrizione è presente in Appendice 1, relative al mix definito nella seguente tabella:

| Figura professionale | Impiego |
|-------------------------|---------|
| Capo Progetto | 10% |
| Data Architect | 20% |
| Specialista di tematica | 30% |
| Sviluppatore | 40% |

C_{fase2}: Costo Fase 2. Produzione e metadattazione di dati a livello 3

La modalità di remunerazione della Fase 2 è “a corpo” sulla base del numero e della classe di dataset prodotti e metadattati di livello tre.

Ai fini della valutazione economica è richiesta una quotazione [€/dataset] per ciascuno dei seguenti elementi:

- Produzione e metadattazione singolo dataset di Classe A (formato CSV)
- Produzione e metadattazione singolo dataset di Classe B (formati XML e JSON)
- Produzione e metadattazione singolo dataset di Classe C (formati di dati geospaziali shapefile, GeoJSON, GML, KML 2.2)

C_{fase3}: Costo Fase 3. Produzione di dati di livello 4 e 5

La Modalità di remunerazione della Fase 3 è a “corpo” sulla base del numero di dataset realizzati di livello 4 e di livello 5.

La seguente formula è utilizzata per il calcolo del costo della Fase 3:

$$C_{fase3} = \text{CorrLivello4} + \text{CorrLivello5}$$

dove,

$$\text{CorrLivello4} = \text{CorrmodelOnt} + \text{PrRDF}$$

$$\text{CorrLivello5} = \text{\#fonti} * \text{PrLink}$$

Di seguito l'evidenza della modalità di remunerazione per ogni voce di costo presentata nella formula relativa alla Fase 3:

“CorrmodelOnt” - Modellazione dell'ontologia

La Modalità di remunerazione dell'attività di modellazione dell'ontologia è a “corpo [gg/p]” in base ai giorni persona concordati ed erogati dal Fornitore.



Ai fini della valutazione economica, il Fornitore dovrà indicare la quotazione espressa in [€/giorno] per ognuna delle figure professionali, la cui descrizione è presente in Appendice 1, relative al mix definito nella seguente tabella:

| Figura professionale | Impiego |
|-------------------------|---------|
| Capo Progetto | 5% |
| Data Architect | 25% |
| Specialista di tematica | 15% |
| Data scientist | 25% |
| Sviluppatore | 30% |

“PrRDF” - Produzione RDF

La modalità di remunerazione è “a corpo” in base al numero di RDF prodotti (con produzione delle triple relative all’arricchimento).

Ai fini della valutazione economica è richiesta una quotazione [€/RDF] per le attività di produzione di un singolo RDF.

“Prlink” - Interlinking del dataset con una singola fonte

La modalità di remunerazione è “a corpo” in base al numero di fonti (#fonti) diversi ai quali collegare i dataset da produrre sottoforma di Linked Open Data; comprende sia dataset interni all’Amministrazione, sia esterni disponibili nel Web dei dati.

Ai fini della valutazione economica è richiesta una quotazione [€/interlinking] per le attività di collegamento del dataset a una singola fonte interna/esterna (*interlinking*).

Fase 4. Pubblicazione dei dati

In base alla modalità di erogazione scelta per questa Fase, le seguenti formule sono utilizzate per il calcolo del costo della Fase 4:

- “as-a-Service”: $C_{fase3} = Corrpub + CorrSPARQL_MV$
- “on-premise”: $C_{fase3} = Corrpub + CorrSPARQL$

Ai fini della valutazione economica è richiesta una quotazione per ciascuno dei seguenti parametri:

“Corrpub” - Produzione sezioni/pagine web o API

Comprende i costi per la realizzazione della sezione o delle singole pagine Web e i costi per la realizzazione e l’esercizio di API che consentano un’interrogazione diretta ai dati pubblicati di livello 3 qualora l’Amministrazione Committente abbia richiesto tale funzionalità.

La Modalità di remunerazione è a “corpo [gg/p]” in base ai giorni persona concordati ed erogati dal Fornitore.



Ai fini della valutazione economica, il Fornitore dovrà indicare la quotazione espressa in [€/giorno] per ognuna delle figure professionali, la cui descrizione è presente in Appendice 1, relative al mix definito nella seguente tabella:

| Figura professionale | Impiego |
|----------------------|---------|
| Capo progetto | 10% |
| Analista Funzionale | 30% |
| Sviluppatore | 60% |

“CorrSPARQL_MV” - Esposizione SPARQL end-point as-a-Service

Canone per l’esposizione di uno SPARQL *end-point* presso il centro servizi del Fornitore comprensivo dei costi infrastrutturali (macchine virtuali) e di esercizio necessari per la sua erogazione. La modalità di remunerazione è a “canone (su base annuale)”.

Ai fini della valutazione economica dovrà essere presentata una quotazione [€/anno] per il canone annuale dello SPARQL *end-point*.

“CorrSPARQL” - Esposizione SPARQL end-point on-premise

Canone annuale che include i costi di installazione e configurazione di uno SPARQL *end-point* presso l’Amministrazione e i costi per la necessaria assistenza tecnica.

Ai fini della valutazione economica dovrà essere presentata una quotazione [€/anno] per il canone annuale dello SPARQL *end-point* installato “*on-premise*”.

Fase 5. Aggiornamento e conservazione

La modalità di remunerazione della Fase 5 è “a canone (su base annuale)” per le attività di aggiornamento e conservazione dei dati.

Tale canone non è soggetto a valutazione economica e il costo annuale è automaticamente calcolato nella misura del 5% rispetto all’importo totale consuntivato per il servizio Open Data ad esclusione della Fase 1.



1.3. Tipologia Servizi L3 - Big Data

Di giorno in giorno cresce rapidamente la produzione digitale di informazioni riguardanti ogni sorta di attività umana, da dati non strutturati a dati amministrativi di dettaglio. L'entità di questo fenomeno aumenta esponenzialmente senza presentare cenni di arresto, anche per via dell'uso pervasivo di applicazioni legate al Web come, ad esempio, i *Social Network*, i *Social Media*, l'*Internet of Things*, le *Smart City*, ecc.

In tali contesti risulta di fondamentale importanza la possibilità di elaborare e correlare questa enorme mole di dati al fine di estrarne valore, con l'obiettivo finale di supportare efficacemente i processi amministrativi e di business delle Amministrazioni Pubbliche. Le classiche metodologie e tecnologie di memorizzazione, gestione e analisi dei dati possono a tale scopo rivelarsi inefficaci, non potendo assicurare sempre performance adeguate. Per far fronte a queste nuove necessità, si stanno quindi affermando negli ultimi anni metodologie e tecnologie innovative riferite col termine Big Data.

Al fine di consentire alle Amministrazioni di usufruire di queste innovative metodologie e tecnologie di Big Data, nell'ambito della presente fornitura sono resi disponibili i seguenti due servizi, suddivisi in base alla finalità di utilizzo:

- **L3.S6 - Servizio di supporto alla memorizzazione dei Big Data.** Tale servizio consente alle Amministrazioni di usufruire di supporto da parte del Fornitore per la memorizzazione dei Big Data, siano essi dati interni e/o esterni all'Amministrazione, al fine di poter gestire le tipiche complessità di persistenza e scalabilità dei sistemi di gestione dei Big Data;
- **L3.S7 - Servizio di supporto all'analisi dei Big Data.** Questo servizio di analisi dei Big Data consente all'Amministrazione di usufruire di un supporto del Fornitore per elaborare e correlare dati interni e/o esterni tramite appositi modelli e strumenti di analisi dei Big Data, per individuare trend, pattern nascosti e nuove correlazioni tra i dati, al fine di abilitare l'Amministrazione stessa nei suoi processi interni decisionali e amministrativi.

L'Amministrazione Committente potrà acquisire uno o entrambi i servizi secondo le proprie necessità.

È esclusa dai servizi Big Data la fornitura dei seguenti sistemi e software:

- sistemi di business intelligence;
 - sistemi di basi di dati relazionali;
 - sistemi di *data warehouse* configurati con sistemi relazionali;
 - sistemi di basi di dati NewSQL (o sistemi relazionali di nuova generazione).
- Questi sistemi consentono di gestire basi di dati strutturate relazionali e



distribuite in modalità “*shared nothing*” supportando le proprietà transazionali ACID;

- sistemi di basi di dati a oggetti;
- sistemi di basi di dati XML;
- sistemi di *triple* o *quad store*. Questi sistemi memorizzano i dati strutturati sotto forma di triple (unità informative formate da tre componenti detti “soggetto”, “predicato” e “oggetto”). Generalmente questa tipologia di sistema è utilizzata per memorizzare dati nel formato RDF.

1.3.1. Servizio L3.S6 - Supporto alla memorizzazione dei Big Data

1.3.1.1. Descrizione del Servizio

Il servizio in oggetto è articolato nei seguenti sotto-servizi:

- **Sotto-servizio base “M”** - Attivazione di sistemi di supporto alla memorizzazione dei Big Data, sia in modalità “*as-a-Service*” che “*on-premise*”.
- **Sotto-servizi accessori e opzionali, acquisibili a discrezione dell’Amministrazione:**
 - “A” - Assessment dello scenario applicativo per l’individuazione dei sistemi di gestione dei Big Data da acquisire;
 - “E” - Conduzione dei sistemi software di gestione dei Big Data acquisiti e installati “*on-premise*” tramite il sotto-servizio base “M”;
 - “O” - Attività di “*configurazione avanzata*” dei sistemi di supporto alla memorizzazione attivati tramite il sotto-servizio base “M”, sia in modalità “*as-a-Service*” che “*on-premise*”.

L’acquisizione dei sotto-servizi accessori e opzionali è vincolata all’acquisizione del sotto-servizio base.

Sotto-servizio base “M” - Attivazione di sistemi di supporto alla memorizzazione dei Big Data

Il servizio in oggetto permette alle Amministrazioni di usufruire di un supporto del Fornitore per l’attivazione di sistemi Big Data distribuiti e scalabili, finalizzati alla memorizzazione e all’archiviazione dei dati che ricadono in almeno una delle seguenti fattispecie:

- **Data Volume** - tale casistica prevede che i dati, a causa del loro elevato volume, non possano essere memorizzati in sistemi centralizzati con tecnologie e database tradizionali, e che quindi necessitano di sistemi distribuiti. (Esempio di applicazione: “*click-stream analysis*”, memorizzazione di file di log dell’ordine dei TeraByte per l’analisi del comportamento dell’utente durante la navigazione



delle pagine di un sito Web istituzionale);

- **Data Velocity** - tale aspetto indica dati generati a velocità superiori alle attuali capacità dei tradizionali database. (Esempio di applicazione: “*sentiment analysis*”, gestione di un flusso di dati *real-time* proveniente dai social network al fine di analizzare il gradimento corrente degli utenti verso un determinato servizio dell’Amministrazione);
- **Data Variety** - questa terza fattispecie indica la variazione dello schema dei dati e/o la presenza di differenti tipi di dati oltre alla tradizionale struttura “relazionale” dei dati tipica dei RDBMS. I dati da gestire devono quindi essere anche in forma non strutturata senza schema (es. foto, video, testi, ecc) o semi-strutturata (es. documenti xml, dati di sensori in formato json, file di log in formato csv, ecc.). (Esempio di applicazione: correlazione di dati “strutturati” provenienti dai database interni all’Amministrazione con dati “semi-strutturati” provenienti dall’esterno).

Al fine di soddisfare le diverse casistiche e problematiche che si possono presentare nella gestione dei Big Data, il servizio consente alle Amministrazioni di attivare uno o più sistemi di supporto alla memorizzazione scegliendo tra le seguenti categorie:

- **Sistemi di gestione dati mediante architetture di file system distribuiti** (es. HDFS). Tali sistemi devono prevedere meccanismi trasparenti all’utente di gestione delle ridondanze dei dati finalizzati alla tolleranza ai guasti (i.e. gestione *failover*), dell’accesso ai dati secondo il paradigma di programmazione *MapReduce* (es. *Apache Hadoop*) e/o attraverso un’interfaccia logica di tipo *SQL-like* (es. *Apache Pig*). Questi sistemi sono progettati per una scalabilità orizzontale, ossia per aumentare le prestazioni del sistema tramite incremento dei nodi del cluster infrastrutturale;
- **Sistemi non relazionali a scalabilità orizzontale per la gestione dei dati tipicamente in modalità distribuita, comunemente indicati con il termine NoSQL.** Le categorie di sistemi NoSQL a disposizione per l’Amministrazione sono:
 - **Key-Value Store:** sistemi NoSQL in cui i dati sono organizzati secondo una mappa chiave-valore e l’accesso ai dati può essere effettuato solo attraverso la chiave;
 - **Document Store:** sistemi NoSQL che memorizzano i dati tramite unità informative distinte e indipendenti dette “documenti” in forma “semi-strutturata”. Nello specifico, i *document store* devono almeno gestire i dati in formato JSON;
 - **Extensible Record Store o Wide-Column Store:** sistemi NoSQL in cui i dati sono



organizzati tramite famiglie statiche di attributi ma che permettono l'aggiunta e la rimozione dinamica di singoli attributi all'interno di una famiglia;

- *Graph Database Management System*: sistemi NoSQL che memorizzano nativamente i dati modellati sotto forma di strutture dati a grafo. In particolare, questi sistemi gestiscono database caratterizzati dalla proprietà di *index-free-adjacency* (i.e. non si fa uso di indici globali per determinare la presenza di archi del grafo dei dati) e secondo il modello dei dati chiamato *property graph* (i.e. nodi e archi contengono una mappa chiave-valore).

Le modalità di erogazione del sotto-servizio obbligatorio “Attivazione Sistemi di supporto alla memorizzazione” sono le seguenti:

- **“as-a-Service”** - In tale modalità i sistemi di supporto alla memorizzazione selezionati dall'Amministrazione sono erogati dal Fornitore presso il proprio Centro Servizi per la fruizione da remoto.
- **“on-premise”** - Questa seconda modalità prevede l'erogazione da parte del Fornitore di tutte le attività sistemistiche di installazione, configurazione e deploy dei sistemi di gestione dati scelti dall'Amministrazione utilizzando le infrastrutture e le eventuali licenze software messe a disposizione dall'Amministrazione stessa.

In caso di erogazione del sotto-servizio in modalità “as-a-Service”, sono incluse le seguenti attività a carico del Fornitore:

- attivazione e completa gestione, in modo totalmente trasparente per l'Amministrazione, dell'infrastruttura hardware e software necessaria per l'erogazione del servizio;
- gestione applicativa (manutenzione correttiva ed evolutiva), in modo totalmente trasparente per l'Amministrazione, dei sistemi di supporto alla memorizzazione;
- monitoraggio del servizio;
- supporto tecnico H24x7 e Help Desk dedicato.

In caso di erogazione del sotto-servizio in modalità “on-premise”, sono previste invece le seguenti attività:

- attività sistemistiche di istanziazione dei sistemi software di gestione dati sulle macchine fisiche/virtuali messe a disposizione dall'Amministrazione. Per singola “istanziazione” si intende l'installazione e la configurazione di un singolo sistema di gestione dati su una singola macchina fisica/virtuale;
- collaudo dei sistemi di supporto alla memorizzazione installati;



- assistenza per l'avvio in esercizio.

Sotto-servizio opzionale "A" - Assessment dello scenario applicativo

Il servizio di "Supporto alla memorizzazione dei Big Data" offre opzionalmente all'Amministrazione di usufruire di un servizio professionale fornito in modalità "on-premise" da parte del Fornitore per l'assessment finalizzato alla determinazione dei sistemi di gestione dati ottimali per le esigenze dell'Amministrazione inerenti la memorizzazione e l'archiviazione dei Big Data e per lo scenario applicativo di utilizzo.

Tale servizio professionale prevede le seguenti attività:

- 1 Analisi iniziale delle esigenze dell'Amministrazione e della situazione attuale;
- 2 Realizzazione di uno Studio di Fattibilità tecnico-economica finalizzato ad indirizzare al meglio l'intervento di attivazione dei sistemi di supporto alla memorizzazione dei Big Data. Tramite lo Studio sarà possibile per l'Amministrazione ottenere tutte le informazioni necessarie per **decidere sull'opportunità** di intraprendere l'acquisto dei sistemi di supporto alla memorizzazione dei Big Data sulla base delle esigenze individuate.
- 3 Redazione di un Documento finale di Fattibilità contenente almeno le seguenti informazioni:
 - Obiettivi dell'Amministrazione;
 - Macro-requisiti dell'Amministrazione;
 - Proposta di almeno due possibili implementazioni del sotto-servizio di Supporto alla memorizzazione dei Big Data che possono differire per tipologia di sistema, tipologia di configurazione e strumenti software utilizzati, con indicazione delle caratteristiche di erogazione del servizio;
 - Descrizione di ognuna delle due implementazioni:
 - Numero e tipologia di sistemi di gestione dati da acquisire e da attivare;
 - Stima del dimensionamento dei dati da gestire mensilmente tramite i sistemi di gestione dei Big Data;
 - Dimensionamento del cluster distribuito di macchine in termini di numero di nodi da attivare, risorse hardware e software minime per ogni nodo del cluster;
 - Descrizione degli eventuali sotto-servizi opzionali "O" da attivare per attività di "configurazione avanzata" dei sistemi di gestione dati attivati tramite il sotto-servizio base (es. configurazione della modellazione dati, configurazione indicizzazione secondari, ecc.);
 - Stima dei tempi e dell'effort necessari per l'attivazione dei Sistemi di gestione dati.



Sotto-servizio opzionale “E” - Conduzione dei sistemi software di supporto alla memorizzazione installati “on-premise”

Il sotto-servizio in oggetto comprende la conduzione di ogni singola istanziazione dei sistemi di gestione dati installati e configurati sulle macchine fisiche/virtuali messe a disposizione dall’Amministrazione.

Tale sotto-servizio potrà essere acquisito solo subordinatamente all’acquisizione del servizio base “M - Attivazione Sistemi di supporto alla memorizzazione” in modalità “on-premise”.

Il sotto-servizio di conduzione dovrà essere obbligatoriamente erogato in modalità di esecuzione **continuativa** “on-premise”.

Sotto-servizio opzionale “O” - Attività opzionali di “configurazione avanzata” dei sistemi di gestione Big Data

Il sotto-servizio in oggetto offre opzionalmente all’Amministrazione di usufruire di attività finalizzate alla “configurazione avanzata” dei sistemi di gestione acquistati (sia in modalità “on-premise” che “as-a-Service”) al fine di soddisfare specifiche esigenze dell’Amministrazione. Tali attività possono essere acquisite a discrezione dell’Amministrazione e/o possono risultare necessarie, a valle dell’Assessment, per la completa ed efficace configurazione della soluzione di memorizzazione dati.

Le attività previste e attivabili dall’Amministrazione sono di seguito elencate e classificate in due tipologie:

Tipologia A - Attività di modellazione/migrazione dati

- Configurazione della modellazione dei dati da gestire tramite i sistemi di supporto alla memorizzazione attivati con il servizio in oggetto, nel caso di basi dati di nuova creazione;
- Configurazione della migrazione dei dati da una base dati esistente presso il CED dell’Amministrazione a una base dati implementata usando i sistemi di supporto alla memorizzazione attivati con il servizio in oggetto.

Tipologia B - Attività sistemistiche

- Implementazione/attivazione di meccanismi che abilitino il partizionamento e la distribuzione dei dati secondo metodi non automatici (i.e. meccanismi che necessitano di una configurazione manuale sui sistemi di supporto alla memorizzazione);
- Installazione e configurazione del software che consente la computazione dei dati secondo il paradigma MapReduce;
- Installazione e configurazione del software necessario per consentire l’accesso ai dati tramite interfacciamento SQL-like, ove il sistema di gestione



dati lo consenta;

- Configurazione di meccanismi di indicizzazione secondaria sui dati, ove il sistema di gestione dati lo consenta;
- Configurazione di meccanismi transazionali per l'accesso ai dati, in conformità ai gradi di libertà offerti dal sistema di gestione dati.

1.3.1.2. Requisiti funzionali

Nell'ambito del servizio "Supporto alla memorizzazione dei Big Data", il Fornitore deve garantire almeno i seguenti requisiti funzionali:

- Possibilità di scegliere tra le due seguenti modalità di installazione del sistema di gestione dei Big Data afferente alle categorie NoSQL:
 - **Installazione in modalità "server"**, ossia è prevista l'installazione del sistema NoSQL in modo da permettere al software applicativo di accedere ai dati tramite interfacce API;
 - **Installazione in modalità "embedded"**, ossia il sistema NoSQL di gestione dei dati è interno al software applicativo che fa uso dei dati.
- Supporto alle operazioni CRUD sui dati memorizzati dai sistemi di gestione dei Big Data, ossia supporto alle operazioni utente "*Create*" (creazione delle unità informative), "*Read*" (lettura dei dati), "*Update*" (aggiornamenti dei dati) e "*Delete*" (cancellazione dei dati);
- Disponibilità, per i sistemi di gestione dei Big Data installati in modalità "server", di strumenti basati su interfacce grafiche *user-friendly* (interfacce *web-based* o client/server) per l'interazione tra l'utente e i dati memorizzati. Tali strumenti devono consentire almeno le operazioni CRUD sui dati.

1.3.1.3. Requisiti tecnici e tecnologici

Nell'ambito del servizio "Supporto alla memorizzazione dei Big Data", il Fornitore deve garantire almeno i seguenti requisiti tecnici e tecnologici:

- In caso di erogazione del servizio in modalità "*as-a-Service*", le singole macchine fisiche messe a disposizione dal Fornitore presso il proprio Centro Servizi devono avere uno spazio disco massimo di 1TB;
- In caso di erogazione del servizio in modalità "*as-a-Service*", l'infrastruttura virtuale messa a disposizione dal Fornitore, comunque riservata alla singola Amministrazione, può essere ospitata da una infrastruttura hardware comune e condivisa tra le Amministrazioni;



- Possibilità di scegliere, per ogni sistema di gestione dei Big Data, tra le seguenti tipologie di memorizzazione dei dati:
 - configurazione in modalità di archiviazione su disco, ossia utilizzando supporti non volatili (persistenti),
 - configurazione in modalità di archiviazione “in-memory”, ossia utilizzando supporti di memoria volatile.
- Possibilità di abilitazione di modalità automatiche di distribuzione dei dati nel caso di sistemi di gestione dei Big Data installati in modalità “server” su un cluster distribuito. Le modalità automatiche previste di distribuzione dei dati devono essere almeno una tra la replicazione, lo sharding o altre forme di partizionamento orizzontale dei dati su un cluster distribuito di macchine.
- Disponibilità di interfacce API per le operazioni CRUD sui dati.

1.3.1.4. Tipologia del Servizio

Il servizio “Supporto alla memorizzazione dei Big Data” può essere erogato in una delle seguenti modalità:

- Modalità “continuativa” di tipo “*as-a-Service*”. In tale modalità i sistemi Big Data di gestione dati acquistati dall’Amministrazione sono messi a disposizione dal Fornitore presso il proprio Centro Servizi;
- Modalità “progettuale” di tipo “*on-premise*”. Questa seconda modalità prevede l’erogazione da parte del Fornitore di tutte le attività sistemistiche di installazione, configurazione e deploy dei sistemi di gestione dati scelti dall’Amministrazione utilizzando le infrastrutture e il software messi a disposizione dall’Amministrazione stessa.

1.3.1.5. Parametri di valutazione economica

Ai fini della valutazione economica, i relativi parametri sono di seguito esplicitati in base alle due modalità di erogazione previste per il servizio, ossia “*as-a-Service*” e “*on-premise*”.

Valutazione economica per la modalità di erogazione “*as-a-Service*”

Il costo C_s del servizio di “Supporto alla memorizzazione dei Big Data” in modalità “*as-a-Service*” è determinato dalla seguente formula:

$$C_s = A + M_s + O_s$$



Di seguito l'evidenza della modalità di remunerazione per ogni voce di costo presentata nella formula:

“A” - Sotto-servizio opzionale di Assessment dello scenario applicativo

La Modalità di remunerazione del sotto-servizio opzionale di Assessment è a “corpo [gg/p]” in base ai giorni persona concordati ed erogati dal Fornitore.

Ai fini della valutazione economica, il Fornitore dovrà indicare la quotazione espressa in [€/giorno] per ognuna delle figure professionali, la cui descrizione è presente in Appendice 1, relative al mix definito nella seguente tabella:

| Figura professionale | Impiego |
|-------------------------|---------|
| Capo Progetto | 5% |
| Data Scientist | 40% |
| Specialista di Prodotto | 20% |
| Data Architect | 35% |

“M_s” - Attivazione di sistemi di supporto alla memorizzazione dei Big Data

La modalità di remunerazione per i Sistemi di supporto alla memorizzazione “as-a-Service” è a “canone (utilizzo risorse su base mensile)” sulla base del volume di dati [GB/TB] gestiti complessivamente dai Sistemi di supporto alla memorizzazione scelti dall'Amministrazione ed erogati tramite le infrastrutture del Centro Servizi del Fornitore.

Ai fini della valutazione economica deve essere presentata una quotazione [€/mese] per ciascuno dei seguenti elementi:

- [1GB] per Fascia 1: Volume mensile complessivo di dati compreso tra 100GB e 1TB
- [1GB] per Fascia 2: Volume mensile complessivo di dati maggiore di 1TB e fino a 10TB
- [1GB] per Fascia 3: Volume mensile complessivo di dati maggiore di 10TB

L'Amministrazione garantisce la contrattualizzazione con il Fornitore per almeno 1 (un) anno.

“O_s” - Attività opzionali di “configurazione avanzata” dei sistemi di gestione Big Data

La Modalità di remunerazione dei sotto-servizi Opzionali associabili ai Sistemi di supporto alla memorizzazione è a “corpo” in base al numero di attività da realizzare.

Il costo del sotto-servizio “O” è quindi determinato dalla seguente formula:

$$O = \sum_{i=1}^N O_tipoA + \sum_{j=1}^M O_tipoB$$

Dove “O_{tipoA}” è il costo di una singola attività di Tipologia A, “O_{tipoB}” è il costo di



una singola attività di Tipologia B.

Ai fini della valutazione economica deve essere presentata una quotazione [€/attività] per ciascuno dei seguenti elementi:

- Singola attività di “configurazione avanzata” relativa alla Tipologia A (attività di modellazione/migrazione dati)
- Singola attività di “configurazione avanzata” relativa alla Tipologia B (attività sistemistiche)

Valutazione economica per la modalità di erogazione “on-premise”

Il costo C_p del servizio di “Supporto alla memorizzazione dei Big Data” in modalità “on-premise” è determinato dalla seguente formula:

$$C_p = A + M_p + E_p + O_p$$

Di seguito l’evidenza della modalità di remunerazione per ogni voce di costo presentata nella formula:

“A” - Componente opzionale di Assessment dello scenario applicativo

La Modalità di remunerazione della componente opzionale di Assessment è equivalente alla modalità indicata per il costo del servizio “as-a-Service”.

“M_p” - Attivazione di Sistemi di supporto alla memorizzazione dei Big Data

La modalità di remunerazione dei Sistemi di supporto alla memorizzazione in modalità “on-premise” è a “corpo” in base al numero complessivo di singole istanziazioni dei sistemi di gestione dati sui nodi (macchine fisiche o virtuali) messi a disposizione dall’Amministrazione.

Per singola “istanziatura” si intende l’installazione e la configurazione di un singolo sistema di gestione dati su una singola macchina fisica/virtuale.

Si precisa che l’Amministrazione deve mettere a disposizione, presso il proprio Centro di Elaborazione Dati, le macchine fisiche o virtuali con sistema operativo installato.

Ai fini della valutazione economica deve essere presentata una quotazione [€/istanziatura] per una singola istanziazione (installazione e configurazione) del sistema di gestione Big Data su un singolo nodo (macchina fisica/virtuale) dell’Amministrazione.

“E_p” - Conduzione dei sistemi software di supporto alla memorizzazione installati “on-premise”

La modalità di remunerazione del servizio di conduzione e manutenzione software dei Sistemi di supporto alla memorizzazione è a “canone (su base annuale)” per singola istanziazione.

Per singola istanziazione si intende l’installazione e la configurazione di un singolo



sistema di gestione dati su una singola macchina fisica/virtuale.

L'Amministrazione garantisce la contrattualizzazione con il Fornitore per almeno 1 (un) anno di manutenzione.

Ai fini della valutazione economica deve essere presentata una quotazione [€/anno] per il canone di conduzione e manutenzione software comprensivo di supporto tecnico per una singola istanziazione del sistema di gestione dati su un nodo del cluster infrastrutturale a supporto dei sistemi di gestione dati installati *“on-premise”*.

“O_p” - Attività opzionali di “configurazione avanzata” dei sistemi di gestione Big Data

La Modalità di remunerazione delle attività opzionali di “configurazione avanzata” dei sistemi di gestione dati è equivalente alla modalità indicata per il costo del servizio *“as-a-Service”*.

1.3.2. Servizio L3.S7 - Supporto all'analisi dei Big Data

1.3.2.1. Descrizione del Servizio

Il servizio comprende le attività, le metodologie e gli strumenti di supporto alle analisi di dati dell'Amministrazione (fonti dati interne eventualmente gestite tramite i sistemi acquisiti con il servizio L3.S6 di “Supporto alla memorizzazione dei Big Data”), e/o provenienti da fonti esterne all'Amministrazione (es. dati relativi a social network, open data, ecc.).

Obiettivo di tale servizio è l'analisi di tipo Big Data per individuare trend, pattern nascosti e nuove correlazioni tra i dati, al fine di supportare l'Amministrazione nei processi interni decisionali e amministrativi.

Il servizio in oggetto è di natura progettuale ed è composto dalle seguenti quattro Fasi di processo:

- **Fase 1** - Valutazione preliminare delle esigenze dell'Amministrazione per l'analisi dei Big Data;
- **Fase 2** - Acquisizione dei dati provenienti da fonti interne / esterne all'Amministrazione;
- **Fase 3** - Formulazione e implementazione del modello di analisi;
- **Fase 4** - Conduzione della soluzione di analisi realizzata.

Si precisa che l'Amministrazione può acquisire l'intero processo previsto dal servizio in oggetto, oppure può selezionare un sottoinsieme delle fasi, da attivare a seconda delle specifiche esigenze.



Di seguito sono descritte in dettaglio la quattro Fasi previste per il servizio:

Fase 1. Valutazione

La Fase in oggetto consente all'Amministrazione di usufruire di un servizio professionale fornito in modalità “on-premise” da parte del Fornitore per la Valutazione iniziale delle esigenze dell'Amministrazione per l'analisi dei Big Data.

Tale Fase, nel dettaglio, consiste nella raccolta preliminare da parte del Fornitore di esigenze e requisiti espressi dall'Amministrazione per le specifiche attività di analisi dei Big Data. Tali esigenze devono essere poi tradotte dal Fornitore in specifiche funzionali e tecnologiche relative alla soluzione da adottare per l'analisi richiesta dall'Amministrazione.

In caso di scelta di utilizzo da parte dell'Amministrazione di tale Fase di Valutazione, il Fornitore dovrà produrre un documento formale contenente almeno le seguenti informazioni:

- Obiettivi dell'Amministrazione;
- Elenco requisiti di analisi (requisiti funzionali e non funzionali, tra cui requisiti prestazionali, tempi di risposta, ecc.);
- Strumenti tecnici e tecnologici previsti per l'analisi;
- Descrizione di massima della tipologia di analisi prevista;
- Stima di tempi, costi ed effort di realizzazione delle successive Fasi di processo;
- Tipologia prevalente di elaborazione prevista per l'analisi (es. “batch”, “real-time”, “one-shot”, ecc);
- Numero e descrizione dei connettori previsti per l'acquisizione dei dati provenienti da fonti interne / esterne all'Amministrazione, comprensivo della descrizione delle fonti dati;
- Tipologia di infrastruttura necessaria per l'analisi;
- Modalità di export dell'output dell'analisi Big Data.

Il Fornitore, prima di poter procedere con le successive Fasi di processo, deve ottenere dall'Amministrazione l'approvazione formale del documento di Valutazione.

Fase 2. Acquisizione dati

La seconda Fase del servizio consiste nella realizzazione di connettori di import dei dati da analizzare tramite il sistema di Analisi Big Data e/o di connettori di export del risultato delle analisi. I dati da acquisire possono provenire da molteplici fonti eterogenee interne e/o esterne all'Amministrazione come ad esempio i data warehouse interni, i database relazionali interni, dati relativi ai social network, dati provenienti da



dispositivi di sensoristica, ecc., ed essere presenti in qualsiasi forma (es. strutturata, semi-strutturata e non-strutturata).

Per ogni fonte dati oggetto di analisi, l'Amministrazione richiede al Fornitore la realizzazione di uno specifico connettore, ossia di un'interfaccia software in grado di catturare e/o leggere i dati provenienti da una sorgente dati interna o esterna al Centro di Elaborazione Dati dell'Amministrazione e di effettuare il caricamento di tali dati nel sistema di analisi Big Data.

Per tale Fase di Acquisizione dati sono previste le seguenti due Classi di Connettori:

- **Classe A - Connettori "standard"** per import di dati. Sono inclusi in questa classe i connettori per l'import di dati dai sistemi di gestione dati acquisiti tramite il Servizio L3.S6 "Supporto alla memorizzazione dei Big Data", il riuso di connettori già realizzati presso altre Amministrazioni, e connettori basati sull'utilizzo di apposite interfacce software API messe a disposizione da sistemi esterni quali ad esempio social network (es. Twitter, Facebook, ecc.). Sono esclusi dal servizio in oggetto eventuali costi aggiuntivi di "*Licensing fee*" previsti per l'accesso a dati non pubblici/storici messi a disposizione da siti web/social network. Nella realizzazione del connettore "standard" sono incluse solo attività di filtraggio e campionamento dei dati in ingresso.
- **Classe B - Connettori "custom"** realizzati per l'import dei dati da sistemi interni all'Amministrazione, per l'import di dati da sistemi di sensoristica/macchine oppure per l'import di dati tramite attività di estrazione dati (i.e. *web crawling/scraping*) da siti web che non mettono a disposizione apposite interfacce API. Nella realizzazione del connettore "custom" sono incluse attività di filtraggio, campionamento, trasformazione e pulizia dei dati in ingresso.

Fase 3. Formulazione del modello di analisi

La Fase in oggetto consente all'Amministrazione di usufruire di un servizio professionale da parte del Fornitore per la formulazione e l'implementazione di un modello, algoritmo o programma (di seguito il "modello") contenente l'intelligenza necessaria per analizzare i dati secondo le specifiche concordate con l'Amministrazione.

La Fase in oggetto prevede le seguenti due modalità di erogazione:

- "*as-a-Service*" - in questa modalità il Fornitore mette a disposizione presso il proprio Centro Servizi una infrastruttura cluster, hardware e software, su cui effettuare l'implementazione del modello;
- "*on-premise*" - in questa modalità l'Amministrazione mette a disposizione le macchine fisiche/virtuali necessarie per la predisposizione da parte del Fornitore



di una infrastruttura cluster, hardware e software, su cui effettuare l'implementazione del modello.

La formulazione e implementazione del modello di analisi può appartenere a una delle tre differenti categorie di analisi:

- **Data Mining:** include classiche analisi di Data Mining che fanno uso sia di tecniche di Machine Learning che di metodi statistici. Sono incluse in questa categoria le stesse tecniche applicate ai dati strutturati a grafo, ossia tecniche di *graph mining*.
- **Stream Processing:** include il *real-time analytics*, lo *stream analytics* e il *complex event processing* (i.e. attività di tracciatura, analisi e processamento di dati attivate a seguito di specifici eventi). Queste analisi sono spesso utilizzate per le rilevazioni scientifiche e ambientali (ad esempio dai dati di sensoristica), nell'ambito delle *smart city*, nei sistemi di rilevamento delle frodi e degli attacchi di sicurezza. Questa categoria di analisi deve essere in grado di processare dati dinamicamente, ossia in tempo reale o quasi in tempo reale;
- **Text analysis:** analisi di dati testuali non strutturati che include il *natural language processing*, la *sentiment analysis*, la *trend analysis*.

Si precisa che la tipologia di analisi scelta per il modello può essere ibrida rispetto alle categorie sopraccitate. Ad esempio, possono essere applicate analisi di tipo *Text analysis* all'interno di un processo di analisi *Data Mining*. Ad ogni modo, deve sempre essere presa in considerazione una tipologia prevalente di analisi.

La formulazione potrà avvenire tramite l'utilizzo di best practices o tramite l'utilizzo di modelli di auto-apprendimento come nel caso *Machine Learning*.

La Fase in oggetto comprende, infine, le attività di configurazione e attivazione della modalità di export dell'output dell'analisi Big Data. Le possibili modalità previste di export dell'output sono le seguenti:

- Utilizzo di un formato tabellare standard (es. CSV) o compatibile con altri sistemi di reporting/visualizzazione dati o compatibile con i sistemi gestionali dell'Amministrazione;
- Utilizzo di un formato aperto (es. XML, RDF, ecc.) per l'output previsto dall'analisi o altro formato richiesto dall'Amministrazione. Tale modalità è consigliabile per elaborazioni il cui volume di dati in output è contenuto (dell'ordine dei MegaByte) e per elaborazioni "batch" o "one-shot";
- Utilizzo di un Sistema di gestione dati acquisito dall'Amministrazione tramite il servizio L3.S6 "Supporto alla Memorizzazione dei Big Data". In tale caso deve



essere realizzato almeno un connettore di export dell'output di analisi verso il Sistema di gestione dati;

- Utilizzo di storage tradizionale acquisibile in modalità “as-a-Service” tramite il Servizio “IaaS: Virtual Storage” del Lotto 1 o tramite servizi offerti da altri fornitori. In tale caso deve essere realizzato almeno un connettore di export dell'output di analisi verso lo storage “as-a-Service”.

Fase 4. Conduzione della soluzione di analisi

Questa quarta Fase consiste nella conduzione applicativa del modello formulato e implementato nella Fase 3 precedente. I dati sulla quale la computazione viene eseguita sono quelli provenienti dalle fonti dati acquisite tramite la Fase 2 di Acquisizione dati interni e/o esterni all'Amministrazione.

Il risultato di questa Fase è l'individuazione di trend, pattern nascosti e nuove correlazioni tra i dati, al fine di abilitare l'Amministrazione nei suoi processi interni decisionali e amministrativi.

In caso di modalità di erogazione del servizio “as-a-Service”, la Fase comprende i costi relativi all'erogazione di una infrastruttura di calcolo messa a disposizione dal Fornitore presso il proprio Centro Servizi per l'implementazione del modello di analisi. In caso di erogazione del servizio in modalità “as-a-Service”, è possibile acquistare la Fase 4 in oggetto solamente insieme alla Fase 3 di Formulazione.

Sono previste nel servizio le seguenti tre possibili tipologie di cluster infrastrutturali “base”, eventualmente scalabili orizzontalmente (“scale-out”) tramite aggiunta di singoli nodi. Si precisa che la necessità di aggiunta di nodi all'infrastruttura “base” scelta dovrà essere comprovata da apposita documentazione e approvata dall'Amministrazione. Inoltre, per ogni tipologia di infrastruttura “base” è possibile solamente aggiungere nodi della stessa tipologia che soddisfino i requisiti minimi indicati al paragrafo §1.3.2.3 (es. per una infrastruttura “base” di tipo “general purpose” è possibile solamente aggiungere nodi “general purpose”).

| Tipologia Infrastruttura “Base” | Descrizione Infrastruttura | Numero di nodi richiesti | Requisiti Minimi nodo |
|---------------------------------|--|--------------------------|---|
| Cluster “General Purpose” | Infrastruttura base composta da un singolo nodo bilanciato in termini di risorse disco, ram, cpu | 1 | Requisiti tecnici e tecnologici minimi per i singoli indicati al paragrafo §1.3.2.3 |
| Cluster “CPU-intensive” | Infrastruttura base dedicata a elaborazione/processamento dati dove è predominante il consumo di CPU (es. text mining) | 5 | |



| Tipologia Infrastruttura "Base" | Descrizione Infrastruttura | Numero di nodi richiesti | Requisiti Minimi nodo |
|---|---|--------------------------|-----------------------|
| Cluster "RAM-intensive" di Fascia Bassa | Infrastruttura base dedicata a elaborazione/processamento dati dove è predominante il consumo di memoria (es. real-time processing) | 5 | |
| Cluster "RAM-intensive" di Fascia Alta | | 10 | |

In caso di modalità di erogazione del servizio "as-a-Service", il Fornitore potrà fornire macchine con prodotti software open source, mentre in caso di modalità "on-premise" sarà l'Amministrazione a fornire le licenze necessarie per la predisposizione dell'infrastruttura di calcolo.

1.3.2.2. Requisiti funzionali

Nell'ambito del servizio "Supporto all'analisi dei Big Data", il Fornitore deve garantire almeno i seguenti requisiti funzionali:

- Il connettore da realizzare durante la Fase 2 "Acquisizione dati" deve essere in grado di catturare tutti i dati in ingresso, ossia si deve evitare l'information loss. Il connettore deve, quindi, supportare il throughput della fonte dati;
- Durante la Fase 2 "Acquisizione dati" deve essere possibile filtrare alcuni dati in base ad alcune caratteristiche specificate dall'Amministrazione;
- Possibilità di campionare, durante la Fase 2 con dati "in-stream", i dati in input sulla base di caratteristiche specificate dall'Amministrazione;
- Le tecniche di analisi nella fase di formulazione devono appartenere almeno a una delle seguenti tipologie: classificazione, clustering, regressione e correlazione statistica, reti neurali, alberi di decisione;
- In caso l'Amministrazione metta a disposizione del fornitore strumenti, software o hardware, è compito del Fornitore valutare l'idoneità degli stessi sulla base delle funzionalità e delle prestazioni richieste per le analisi.

1.3.2.3. Requisiti tecnici e tecnologici

Nell'ambito del servizio "Supporto all'analisi dei Big Data", il Fornitore deve garantire almeno i seguenti requisiti tecnici e tecnologici:

- Devono essere previste le seguenti modalità di acquisizione dei dati:
 - Acquisizione "real-time" - il connettore deve essere in grado di catturare i dati in ingresso in tempo reale;



- Acquisizione “batch” - il connettore deve essere configurato per catturare i dati a intervalli regolari e definiti (es. una volta al giorno a un’ora prestabilita) oppure configurato per analisi “one-shot”;
- L’architettura tecnica da implementare per il servizio di analisi dovrà garantire il rispetto dei requisiti non funzionali condivisi con l’Amministrazione;
- In caso di erogazione “as-a-Service” del servizio, l’infrastruttura dovrà essere predisposta presso i Centri Servizi conformi alle specifiche di cui al “Capitolato tecnico - parte generale”;
- In caso di erogazione “as-a-Service” del servizio, l’infrastruttura dovrà permettere l’incremento delle risorse di elaborazione tramite aggiunta di nodi (“scale-out”);
- In caso di erogazione “as-a-Service” del servizio, dovranno essere garantiti i seguenti requisiti per i singoli nodi di elaborazione:

| Tipologia nodo | Requisito CPU | Requisito RAM | Requisito Disco |
|--------------------------------------|-----------------------------|----------------------------------|--|
| Nodo “General Purpose” | Almeno CPU [2Ghz] dual-core | Memoria RAM minima pari a 8 GB | Disco di tipo SATA o SAS da almeno 7.2k rpm con capienza minima pari a 500GB |
| Nodo “CPU-intensive” | Almeno CPU [2,4Ghz] 16 core | Memoria RAM minima pari a 32 GB | Disco di tipo SATA o SAS da almeno 7.2k rpm con capienza minima pari a 500GB |
| Nodo “RAM-intensive” di Fascia Bassa | Almeno CPU [2Ghz] 8 core | Memoria RAM minima pari a 64 GB | Disco di tipo SATA o SAS da almeno 7.2k rpm con capienza minima pari a 500GB |
| Nodo “RAM-intensive” di Fascia Alta | Almeno CPU [2,4Ghz] 16 core | Memoria RAM minima pari a 128 GB | Disco di tipo SATA o SAS da almeno 7.2k rpm con capienza minima pari a 500GB |

1.3.2.4. Tipologia del Servizio

Il servizio “Supporto all’analisi dei Big Data” può essere erogato in una delle seguenti modalità:

Classificazione del documento: Consip Public

Procedura ristretta, suddivisa in 4 Lotti, per l’affidamento dei servizi di Cloud Computing, di Sicurezza, di Realizzazione di Portali e Servizi on-line e di Cooperazione Applicativa per le Pubbliche Amministrazioni - ID SIGEF 1403

Capitolato Tecnico Lotto 3



- **Modalità “continuativa” di tipo “as-a-Service”**. In tale modalità i sistemi hardware e le licenze software open source sono messi a disposizione dal Fornitore presso il proprio Centro Servizi;
- **Modalità “progettuale” di tipo “on-premise”**. Questa seconda modalità prevede l'erogazione da parte del Fornitore di tutte le attività sistemistiche di installazione, configurazione e deploy dei sistemi di analisi utilizzando le infrastrutture hardware e le eventuali licenze software messe a disposizione dall'Amministrazione stessa.

1.3.2.5. Parametri di valutazione economica

Ai fini della valutazione economica, i relativi parametri sono di seguito esplicitati in base alle quattro Fasi di processo e in base alle due modalità di erogazione previste per il servizio, ossia “as-a-Service” e “on-premise”.

Valutazione economica per la modalità di erogazione “as-a-Service”

Il costo C_s del servizio di “Supporto all’analisi dei Big Data” in modalità “as-a-Service” è determinato dalla seguente formula:

$$C_s = V + Q_s + F_s + E_s$$

Di seguito l’evidenza della modalità di remunerazione per ogni voce di costo presentata nella formula:

“V” - Fase 1. Componente opzionale di Valutazione

La Modalità di remunerazione della componente opzionale di Valutazione è a “corpo [gg/p]” in base ai giorni persona concordati ed erogati dal Fornitore.

Ai fini della valutazione economica, il Fornitore dovrà indicare la quotazione espressa in [€/giorno] per ognuna delle figure professionali, la cui descrizione è presente in Appendice 1, relative al mix definito nella seguente tabella:

| Figura professionale | Impiego |
|----------------------|---------|
| Capo Progetto | 10% |
| Data Scientist | 40% |
| Data Architect | 30% |
| Analista funzionale | 20% |

“Q_s” - Fase 2. Realizzazione connettori per import dati

La modalità di remunerazione prevista per tale Fase di Acquisizione dati tramite realizzazione di connettori è a “corpo” sulla base del numero di connettori realizzati.

Ai fini della valutazione economica deve essere presentata una quotazione [€/connettore] per ciascuno dei seguenti elementi:

- Realizzazione e attivazione singolo connettore di tipo **Classe A “Connettore Standard”**



- Realizzazione e attivazione singolo connettore di tipo **Classe B “Connettore Custom”**

“F₃” - Fase 3. Formulazione del modello di analisi

La Modalità di remunerazione delle attività di Formulazione e Implementazione del modello di analisi presso il Centro Servizi del Fornitore è a “corpo [gg/p]” in base ai giorni persona concordati ed erogati dal Fornitore.

Ai fini della valutazione economica, il Fornitore dovrà indicare la quotazione espressa in [€/giorno] per ognuna delle figure professionali, la cui descrizione è presente in Appendice 1, relative al mix definito nella seguente tabella:

| Figura professionale | Impiego |
|----------------------|---------|
| Capo Progetto | 5% |
| Data Scientist | 35% |
| Data Architect | 20% |
| Sviluppatore | 40% |

“E₄” - Fase 4. Esercizio e conduzione della soluzione di analisi (opzionale)

La fase in oggetto è opzionale e attivabile su richiesta dell’Amministrazione.

La modalità di remunerazione di questa Fase è composta dai seguenti due elementi di remunerazione:

- a “consumo (utilizzo risorse su base giornaliera)” per l’utilizzo del cluster infrastrutturale messo a disposizione dal Fornitore (infrastruttura “base” più eventuali nodi aggiunti al cluster);
- a “canone (su base trimestrale)” in base ai giorni persona/mese concordati tra il Fornitore e l’Amministrazione per le attività di conduzione applicativa della soluzione di analisi realizzata nella Fase 3. Il numero di giorni persona concordati su base trimestrale, determinati in relazione alla complessità della soluzione e alla frequenza di analisi richiesta dall’Amministrazione, non può in ogni caso superare il tetto massimo del 5% su base trimestrale, ovvero il 20% annuo. Il numero di giorni persona concordati è eventualmente rinegoziabile ogni 3 mesi.



Ai fini della valutazione economica della Fase 4 deve essere presentata una quotazione [€/giorno] distinta per l'utilizzo giornaliero delle seguenti tipologie di infrastrutture "base" previste nel servizio:

| Tipologia Infrastruttura "Base" | Descrizione Infrastruttura | Numero di nodi richiesti | Requisiti Minimi nodo |
|---|---|--------------------------|---|
| Cluster "General Purpose" | Infrastruttura base composta da un singolo nodo bilanciato in termini di risorse disco, ram, cpu | 1 | Requisiti tecnici e tecnologici minimi per i singoli indicati al paragrafo §1.3.2.3 |
| Cluster "CPU-intensive" | Infrastruttura base dedicata a elaborazione/processamento dati dove è predominante il consumo di CPU (es. text mining) | 5 | |
| Cluster "RAM-intensive" di Fascia Bassa | Infrastruttura base dedicata a elaborazione/processamento dati dove è predominante il consumo di memoria (es. real-time processing) | 5 | |
| Cluster "RAM-intensive" di Fascia Alta | | 10 | |

Il costo [€/giorno] di ogni singolo nodo aggiunto al cluster infrastrutturale di riferimento è pari a $\frac{1}{n}$ del prezzo offerto per il cluster infrastrutturale "base", dove "n" è il numero di nodi previsti nel cluster infrastrutturale "base" (es. di calcolo: Aggiunta di un nodo all'infrastruttura base "CPU-intensive". Se il prezzo offerto per l'infrastruttura base "CPU-intensive" è pari a 10, l'aggiunta di un nodo "CPU-intensive" all'infrastruttura è pari a $\frac{1}{5}$ di 10, ossia è pari a 2).

Il Fornitore dovrà, infine, indicare la quotazione espressa in [€/giorno] per il canone trimestrale opzionale e relativo alle attività di conduzione applicativa della soluzione di analisi, per ognuna delle figure professionali, la cui descrizione è presente in Appendice 1, relative al mix definito nella seguente tabella:

| Figura professionale | Impiego |
|----------------------|---------|
| Capo Progetto | 5% |
| Data Scientist | 5% |
| Data Architect | 15% |
| Sviluppatore | 25% |
| Sistemista | 50% |

Valutazione economica per la modalità di erogazione "on-premise"

Il costo C_p del servizio di "Supporto all'analisi dei Big Data" in modalità "on-premise" è determinato dalla seguente formula:



$$C_s = V + Q_p + F_p + E_p$$

Di seguito l'evidenza della modalità di remunerazione per ogni voce di costo presentata nella formula:

“V” - Fase 1. Componente opzionale di Valutazione

La Modalità di remunerazione della componente opzionale di Valutazione è equivalente alla modalità indicata per il costo del servizio “*as-a-Service*”.

“Q_p” - Fase 2. Realizzazione connettori per import dati

La modalità di remunerazione prevista per tale Fase di Acquisizione dati “*on-premise*” è equivalente alla modalità indicata per il costo del servizio “*as-a-Service*”.

“F_p” - Fase 3. Formulazione del modello di analisi

La Modalità di remunerazione delle attività di formulazione e implementazione del modello di analisi presso il Centro di Elaborazione Dati dell'Amministrazione è equivalente alla modalità indicata per il costo del servizio “*as-a-Service*”.

“E_p” - Fase 4. Esercizio e conduzione della soluzione di analisi (opzionale)

La fase in oggetto è opzionale e attivabile su richiesta dell'Amministrazione.

La modalità di remunerazione della conduzione applicativa è a “canone (su base trimestrale)” ed è equivalente alla modalità di remunerazione indicata nel caso “*as-a-Service*” di Fase 4.

- FINE DEL DOCUMENTO -