



Scheda Progetto - SDF  
Sistema Statistico  
Regionale Datawarehouse



REGIONE  
LAZIO



# Scheda Progetto

Sistema Statistico Regionale

Portale e Datawarehouse

programmazione regionale

v. 1.0 del 11/12/2018

LAZIOCREA S.P.A.

SEDE LEGALE E AMMINISTRATIVA: VIA DEL SERAFICO, 107 00142 ROMA -  
T +39 06 515631 F +39 06 51563611

UFFICI: VIA ADELAIDE BONO CAIROLI, 68 - ROMA 00145 -  
T +39 06 51689800 - F +39 06 51892207

## INDICE

1	Introduzione – scenario attuale.....	4
1.1	Linea Di Finanziamento .....	4
2	Obiettivi.....	5
3	Requisiti.....	6
3.1	Requisiti funzionali.....	6
3.2	Caratteristiche del sistema .....	6
3.2.1	Integrabilità di nuove fonti informative.....	6
3.2.2	Acquisizione e riordino dei dati pregressi .....	6
3.2.3	Integrazione banche dati e applicativi esistenti .....	7
3.2.4	Cruscotto di monitoraggio .....	7
3.2.5	Gestione dei metadati .....	7
3.2.6	Preparazione dei dati per inserimento nel DWH.....	7
3.2.7	Analisi ed elaborazione dei dati .....	8
3.2.8	Rispetto della normativa sulla privacy.....	8
3.2.9	Attori e ruoli.....	8
3.2.10	Open Data .....	9
3.2.11	Multilingua .....	9
4	Soluzione proposta .....	10
4.1	FASE 1 – Analisi dati e sorgenti informative .....	10
4.2	FASE 2 – Installazione e personalizzazione piattaforma DWH.....	11
4.3	FASE 3 – Integrazione fonti e ripresa dati.....	11
4.4	FASE 4 - Realizzazione portale, report, dashboard, analisi ed elaborati	12

	<p>Scheda Progetto - SDF Sistema Statistico Regionale Datawarehouse</p>	
---	---	--

4.4.1	Portale .....	12
4.4.2	Report.....	13
4.4.3	Dashboard.....	13
4.4.4	Analisi dimensionali.....	13
4.4.5	Elaborati vari .....	14
4.5	FASE 5 - Formazione e avvio operativo.....	15
4.6	FASE 6 - Manutenzione MAC + MEV.....	16
5	Dimensionamento fisico degli ambienti HW.....	17
6	Sorgenti informative da integrare.....	18
7	Benefici attesi.....	19
8	Cronoprogramma di massima delle attività.....	20
9	Analisi dei costi.....	21

	<p>Scheda Progetto - SDF Sistema Statistico Regionale Datawarehouse</p>	
---	---	--

# 1 Introduzione – scenario attuale

L'Ufficio Sistema Statistico Regionale ha tra i propri compiti quello di supportare le politiche di programmazione regionale mediante l'analisi degli archivi statistici di competenza e/o afferenti alla Regione Lazio. A tal proposito l'USSR si è dotata di alcuni strumenti informatici per poter effettuare una serie di comuni operazioni sulle sorgenti dati come: estrazioni, incrocio di dati, elaborazioni, ecc.

Quello che è stato constatato dal personale dell'USSR è che tali strumenti non riescono a far fronte alle nuove esigenze dell'ufficio in quanto non consentono di poter svolgere attività complesse sui dati e inoltre non consentono di garantire degli standard minimi di sicurezza nell'accesso alle informazioni.

L'avvento degli "open data", la continua e progressiva digitalizzazione della Pubblica Amministrazione e la necessità delle istituzioni di cooperare e quindi di intercambiare informazioni, ha portato l'USSR ad aver bisogno di un sistema che si avvicinasse il più possibile alle tecnologie di Datawarehousing e Datamining piuttosto che a dei semplici sistemi di data analysis.

Il presente studio ha proprio l'obiettivo di definire un nuovo sistema informatico in grado di rispondere pienamente alle esigenze e alle aspettative dell'USSR.

## 1.1 Linea Di Finanziamento

Il presente Intervento è volto a migliorare i sistemi informativi regionali finalizzati ad innovare e rafforzare i procedimenti amministrativi; l'intervento proposto per il PRA dovrà essere finanziato a valere del POR FSE Asse IV (OT 11).

	<p>Scheda Progetto - SDF Sistema Statistico Regionale Datawarehouse</p>	
---	---	--

## 2 Obiettivi

L'USSR si vuole dotare di uno strumento informatico (Datawarehouse) fruibile on line (web-oriented) che consenta il conseguimento dei seguenti macro-obiettivi:

- Centralizzare e sistematizzare le sorgenti dati della Regione Lazio e/o rese disponibili all'interno della Regione Lazio (archivio ISTAT, archivi provenienti da altre Pubbliche Amministrazioni, archivi riferibili direttamente o indirettamente alla Regione Lazio);
- Analizzare e tenere sotto controllo una serie di indicatori (economici, sociali, ecc) di interesse per la Regione Lazio;
- Fornire un supporto decisionale alle scelte delle politiche regionali;
- Fornire un supporto per la trend-analysis, il data-mining, i predictive-models nonché l'individuazione della correlazione tra fenomeni.

	<p>Scheda Progetto - SDF Sistema Statistico Regionale Datawarehouse</p>	
---	---	--

## 3 Requisiti

Di seguito vengono riportati i requisiti principali che devono essere soddisfatti dal sistema DWH.

### 3.1 Requisiti funzionali

Il DWH deve mettere a disposizione degli utenti una serie di strumenti in grado di eseguire le seguenti funzioni:

- Report
- Analisi OLAP
- Mappe tematiche
- Query di dati e metadati
- Visualizzazione, download e upload di dati

### 3.2 Caratteristiche del sistema

Il DWH deve consentire l'espletamento di un insieme di attività che sono considerate di vitale importanza per il Committente. Tali caratteristiche rientrano all'interno di uno scenario che vede il DWH come elemento che accentra tutte le informazioni eterogenee provenienti da più parti, le elabora e le mette a disposizione in diverse forme e modalità ai vari attori del sistema.

Vengono di seguito riportate tutte le attività necessarie all'avvio, gestione e utilizzo del DWH.

#### 3.2.1 Integrabilità di nuove fonti informative

Il DWH deve permettere di integrare nuove fonti informative di vario tipo: dati strutturati (es: database, tabelle relazionali, ...) e dati semi-strutturati (es: xml, csv, ...). L'integrazione di tali fonti può essere resa completamente automatizzata (dove possibile), semi-automatica o in alternativa, qualora il processo non sia automatizzabile, completamente manuale.

#### 3.2.2 Acquisizione e riordino dei dati pregressi

L'Ufficio Statistico Regionale è attivo da molto tempo e pertanto ha a disposizione una serie molto vasta di dati ed elaborazioni provenienti da vari archivi. Questo patrimonio informativo costituisce una risorsa molto importante per l'ufficio in quanto può essere utilizzato come base di confronto per le elaborazioni presenti e passate. A tal proposito il

	<p>Scheda Progetto - SDF Sistema Statistico Regionale Datawarehouse</p>	
---	---	--

DWH dovrà essere infasato all'atto della realizzazione, di tutti i dati che l'ufficio metterà a disposizione: dati attuali e storici, elaborazioni, ecc. A tal proposito potrebbe essere necessario un riordino delle informazioni dal momento che allo stato attuale, non essendo centralizzate, potrebbero presentare delle ridondanze e delle incongruenze che il DWH dovrebbe risolvere (dove possibile).

### 3.2.3 Integrazione banche dati e applicativi esistenti

Oltre ai dati storici e alle varie sorgenti informative già esistenti e in uso presso l'USR, è stato espressamente richiesto di prevedere l'integrazione con i seguenti sistemi:

- DEMOSTAR → dati e indicatori sulla popolazione del Lazio
- ATTIVITA' DEL TAVOLO TECNICO → attività statistiche regionali. In particolare:
  - o SCHEDE ATTIVITA' STATISTICHE REGIONALI
  - o TAVOLI TECNICI DELLE ATTIVITA' STATISTICHE REGIONALI

### 3.2.4 Cruscotto di monitoraggio

Il DWH deve poter definire un insieme di indicatori per consentire il monitoraggio di alcune tematiche di interesse per il Committente. La necessità è quella di monitorare lo stato di avanzamento del Programma Operativo in merito alla realizzazione degli interventi FESR FSE 2014-2020

Gli indicatori possono essere così classificati:

- Indicatori finanziari
- Indicatori di realizzazione
- Indicatori di risultato
- Indicatori di impatto.

### 3.2.5 Gestione dei metadati

La gestione dei metadati all'interno del DWH deve consentire di contestualizzare il dato per comprenderne fino in fondo il significato e la validità. Sono considerati metadati informazioni come: la sorgente di acquisizione, la metrica, intervallo di validità, la data in cui è stato selezionato il campione, ecc.

### 3.2.6 Preparazione dei dati per inserimento nel DWH

L'integrazione di una nuova sorgente informativa all'interno del DWH deve prevedere una serie di operazioni che permettano di gestire i dati acquisiti per successivi processamenti ed elaborazioni. A tal fine è pertanto indispensabile che il DWH consenta di poter definire una serie di operazioni (pre-processamento) necessarie a "preparare" i dati prima della loro acquisizione all'interno del DWH. Le operazioni più comuni devono avere come fine: il

	<p>Scheda Progetto - SDF Sistema Statistico Regionale Datawarehouse</p>	
---	---	--

raffinamento dei dati (pulizia), normalizzazione/denormalizzazione, aggregazione/disaggregazione, ecc.

### 3.2.7 Analisi ed elaborazione dei dati

Le sorgenti informative integrate nel DWH e i dati in esso acquisiti contribuiscono alla principale delle funzioni del DWH ovvero: l'analisi e l'elaborazione dei dati. Il DWH deve pertanto consentire l'incrocio dei dati, l'elaborazione e l'estrapolazione di report, il tutto nel modo più semplice e flessibile possibile, garantendo dei tempi di risposta sufficientemente rapidi.

### 3.2.8 Rispetto della normativa sulla privacy

L'attuale normativa sulla privacy (L.196/03) prevede una serie di restrizioni per quanto riguarda la fruizione di dati sensibili e personali. Il DWH dovrebbe consentire l'abilitazione/inibizione di tutte le restrizioni possibili per l'accesso al dato al fine di non contravvenire a quanto disposto dalla normativa vigente. In particolare si vuole inibire l'export massivo dei dati e l'inibizione di alcuni dati in funzione di alcune tipologie di utenti.

### 3.2.9 Attori e ruoli

Il DWH deve essere utilizzato da un insieme di attori che accedono alla piattaforma informatica con ruoli ed obiettivi differenti. Per fare ciò è indispensabile che il DWH offra un meccanismo di profilazione in grado di definire, per ogni ruolo, le risorse a disposizione (funzioni, aree applicative, elaborati, ecc). Una volta definiti i ruoli, tutti gli utenti saranno assegnati ad un profilo ed opereranno nei limiti imposti dal profilo utilizzato.

Nello scenario del SSR è possibile individuare i seguenti attori:

- Ufficio Statistico Regionale → Sono i principali utilizzatori del sistema, ovvero coloro che devono alimentare e gestire il DWH al fine di rendere fruibili i dati per lo stesso Ufficio e per gli altri attori del processo.
- Direzioni regionali → Sono utenti appartenenti allo stesso dominio di competenza (Regione Lazio) ma con funzioni e operatività specifiche.
- Agenzie ed Enti strumentali → Sono utenti appartenenti a domini differenti da quello regionale ma che possono aver necessità di accesso su singole e specifiche tematiche a seconda della competenza dell'Agenzia o dell'Ente.
- Cittadini ed Imprese → Sono utenti che accedono ai dati di pubblico interesse per varie finalità (statistiche, studi di settore, ecc)



	<p>Scheda Progetto - SDF Sistema Statistico Regionale Datawarehouse</p>	
---	---	--

### 3.2.10 Open Data

Gli Open Data si stanno affermando come formato di interscambio in moltissime realtà della PA. Open Data significa offrire a Cittadini, Aziende e altre Pubbliche Amministrazioni una fonte dati “ufficiale” e utilizzabile per molteplici scopi (statistici, elaborativi, ecc).

Il DWH dovrà consentire la pubblicazione e il download (da parte di utenti anonimi) di un sottoinsieme di dati predefinito in formato Open Data.

### 3.2.11 Multilingua

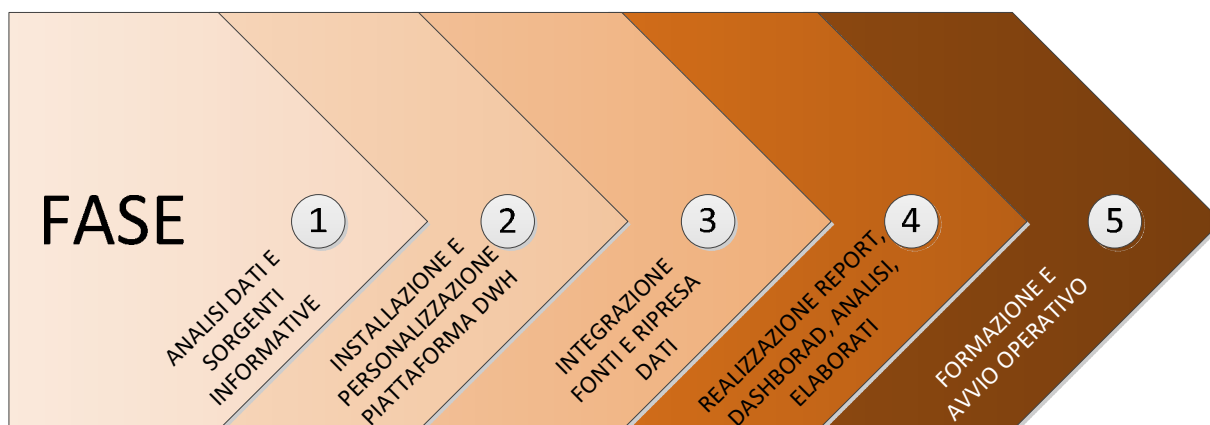
Si ritiene necessario che la gestione dei dati e metadati sia pensato in ottica multilingua (italiano e inglese) per consentire l’accesso alle informazioni da parte di tutti i possibili fruitori delle informazioni.

## 4 Soluzione proposta

Dall'analisi dei requisiti esposti dal Committente emerge chiaramente la necessità di un sistema che implementi pienamente una soluzione di Business intelligence ovvero un complesso meccanismo in grado di acquisire, organizzare, processare ed elaborare dati (in vari formati) provenienti da fonti eterogenee.

Il rispetto dei requisiti indicati deve necessariamente passare attraverso un percorso fatto di varie FASI, ognuna con un obiettivo preciso ed un output necessario come input per la fase successiva.

Di seguito viene schematizzato il processo che è stato ipotizzato per poter realizzare il sistema di BI per l'USR:



Vengono descritte le singole fasi per focalizzare l'attenzione su tutte le principali attività tecniche e organizzative necessarie alla buona riuscita del progetto e

### 4.1 FASE 1 - Analisi dati e sorgenti informative

Questa fase ha l'obiettivo di censire tutta la mole di dati, elaborati, sorgenti informative, ecc che dovranno confluire all'interno del sistema di BI.

Si ritiene indispensabile un intenso confronto tra gli analisti/esperti BI e i referenti dell'USR al fine di definire completamente il dominio iniziale dei dati che costituirà il datawarehouse su cui si svilupperà il progetto nelle fasi successive.

Oltre all'analisi dei dati e delle sorgenti informative, sarà necessario definire in questa fase tutti gli utenti e i relativi profili di coloro che dovranno, a vario titolo, utilizzare il sistema di BI. Questo serve prevalentemente per organizzare l'accesso alle informazioni e alle funzionalità del sistema nella maniera più confacente al fine di rispettare tutti i vincoli sulla privacy e sicurezza indicati dal Committente.

	<p>Scheda Progetto - SDF Sistema Statistico Regionale Datawarehouse</p>	
---	---	--

L'output della Fase-1 è pertanto costituito da un insieme di documenti che definiscono i dati da acquisire, le modalità con cui tali operazioni saranno effettuate, il modello dei dati con cui i dati devono essere organizzati, i profili che accederanno ai dati e alle funzionalità del sistema.

## 4.2 FASE 2 – Installazione e personalizzazione piattaforma DWH

Questa fase ha l'obiettivo di implementare la piattaforma tecnologica di base che deve ospitare i dati, le funzionalità, i profili, le utenze ecc, ovvero il DWH come sistema nel suo complesso.

E' prevista l'installazione del/i software, la configurazione e la personalizzazione. Quest'ultima verrà realizzata a partire dall'analisi svolta al punto precedente.

Il risultato della fase 2 è costituito da un ambiente tecnologico (hardware/software) predisposto e configurato per gestire tutte le attività successive di integrazione, analisi e reportistica.

## 4.3 FASE 3 – Integrazione fonti e ripresa dati

Una volta predisposto l'ambiente che ospita il DWH si passa alla fase di integrazione e ripresa dei dati.

L'integrazione consiste nel reperire tutte le informazioni di interesse dalla fonte dati da integrare per inserirle all'interno di un modello logico in cui i dati devono essere organizzati. Il modello dei dati è il risultato dell'analisi della Fase I.

Per ogni singola sorgente informativa è necessario procedere secondo il seguente processo:

1. Individuare tutti i dati di interesse, la loro ubicazione, il formato di partenza, la quantità dei dati, le dimensioni fisiche, ecc
2. Definire eventuali meccanismi di pulizia e/o recovery di dati mancanti
3. Definire eventuali procedure di verifica e/o validazione dei dati
4. Definire il modello finale all'interno del quale i dati dovranno essere inseriti
5. Definire la frequenza di aggiornamento dei dati dalla sorgente e la tipologia di popolamento (incrementale, riscrittura, ecc), la storicizzazione dei dati, ecc
6. Definire i metadati della sorgente al grado di granularità che si ritiene opportuno
7. Realizzare tutte le procedure informatiche di estrazione, trasformazione e caricamento dei dati (ETL) secondo le specifiche delineate nella Fase I e nei punti precedenti

Seguire in maniera accurata tutti i passaggi indicati consente di sfruttare al meglio e mantenere sempre costantemente allineati i dati del DWH con quelli provenienti dalle sorgenti informative. Uno dei principali problemi riguardanti l'integrazione è proprio quello di non aver realizzato una procedura di ETL robusta. Eventuali problematiche in questa fase si riflettono sulla qualità e affidabilità dei dati utilizzati per le attività successive.

	<p>Scheda Progetto - SDF Sistema Statistico Regionale Datawarehouse</p>	
---	---	--

I dati che sono presenti come dati storici all'interno dell'USR devono essere acquisiti "una tantum" per consentire la realizzazione di analisi o confronti di tipo statistico.

La gestione dei dati storici deve essere gestita come attività "spot" necessaria cioè unicamente a popolare il DWH di tutti i dati ritenuti di interesse e che già appartengono al patrimonio informativo dell'USR. Per tale tipologia di dati non verranno utilizzate procedure schedate di popolamento del DWH in quanto l'immissione dei dati avverrà tramite la realizzazione di apposite procedure (batch) il cui utilizzo ed obiettivo è mirato esclusivamente all'acquisizione del set individuato.

#### 4.4 FASE 4 - Realizzazione portale, report, dashboard, analisi ed elaborati

Tutto il lavoro indicato e riassunto nelle fasi precedenti trova il suo fine ultimo nella realizzazione degli elaborati che, a partire dai dati immagazzinati nel DWH, estrapolano informazioni attraverso meccanismi più o meno semplici di interpolazione, aggregazione, correlazione, ecc.

E' possibile individuare 4 tipologie di elaborati a partire dai dati "grezzi" del DWH.

##### 4.4.1 Portale

La struttura di LAZIOcrea coinvolta nel progetto, per la realizzazione di siti e portali Internet utilizza tecnologia open-source e adotta sistemi di content management basati su classiche architetture LAMP (Linux, Apache, PHP, MySQL).

Il nuovo sito sarà realizzato utilizzando una piattaforma di content management open source su architettura LAMP (Drupal 8). Il software in oggetto, è oggi utilizzato su vasta scala a livello mondiale per la realizzazione di siti a carattere informativo e di servizi in virtù delle sue caratteristiche di flessibilità, possibilità di personalizzazione e orientamento al mondo mobile.

Tale soluzione è già stata utilizzata per progetti con caratteristiche analoghe e offre le garanzie necessarie alla realizzazione del progetto nella tempistica prevista.

A regime, il personale indicato da Regione Lazio sarà opportunamente formato in modo da poter essere autonomo nella gestione dei contenuti. Gli interventi di manutenzione evolutiva del sito saranno gestiti da LAZIOcrea.

Il sito sarà predisposto sarà di tipo responsive, basato sui linguaggi attualmente in uso per lo sviluppo di siti web (html5 , css3, javascript).

Sarà mantenuto l'attuale sistema di monitoraggio basato su Google Analytics (il codice di tracciamento inglobato in tutte le singole pagine identificate da un titolo univoco). Il sito sarà

	<p>Scheda Progetto - SDF Sistema Statistico Regionale Datawarehouse</p>	
---	---	--

dotato di uno specifico componente per l'ottimizzazione in ottica SEO (Search Engine Optimization).

#### 4.4.2 Report

Il report rappresenta una estrapolazione di dati all'interno del DWH attraverso un sistema di filtri e condizioni. Generalmente il report prevede una struttura predefinita dove è possibile eventualmente impostare una serie di variabili che caratterizzano il particolare report (es intervallo temporale, ...).

Il report viene spesso arricchito attraverso degli elementi grafici come diagrammi a torte, istogrammi, grafici con linee, anelli, bolle, superfici ecc.

Altra caratteristica dei report è che, essendo pensato per un utilizzo da parte di un utente finale o di terze parti, può essere esportato in diversi formati (excel, csv, pdf, ecc) e può essere schedulato per essere prodotto a intervalli di tempo prestabiliti (es ogni primo del mese) e fruito mediante vari metodi (es. accesso e download da un'area riservata, inviato automaticamente per posta elettronica, ecc).

#### 4.4.3 Dashboard

Una dashboard rappresenta un cruscotto di monitoraggio ovvero uno strumento che serve per tenere costantemente sotto controllo determinati indici che sono precalcolati, o calcolati in tempo reale, dal DWH. Generalmente i valori mostrati in una dashboard non hanno bisogno di essere esportati perché l'obiettivo in questo caso non è quello di rendere disponibili queste informazioni a terze parti. Il monitoraggio effettuato mediante una dashboard consente infatti di controllare rapidamente se si verificano delle variazioni significative o vengono superati dei livelli di soglia di alcuni valori caratterizzanti la dimensione o le dimensioni che sono di interesse. La dashboard non è utile per effettuare analisi di dettaglio o per accedere ai singoli valori dei dati. Nella dashboard sono presenti in genere valori che scaturiscono da un insieme anche molto complesso di incroci e elaborazione di dati grezzi.

#### 4.4.4 Analisi dimensionali

Le analisi dimensionali rappresentano uno dei più significativi valori aggiunti del DWH in quanto consentono di analizzare un insieme di metriche (dimensioni) utilizzando delle tecniche che consentono non solo di incrociare tra di loro i dati ma anche di "navigare" all'interno di essi.

Appartengono all'analisi dimensionale le seguenti tecniche di rendering dei dati:

	<p>Scheda Progetto - SDF Sistema Statistico Regionale Datawarehouse</p>	
---	---	--

- Pivoting → è l'operazione di rotazione delle dimensioni di analisi;
- Slicing/Dicing → è l'operazione di estrazione/aggregazione di un subset di informazioni;
- Drill-down → è l'operazione di "esplosione" del dato nelle sue determinanti;
- Drill-across → è l'operazione mediante la quale si naviga attraverso uno stesso livello nell'ambito di una gerarchia;
- Drill-through → è l'operazione simile al drill-down solamente che le informazioni richieste vengono prese direttamente dai sistemi esterni (es. OLTP) che gestiscono il dato di dettaglio.

#### 4.4.5 Elaborati vari

Oltre alle analisi dimensionali e alle varie tipologie di report a disposizione dell'utente, è necessario poter impostare sul sistema la produzione di elaborati vari avendo la possibilità di esportarne il risultato (e/o renderlo disponibile in un apposito spazio web) in un qualunque formato (csv, txt, xml, ecc).

Tutto ciò rende possibile ad esempio poter esportare ad intervalli periodici dei tabulati che possono essere necessari a terze parti o a processi informatici per successive elaborazioni.

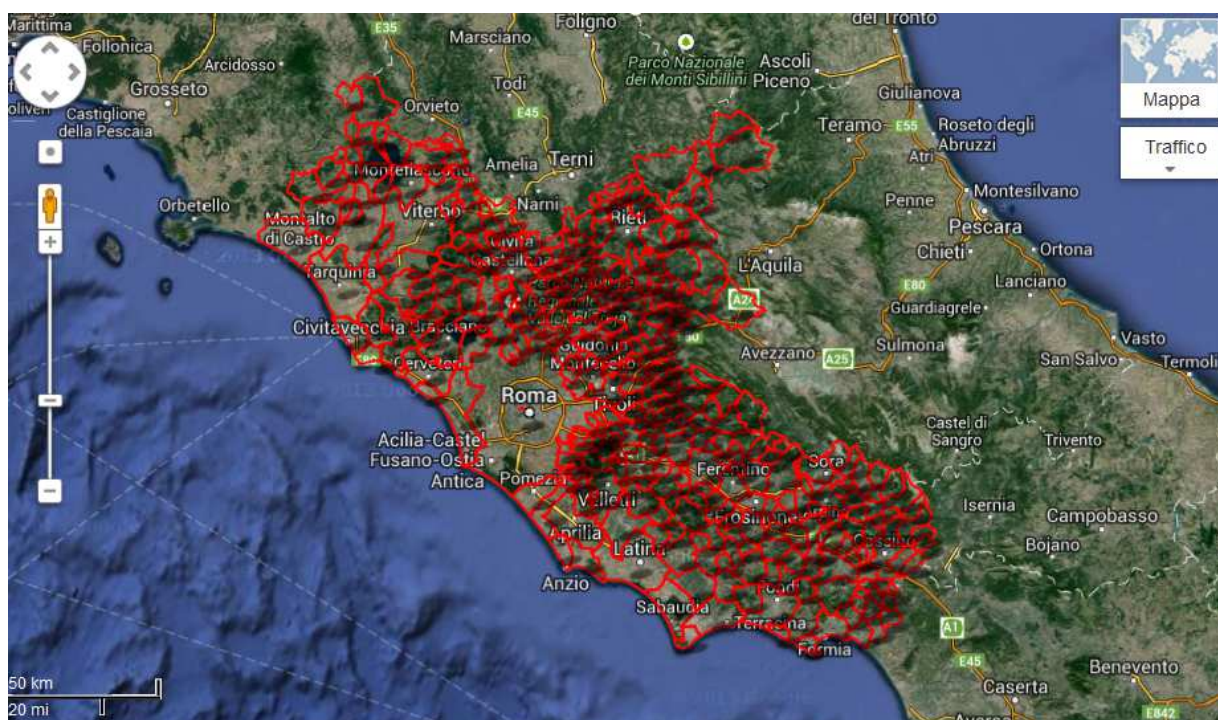
##### 4.4.5.1 Export per sistemi cartografici

Si vuole che alcuni report possano essere redatti in formato compatibile per la cartografia (es. KML - Keyhole Markup Language).

L'obiettivo che si vuole perseguire è quello di visualizzare i dati di uno o più report avvalendosi di sistemi e layer cartografici. Questo è utile per evidenziare la localizzazione delle varie dimensioni e metriche all'interno del report con un'ottica territoriale. L'ausilio di colori, scritte, icone, ecc dovrà consentire di rendere i report più chiari e completi.

Per quanto riguarda la dimensione territoriale si vuole che la granularità sia al massimo a livello comunale (per i comuni medio/piccoli) e a livello di sezioni di censimento o municipi (per le grandi città).





(es:<https://maps.google.it/maps?q=https:%2F%2Fsites.google.com%2Fsite%2Flocalmapping%2FHome%2Fkml%2FLazio.kml%3Fattredirects%3D0&hl=it&geocode=+&t=m&z=8>)

#### 4.5 FASE 5 - Formazione e avvio operativo

Tutte le fasi finora descritte sono state pensate per realizzare e predisporre l'ambiente per l'acquisizione, l'elaborazione e il processamento, a vari livelli, delle fonti dati. Chiaramente il sistema, data la complessità, data la variabilità delle sorgenti informative e data la grande eterogeneità delle informazioni da estrarre e processare ha bisogno di continue attività di gestione. Queste ultime hanno l'obiettivo, da un lato, di mantenere il sistema in perfetta efficienza dal punto di vista operativo, dall'altro di ampliare il numero di informazioni ed elaborazioni che possono essere estratte dal sistema.

Per consentire una corretta e completa gestione operativa sono previste delle sessioni di formazione che hanno l'obiettivo di istruire i vari attori del processo mettendoli nelle condizioni di poter operare in perfetta autonomia sulla piattaforma.

La formazione prevede molteplici aree tematiche e varie sessioni di approfondimento tecnico prevalentemente concentrate su:

- Corso base e avanzato di ETL
- Corso base e avanzato di OLAP
- Corso di amministrazione e gestione della piattaforma

Le sessioni saranno svolte presso i locali della Regione Lazio o i locali messi a disposizione da Laziocrea. Il numero di partecipanti può variare fino ad un massimo di 30 persone/sessione

	<p>Scheda Progetto - SDF Sistema Statistico Regionale Datawarehouse</p>	
---	---	--

per consentire a tutti di poter seguire con il dovuto grado di attenzione e per non avere classi troppo numerose.

#### 4.6 FASE 6 - Manutenzione MAC + MEV

Al fine di garantire la continuità operativa del sistema e offrire l'assistenza sia sulle funzionalità presenti che su eventuali funzionalità evolutive richieste dall'USSR, si prevede di offrire il seguente servizio di assistenza e manutenzione:

TIPOLOGIA	DESCRIZIONE	DURATA
<p><b>MAC</b> (Manutenzione adeguativa e correttiva)</p>	<p>Si verifica quando viene individuato un errore nel software sviluppato e/o quando è necessario adeguare la procedura sviluppata alle variazioni delle condizioni e dell'ambiente operativo circostante per le quali viene utilizzata.</p>	<p><b>18 mesi</b> (dalla messa in produzione)</p>
<p><b>MEV</b> (Manutenzione evolutiva)</p>	<p>Si verifica quando nascono nuove esigenze che possono essere soddisfatte realizzando nuove funzionalità nella medesima procedura sviluppando del software 'ad hoc'</p>	<p><b>50 giorni/uomo</b> (a consumo all'interno dei 18 mesi di MAC)</p>



## 5 Dimensionamento fisico degli ambienti HW

Un aspetto da tenere ben presente nei progetti di DWH è il dimensionamento degli archivi che gestiscono i dati nonché la capacità computazionale delle apparecchiature HW che ospitano i sistemi software. Per il progetto verrà messa a disposizione la seguente infrastruttura:

MACCHINA VIRTUALE APPLICATIVA	DATABASE MACHINE
RAM MAX: <b>8 GB</b>	DB ENGINE: <b>POSTGRES 9.x</b>
SPAZIO DISCO MAX: <b>20 GB</b>	DIMENSIONE UTILE MAX: <b>500 GB</b>

Laziocrea utilizza una infrastruttura virtualizzata con nodi bilanciati e clusterizzati per garantire scalabilità dei servizi nonché per prevenire possibili fault di sistema.

I costi sostenuti per l'utilizzo di una architettura come quella descritta sommariamente nella tabella precedente possono essere assorbiti all'interno della "normale" gestione progettuale e operativa di Laziocrea.

Per esigenze diverse, ovvero che richiedano un maggiore e più intensivo utilizzo di risorse HW, anche in corso d'opera, è necessario prevedere una quota per l'acquisizione e l'ampliamento delle risorse indicate da considerare all'interno del presente studio o all'interno di future evoluzioni.

	<p>Scheda Progetto - SDF Sistema Statistico Regionale Datawarehouse</p>	
---	---	--

## 6 Sorgenti informative da integrare

Definire la quantità, qualità e complessità di ogni sorgente informativa da integrare all'interno del sistema descritto è sicuramente un compito arduo dal momento che non sono note, allo stato attuale, tutte le fonti dati di interesse.

I parametri da tenere bene in considerazione sono:

### QUANTITA'

- Numero di record complessivi da elaborare
- Dimensione fisica dei dati

### QUALITA'

- Pulizia e accuratezza dei dati
- Formato dei dati eterogeneo/omogeneo
- Verifiche e validazioni

### COMPLESSITA'

- Frequenza di aggiornamento dei dati
- Numero di dimensioni/metriche interessate
- Riferimenti interni ed esterni tra dati
- Pre/Post processamento dei dati
- Storicizzazione dei dati

Il presente studio prevede l'analisi ed integrazione di un numero massimo di **n°20** sorgenti informative intese come sorgenti di *media* difficoltà. Qualora l'integrazione delle sorgenti informative risulta molto impegnativa, in considerazione dei parametri indicati, verrà ridotto il numero complessivo di sorgenti dati al fine di poter mantenere l'effort entro dei parametri consoni ed accettabili.

	<p>Scheda Progetto - SDF Sistema Statistico Regionale Datawarehouse</p>	
---	---	--

## 7 Benefici attesi

Lo strumento proposto nel presente studio offre tutte le potenzialità per mettere il personale regionale nelle condizioni di poter gestire una grande mole di dati in maniera rapida e strutturata.

Il beneficio derivante dalla realizzazione di un sistema di Datawarehouse come quello descritto, in genere, ha riscontri nell'immediato ma prevede azioni a medio-lungo termine. Sono moltissimi gli ambiti di intervento cui l'analisi dei dati da fonti dati eterogenee può rivelarsi un ottimo strumento soprattutto nel campo della PA e, in particolare, della Regione Lazio.

Tenere sotto controllo voci di spesa, incrociare dati per verificare possibili sprechi e/o inefficienze, creare e monitorare degli indicatori è la chiave per una PA efficiente. Chiaramente la piattaforma rappresenta solamente uno strumento in grado di mettere nelle condizioni il management di optare per una soluzione o per un'altra sulla base di dati reali, opportunamente elaborati. Il supporto alle decisioni realizzato in questo modo evita e mette al riparo da "salti nel buio", ovvero da provvedimenti che si rivelano peggiore dell'anomalia che dovrebbero prevenire.

## 8 Cronoprogramma di massima delle attività

Di seguito un cronoprogramma di massima delle attività descritte:

ATTIVITA'	TEMPISTICA
FASE 1 - Analisi dati e sorgenti informative	3 mesi
FASE 2 - Installazione e personalizzazione piattaforma DWH	1 mese
FASE 3 - Integrazione fonti e ripresa dati	3 mesi
FASE 4 - Realizzazione report, dashboard, analisi ed elaborati	3 mesi
FASE 5 - Formazione e avvio operativo	0,5 mese
FASE 6 - Manutenzione MAC + MEV	12 mesi

## 9 Analisi dei costi

Il presente paragrafo ha lo scopo di rappresentare i costi del progetto e fornire gli elementi per analizzare la congruità economica delle tariffe applicate, ai sensi dell'art. 192, comma 2, del D.Lgs. n. 50/2016 recante il "Regime speciale degli affidamenti in house".

Le attività di coordinamento del progetto e di gestione degli approvvigionamenti (dall'avvio delle procedure di acquisizione al governo delle attività contrattualizzate) saranno svolte da personale interno LAZIOcrea, i cui costi - elaborati sulla base di quanto indicato al paragrafo 3) dell'allegato B) al POA 2018 - rientrano nel costo complessivo di funzionamento della Società, remunerato ai sensi dell'art. 7 del Contratto Quadro di Servizi approvato con D.G.R. n. 891 del 19/12/2017.

In riferimento ai servizi da acquisire sul mercato, al fine di favorire la valutazione della congruità dei costi, è stata effettuata un'analisi comparativa analizzando e raffrontando i costi dei servizi e le tariffe applicate nel presente studio con i costi applicati nelle procedure di gara per l'acquisizione di servizi simili e/o per affidamenti diretti di servizi comparabili a quelli offerti, anche se non completamente sovrapponibili. Le figure professionali ritenute necessarie per l'espletamento delle attività dello studio e l'effort stimato (in termini di giornate-uomo) per ciascun profilo professionale, sono il risultato dell'esperienza maturata dalla scrivente Società nell'implementazione e gestione di sistemi informativi.

I costi complessivi del progetto, e i razionali utilizzati per la loro definizione, sono mostrati nel seguito con particolare riferimento a:

- Prospetto 1 - Costi per l'analisi, progettazione, sviluppo e messa in esercizio del sistema (sviluppi a corpo);
- Prospetto 2 - Costi della manutenzione evolutiva (MEV);
- Prospetto 3 - Sintesi dei costi per l'acquisto di servizi ripartiti per annualità e natura della spesa, distinguendo tra spesa in conto corrente e spese in conto capitale

Per quanto attiene ai servizi professionali, le figure professionali e le tariffe unitarie prese a riferimento sono quelle indicate al punto 4) dell'Allegato B) del POA 2018. In particolare, si è scelto di riferirsi alle tariffe della gara Consip "Disciplinare di gara a procedura aperta, in tre lotti, per l'affidamento dei servizi informatici di sviluppo, manutenzione evolutiva, correttiva ed adeguativa, supporto specialistico e supporto al ridisegno dei processi nell'ambito dei sistemi informativi del dipartimento del tesoro".

Nella tabella sottostante, applicando le suddette tariffe, si rappresenta il costo dei **servizi professionali** volti all'analisi, progettazione, sviluppo e messa in esercizio delle funzionalità

	<p>Scheda Progetto - SDF Sistema Statistico Regionale Datawarehouse</p>				<p>REGIONE LAZIO</p>	
---	---	---	--	---	--------------------------	---

descritte (sviluppi a corpo), al netto delle attività di manutenzione evolutiva, per cui si rimanda alla specifica tabella:

Prospetto n.1			
ANALISI, PROGETTAZIONE, SVILUPPO E MESSA IN ESERCIZIO DEL SISTEMA (SVILUPPI A CORPO)			
FIGURE PROFESSIONALI	Tariffa unitaria in euro per giorno persona (oltre IVA)	GG/UU	Totale
Capo Progetto	€ 409,79	80	€ 32.783,20
Analista Funzionale	€ 327,74	190	€ 62.270,60
Analista Programmatore	€ 217,21	220	€ 47.786,20
Programmatore	€ 188,40	280	€ 52.752,00
Specialista di prodotto/tecnologia	€ 327,74	140	€ 45.883,60
	<b>Totale</b>	<b>910</b>	<b>€ 241.475,60</b>
	<b>Iva</b>		<b>€ 53.124,63</b>
	<b>Totale IVA inclusa</b>		<b>€ 294.600,23</b>

Riguardo alla MEV, nella tabella che segue vengono mostrati le figure professionali, le giornate-uomo ritenute necessarie e i relativi costi. Per quanto attiene alle tariffe unitarie utilizzate, sono state prese a riferimento le medesime tariffe utilizzate per la compilazione del prospetto riguardante le attività di progettazione, sviluppo e messa in esercizio del sistema, rispetto alle figure professionali per le quali è prevista l'erogazione del servizio di MEV.

Prospetto n.2			
MANUTENZIONE EVOLUTIVA (MEV)			
FIGURE PROFESSIONALI	Tariffa unitaria in euro per giorno persona (oltre IVA)	GG/UU	Totale
Capo Progetto	€ 409,79	0	€ 0,00
Analista Funzionale	€ 327,74	15	€ 4.916,10
Analista Programmatore	€ 217,21	60	€ 13.032,60
Programmatore	€ 188,40	120	€ 22.608,00
Specialista di prodotto/tecnologia	€ 327,74	20	€ 6.554,80
	<b>Totale</b>	<b>215</b>	<b>€ 47.111,50</b>
	<b>Iva</b>		<b>€ 10.364,53</b>
	<b>Totale IVA inclusa</b>		<b>€ 57.476,03</b>

Nel prospetto successivo viene mostrata la sintesi, ripartita tra le varie annualità, dei costi per l'acquisto di servizi ripartiti per annualità e natura della spesa, distinguendo tra spesa in conto corrente e spese in conto capitale.

	Scheda Progetto - SDF Sistema Statistico Regionale Datawarehouse				REGIONE LAZIO	
---	--	---	--	---	------------------	---

Prospetto n.3			
COSTI TOTALI DEI SERVIZI			
	Totale	2019	2020
<b>SPESE GARA O ADESIONE A CONVENZIONE CONSIPI</b>			
Adesione a convenzione		2.000,00 €	
<b>SERVIZI ICT E CONSULENZE</b>			
Gestione e manutenzione applicazioni	- €	- €	- €
Sviluppo software e manutenzione evolutiva (MEV)	€ 288.587,10	€ 202.010,97	€ 86.576,13
<b>TOTALE SERVIZI ICT E CONSULENZE</b>	<b>€ 288.587,10</b>		
<b>Totale imponibile</b>	<b>€ 290.587,10</b>	€ 204.010,97	€ 86.576,13
Iva (22%)	€ 63.929,16	€ 44.882,41	€ 19.046,75
<b>Totale IVA inclusa</b>	<b>€ 354.516,26</b>	<b>€ 248.893,38</b>	<b>€ 105.622,88</b>

Segue la tabella di riepilogo dei costi di progetto con la ripartizione degli stessi sulle annualità che caratterizzano l'intervento.

	Totale intervento IVA inclusa	Ripartizione per annualità IVA inclusa	
		2019	2020
SPESE CORRENTI	€ 2.440,00	€ 2.440,00	€ 0,00
CONTO CAPITALE	€ 352.076,26	€ 246.453,38	€ 105.622,88
<b>TOTALE</b>	<b>€ 354.516,26</b>	<b>€ 248.893,38</b>	<b>€ 105.622,88</b>