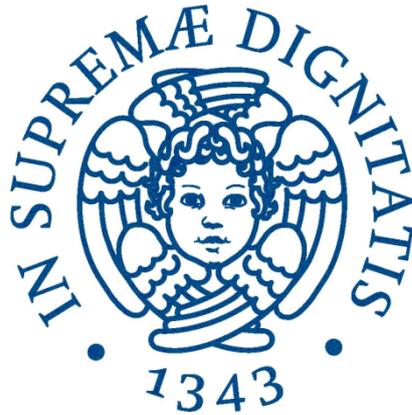


Università di Pisa



Corso di Laurea Specialistica in

Ingegneria Informatica per la Gestione d'Azienda

GAD (Grafo Attributi Dimensionali): un modello concettuale di ausilio alla gestione dei processi del DW RGS

Relatore - Prof. **Roberto Chiavaccini**

Correlatore - Prof. **Enzo Mingozzi**

Relatore esterno - Ing. **Gianluca Ricco**

Candidato - Dott. **Francesco Piroddu**

a.a. 2011/2012

Dettagli contatto:

Francesco Piroddu

email: francescopiroddu@hotmail.com

mobile: +39.348.9333686

Ringraziamenti

Ringrazio la mia famiglia

Ringrazio i miei amici più vicini

Ringrazio tutti coloro che hanno preso parte al mio percorso universitario

Ringrazio tutti coloro che, ad ogni modo, sono stati presenti ed hanno reso possibile la stesura di questo testo, frutto di un obiettivo per il quale ci sono voluti anni di studio.

Sono stati anni importanti: saranno importanti anche quelli futuri

Sommario

Abstract 9

Prefazione 11

Il Data Warehouse ed i Sistemi Informativi Direzionali della Ragioneria Generale dello Stato 15

- 1 Il contesto 16
 - 1.1 Finalità ed obiettivi della Ragioneria Generale dello Stato 16
 - 1.2 I sistemi a supporto dei processi 19
 - 1.2.1 I sistemi informativi gestionali 19
 - 1.2.2 L'evoluzione delle esigenze informative: i sistemi conoscitivi 20
 - 1.2.3 Il sistema informativo per il monitoraggio della spesa pubblica 21
- 2 La strategia 22
 - 2.1 Soluzioni di Business Intelligence 22
 - 2.1.1 Le esigenze "conoscitive" della RGS 22
 - 2.1.2 Tipologie di soluzioni di BI 23
 - 2.1.3 Introduzione al Data Warehouse 24
 - 2.1.4 Introduzione al Sistema Direzionale 27
 - 2.2 Le scelte adottate per la RGS 29
 - 2.2.1 La soluzione di BI della RGS 29
 - 2.2.2 Architettura delle soluzioni di BI della RGS 30
 - 2.2.3 Modelli concettuali dell'informazione 32
 - 2.2.4 Metadati 35
 - 2.2.5 Riconoscimenti della validità della soluzione 36
- 3 La soluzione 37
 - 3.1 Il Data Warehouse RGS 37
 - 3.1.1 Architettura logica e tecnologica 41
 - 3.2 Il Sistema Informativo Direzionale 44
 - 3.2.1 Architettura tecnologica 45
 - 3.2.2 Fonti dati 46
 - 3.2.3 Architettura applicativa 47
 - 3.3 L'Organizzazione 48

- 3.4 I processi di governo e di sviluppo 52
- 4 I risultati 54
 - 4.1 L'informazione nel processo decisionale 54
 - 4.2 Il governo dell'informazione 57
 - 4.2.1 La metodologia per l'individuazione degli indicatori 57
 - 4.2.2 Le fonti informative 59
 - 4.2.3 La rappresentazioni delle informazioni 60
 - 4.3 I benefici 62
- 5 Le prospettive 64

Il Dimensional Fact Model 67

- 1 Concetti di base 70
 - 1.1 Fatto 72
 - 1.2 Misura 72
 - 1.3 Dimensione 72
 - 1.4 Evento primario 73
 - 1.5 Attributo dimensionale 73
 - 1.6 Gerarchia 73
 - 1.7 Evento secondario 74
- 2 Modellazione avanzata 74
 - 2.1 Attributi descrittivi 76
 - 2.2 Gerarchie condivise 76
 - 2.3 Archi multipli 77
 - 2.4 Gerarchie incomplete 77
 - 2.5 Dinamicità 78
 - 2.6 Additività 79
- 3 DFMCASE 80
- 4 Sovrapposizione di schemi di fatto 82
- 5 Dominio dei dati trattati 85

L'applicazione: GAD 87

- 1 Specifica funzionale 88
 - 1.1 Scopo del progetto 88
 - 1.1.1 Descrizione del progetto 88
 - 1.1.2 Tipologie di utenti 88

- 1.1.3 Glossario definizioni ed acronimi *89*
- 1.2 Scelta della soluzione *90*
 - 1.2.1 Premessa *90*
 - 1.2.2 Architettura del sistema *90*
 - 1.2.3 Vincoli *94*
- 1.3 Specifica dei requisiti *94*
 - 1.3.1 Requisiti funzionali *94*
 - 1.3.2 Requisiti non funzionali *97*
 - 1.3.3 Diagramma generale dei casi d'uso *99*
 - 1.3.4 Casi d'uso *100*
 - 1.3.5 Diagramma di sequenza *114*
 - 1.3.6 Diagramma delle classi *116*
 - 1.3.7 Diagramma di Gantt *117*
- 2 Manuale utente *118*
 - 2.1 Schermata Elaborazioni *119*
 - 2.1.1 Schermata Vista generale *120*
 - 2.1.2 Schermata Ricerca avanzata *121*
 - 2.1.3 Schermata Grafo principale *122*
 - 2.1.4 Schermata Dettaglio elemento *123*
 - 2.2 Schermata Visualizzazione query *127*
- 3 Release future *128*
 - 3.1 Modifica della struttura del DW RGS *128*
 - 3.2 Query ontologiche *128*

Conclusioni *133*

Bibliografia *135*

Abstract

Il presente scritto è il risultato dell'attività di analisi e progettazione nell'ambito del progetto GAD (Grafo Attributi Dimensionali) svolto dalla Consip S.p.A per il committente rappresentato dal Ministero dell'Economia e delle Finanze.

L'obiettivo del progetto sarà la realizzazione di un'applicazione web-based che andrà a soddisfare le esigenze di due differenti tipologie di utenti:

- Utente di business: permette a dirigenti e funzionari della Pubblica Amministrazione e della Ragioneria Generale dello Stato (RGS) di navigare (sfruttando il paradigma del modello dimensionale) i metadati relativi alle misure e alle dimensioni di analisi oggetto di interesse per l'amministrazione. La visione integrata ed omogenea dei concetti di interesse e delle relazioni esistenti tra esse sarà di ausilio all'utente nella comprensione della tematica di interesse e nella individuazione delle informazioni necessarie a supportare i processi decisionali;
- Utente tecnico: costituirà uno strumento di ausilio per i capi progetto Consip e i fornitori responsabili dello sviluppo nelle diverse fasi del ciclo di vita del software:
 - progettazione concettuale dei Data Mart del DW RGS;
 - progettazione logica dei Data Mart del DW RGS;
 - progettazione dello strato semantico a supporto dell'utente di business.

La visione integrata fornita dal modello GAD renderà inoltre più efficiente la gestione del sistema DW RGS.

L'applicazione attingerà le informazioni di interesse attraverso un ulteriore livello software, il quale si occuperà di rappresentare la realtà del DW RGS secondo il modello concettuale Dimensional Fact Model (DFM).

Si descriverà inizialmente la situazione attuale in termini di sistemi esistenti e funzionalità offerte agli utenti. Si procederà descrivendo il processo di sviluppo e le attività di progettazione effettuate con ampio riferimento ai sistemi coinvolti.

Prefazione

In questa sezione si descriveranno le macrocomponenti del sistema in esame, riportate in Figura 1, le quali saranno oggetto di analisi più approfondite nei capitoli successivi.

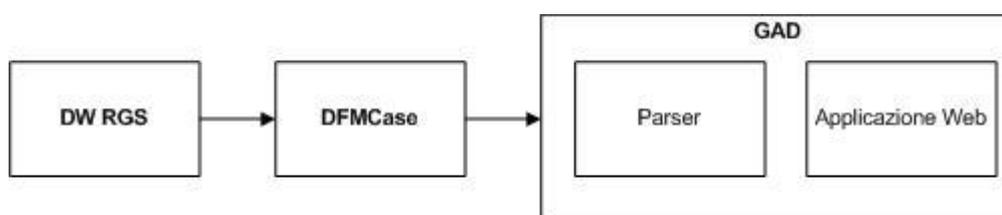


Figura 1

Data Warehouse RGS.

L'architettura per il data warehousing utilizzata, riportata in Figura 2, è costituita di 4 livelli:

- *Livello delle sorgenti.* Il DW RGS utilizza fonti di dati eterogenei provenienti dai vari sistemi informativi della RGS.
- *Livello dell'alimentazione.* I dati memorizzati nelle sorgenti devono essere estratti, ripuliti per eliminare le inconsistenze e completare eventuali parti mancanti, integrati per fondere sorgenti eterogenee secondo uno schema comune. I cosiddetti strumenti ETL (Extraction, Transformation, Loading) permettono di integrare schemi eterogenei, nonché di estrarre, trasformare, ripulire, validare, filtrare e caricare i dati dalle sorgenti nel DW.
- *Livello del warehouse.* Le informazioni vengono raccolte in un singolo "contenitore" centralizzato logicamente: il DW RGS. Esso viene usato come sorgente per costruire i vari data mart, che ne costituiscono una parziale replica, orientati verso specifiche aree della RGS. Accanto al DW, il contenitore dei metadati mantiene informazioni sulle sorgenti, sui meccanismi di accesso, sulle procedure di pulizia e alimentazione, sugli utenti, sugli schemi dei data mart.

Prefazione

- *Livello di analisi.* Permette la consultazione efficiente e flessibile dei dati integrati a fini di stesura di report, di analisi, di simulazione.

Un ulteriore livello, *Livello dei dati riconciliati*, detto anche operational data store, materializza i dati operazionali ottenuti a valle del processo di integrazione e ripulitura dei dati sorgente (si trova dunque nel mezzo della procedura ETL, fra la procedura di *transformation* e quella di *loading*). Esso crea un modello di dati comune di riferimento per l'implementazione del data warehouse, introducendo al contempo una separazione netta tra le problematiche legate all'estrazione e integrazione dei dati dalle sorgenti e quelle inerenti l'alimentazione del DW.

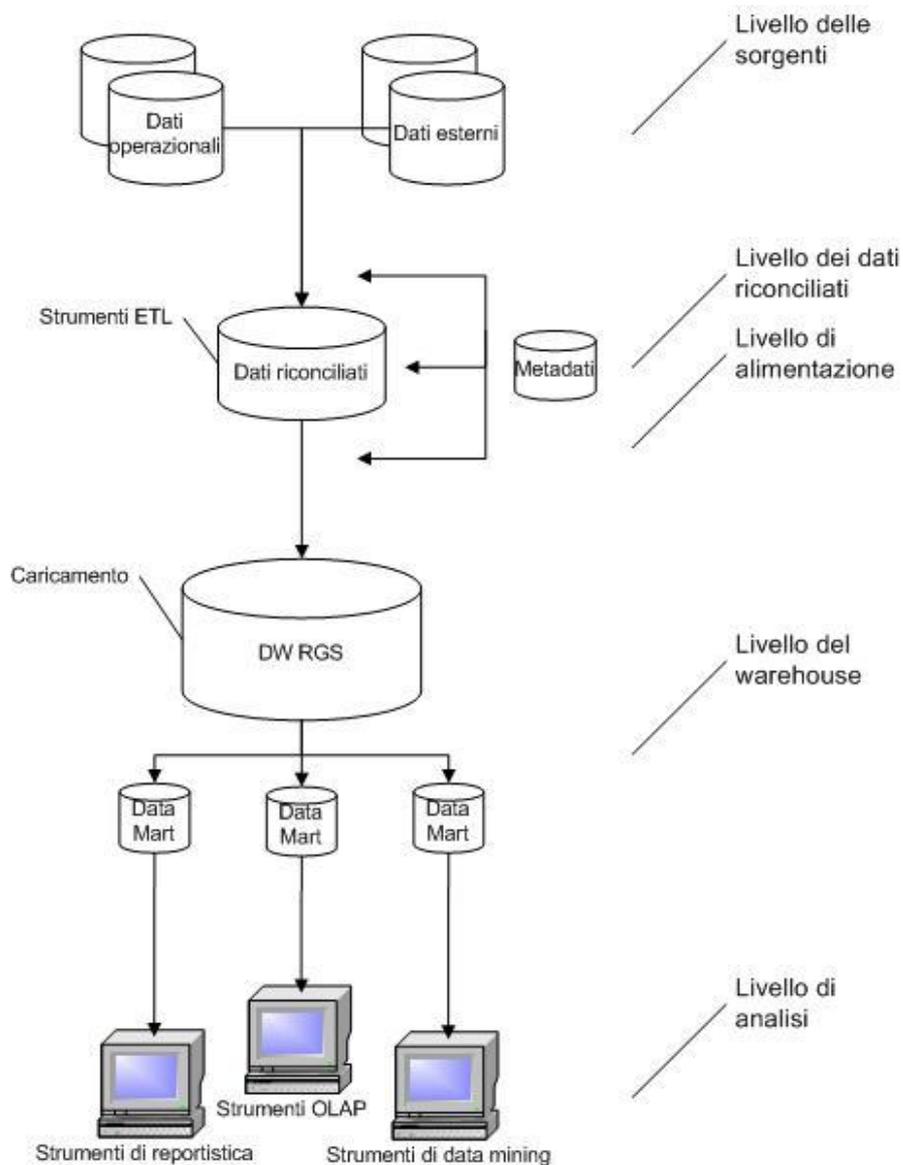


Figura 2

Prefazione

Nel caso in esame, il DW RGS contiene le informazioni relative al dipartimento della RGS (Ragioneria Generale dello Stato), afferente al Ministero dell'Economia e delle Finanze.

Si esamineranno nel dettaglio le caratteristiche dell'architettura su esposte nel Capitolo 1, Il Data Warehouse ed i Sistemi Informativi Direzionali della Ragioneria Generale dello Stato.

DFMCase.

Il DFMCase è un software, realizzato dalla Consip S.p.A., che permette di disegnare lo schema concettuale di un data warehouse secondo il DFM (Dimensional Fact Model). Tramite questo strumento sono stati disegnati gli schemi concettuali di ognuno dei data mart componenti il DW RGS.

Tale software è ad oggi utilizzato dai progettisti dei data mart, addetti alla manutenzione ed alla realizzazione delle evoluzioni degli stessi, per la consultazione degli schemi concettuali dei data mart.

Il software elabora file di tipo *dfmx*. Esso è un file XML avente una grammatica specifica per la rappresentazione di un insieme di schemi di fatto. Il DFMCase permette l'esportazione degli schemi disegnati nello stesso formato su riportato, il *dfmx*. Nella particolare istanza del DW RGS, chiameremo tale fonte dati *dwrqs.dfm*. Essa conterrà l'elenco degli schemi concettuali associati ad ogni data mart.

Si descriverà nel dettaglio il formalismo del DFM nel Capitolo 2, Il Dimensional Fact Model.

GAD.

L'idea che sta alla base del progetto descritto in questo testo è unificare i concetti presenti sui diversi data mart in un unico modello consultabile attraverso l'**Applicazione Web**.

Tale applicazione renderà interattiva la rappresentazione, secondo un'opportuna rivisitazione del modello concettuale DFM, la struttura del DW RGS.

L'applicazione attingerà le informazioni di interesse contenute nel *dwrqs.dfm* attraverso il modulo **Parser**, il quale unificherà le informazioni dei diversi schemi concettuali, ottenendo un

Prefazione

unico nuovo modello dei dati, adatto a rappresentare la realtà che l'Applicazione Web dovrà rappresentare.

Si esaminerà nel dettaglio la struttura di tale macrocomponente nel Capitolo 3, L'applicazione: GAD.

Il Data Warehouse ed i Sistemi Informativi Direzionali della Ragioneria Generale dello Stato

La complessità sempre crescente del contesto in cui la Pubblica Amministrazione è chiamata a muoversi in questo inizio del terzo millennio richiede una crescita continua di strumenti: legislativi, amministrativi, organizzativi e informativi, essenziali per le attività di governo e di indirizzo.

La Ragioneria Generale dello Stato svolge una missione centrale nella programmazione e gestione economico-finanziaria dello Stato, presidiando importanti processi quali ad esempio: la predisposizione e la gestione del Bilancio e dei conti pubblici, il supporto agli organi di Governo nella valutazione degli interventi legislativi, il monitoraggio degli interventi comunitari e la vigilanza e controllo nella gestione delle risorse pubbliche.

A fronte di compiti così vasti e articolati, la RGS si è dotata già da tempo di adeguati sistemi informatici, che si sono man mano specializzati ed evoluti, e che rappresentano oggi un supporto indispensabile per le attività svolte nell'ambito dei processi istituzionali. Su questa base, la sempre maggiore consapevolezza dei fenomeni, unita al potenziale conoscitivo reso disponibile dall'accumulo di un'ampissima banca dati, che raccoglie informazioni storiche, correnti e previsionali, offre l'opportunità per compiere un ulteriore salto di qualità e supportare l'evoluzione e il potenziamento dei metodi di analisi dei fenomeni e di indirizzo delle scelte.

In occasione di due grandi eventi che hanno richiesto interventi di aggiornamento delle banche dati, l'avvento dell'anno 2000 e l'adozione dell'euro, nacque l'idea di riorganizzare il patrimonio conoscitivo in maniera sistematica, avviando un percorso di integrazione e omologazione dei dati provenienti dalle varie fonti, fino alla nascita del Data Warehouse RGS avvenuta nel 2002. RGS e Consip hanno perseguito con convinzione questa scelta, e gli sforzi e la tenacia con cui hanno lavorato in questi anni permettono oggi di registrare il raggiungimento di brillanti risultati, tangibili quotidianamente sia dagli uffici che dai vertici dell'Amministrazione.

Tale sistema, nel suo complesso di processi, metodologie, organizzazione e strumenti informatici, ha delle potenzialità tali da rendere RGS pronta a raccogliere nuove sfide, prima fra tutte, oggi alla massima attenzione del Governo e dell'opinione pubblica, quella relativa al monitoraggio tempestivo ed efficace dell'andamento della spesa pubblica.

1 Il contesto

1.1 Finalità ed obiettivi della Ragioneria Generale dello Stato

La Ragioneria Generale dello Stato (RGS) è il dipartimento del Ministero dell'Economia e della Finanze (MEF) che supporta il Parlamento ed il Governo nelle politiche, nei processi e negli adempimenti di bilancio.

La **vision**, cioè la sua funzione istituzionale della RGS, è la rigorosa programmazione e la corretta gestione delle risorse pubbliche attraverso attività e strumenti di monitoraggio, verifica e analisi degli andamenti della spesa pubblica.

La **mission**, cioè gli obiettivi istituzionali della RGS, si articolano in funzione delle seguenti linee guida:

- assicurare un puntuale ed efficace supporto al Parlamento e al Governo nel processo legislativo e nell'attuazione delle leggi di bilancio;
- vigilare sulla coerenza della finanza pubblica con gli obiettivi previsti dal Patto di Stabilità e con gli obiettivi di crescita interna;
- garantire il continuo miglioramento dei sistemi e delle metodologie di monitoraggio, analisi e certificazione dei conti pubblici;
- supportare la Pubblica Amministrazione con un'efficiente rete di presidi territoriali.

I **valori**, i quali definiscono "l'etica di comportamento" dell'istituzione e identificano la strada maestra, le linee guida da seguire per perseguire la mission, sono:

- Spirito di appartenenza, collaborazione e servizio.
- Capacità di gestire il cambiamento.

- Rigore ed eccellenza professionale.
- Responsabilità e affidabilità.
- Efficienza, centralità degli obiettivi, tensione ai risultati.
- Integrità, indipendenza e riservatezza.



Figura 3

L'Istituto si configura come un'organizzazione complessa composta di una struttura centrale e un'articolazione territoriale (si veda Figura 4).

In particolare la struttura centrale è costituita da 10 Ispettorati Generali (nove dei quali svolgono funzioni operative sul controllo dei conti pubblici, uno è adibito alla gestione del personale ed il rimanente è il Centro di documentazione e studi) mentre la parte territoriale comprende gli Uffici Centrali di Bilancio (UCB) e le Ragionerie Provinciali (o Territoriali) dello Stato (RPS).

Gli UCB, dislocati presso i Ministeri con portafoglio, sono snodi essenziali della RGS presso le amministrazioni centrali, nei confronti delle quali svolgono azioni di supporto e vigilanza, nonché di coordinamento, finalizzate al corretto andamento della gestione finanziaria e all'applicazione delle linee di bilancio.

Le RPS hanno sede in ciascuno dei 103 capoluoghi di Provincia d'Italia e sono il terminale della rete di rapporti della RGS con le amministrazioni decentrate. Esplicano, oltre il controllo preventivo e successivo sugli atti in conformità alla vigente normativa che ne delinea i compiti istituzionali, anche funzioni di monitoraggio sui dati finanziari e sui processi amministrativi delle amministrazioni statali con competenza territoriale e sono di supporto alle amministrazioni stesse nella corretta applicazione delle norme.

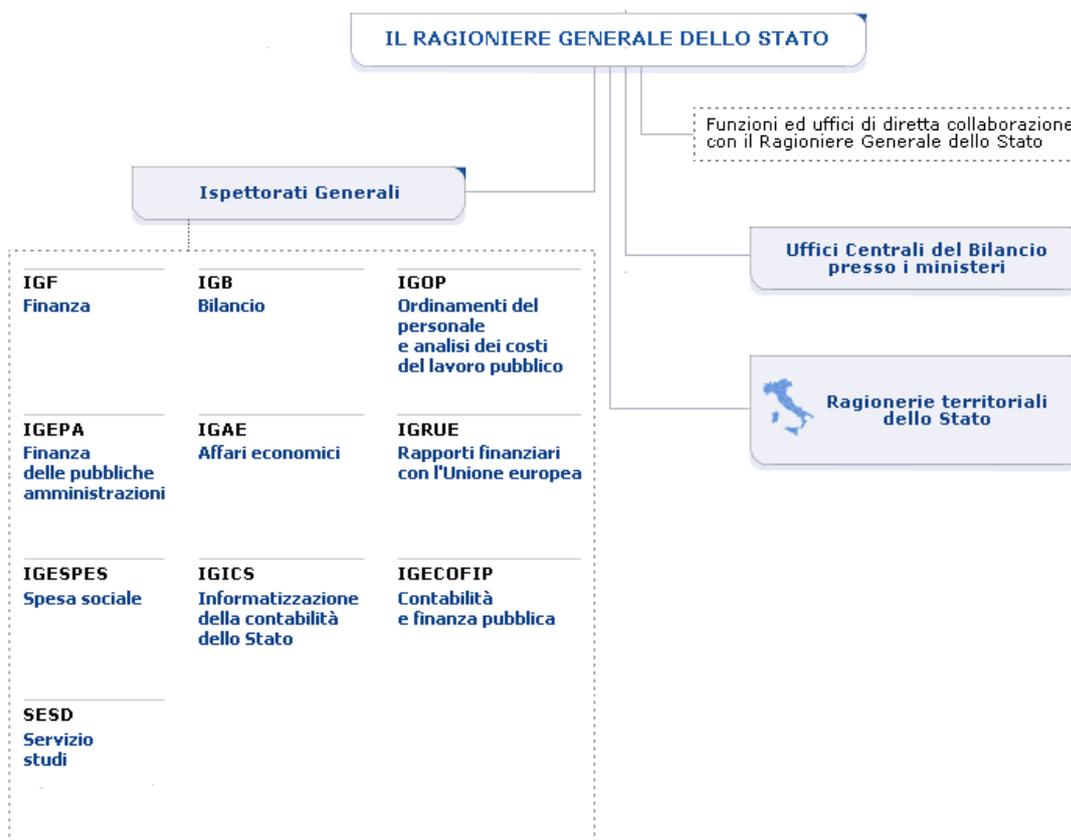


Figura 4

Le funzioni istituzionali della RGS incidono su tutto il processo di gestione della finanza pubblica: dalla programmazione della spesa fino alla verifica del suo andamento effettivo.

La valutazione dell'impatto finanziario dei provvedimenti elaborati dal Governo e dal Parlamento e la relativa verifica di coerenza con gli obiettivi di finanza pubblica sono le principali funzioni che coinvolgono trasversalmente tutti gli Ispettorati cui è richiesta l'elaborazione di pareri sulla fattibilità e rilevanza economico-finanziaria di nuove leggi e provvedimenti nonché la verifica delle relazioni tecniche sui singoli provvedimenti di spesa.

I processi istituzionali e di supporto al funzionamento svolti dalla RGS sono sinteticamente rappresentabili attraverso la catena del valore riportata in Figura 5.

La catena del valore



Figura 5

1.2 I sistemi a supporto dei processi

1.2.1 I sistemi informativi gestionali

La RGS si è dotata, in un quadro organico di investimenti iniziati a partire dalla fine degli anni '70, di sistemi informatici atti a supportare i principali processi propri della sua missione istituzionale.

Oggi la gestione del Bilancio dello Stato è completamente informatizzata, dalla formazione all'assestamento, alla rendicontazione; le manovre di legge finanziaria, compresi gli emendamenti, sono supportate interamente da sistemi informatici che ne recepiscono in tempo reale le variazioni; tutti i pagamenti erogati dalla PA centrale avvengono per via telematica; l'iter di prelegislativa, per la verifica della congruità degli oneri e delle coperture finanziarie delle proposte di legge è interamente supportato da un sistema di workflow informatico.

Importanti progetti di automazione sono poi in corso sul fronte del settore pubblico allargato. Insieme a tante altre applicazioni, questi sistemi vanno a comporre il Sistema Informativo della RGS (SIRGS), che si è evoluto e perfezionato fino a costituire un patrimonio informativo di grande valore, da indirizzare ed integrare ai fini del monitoraggio della finanza pubblica.

1.2.2 L'evoluzione delle esigenze informative: i sistemi conoscitivi

Nel tempo sono emerse, e sono di giorno in giorno più pressanti, nuove esigenze informative, dettate dalla necessità di reagire prontamente e con efficacia ai cambiamenti indotti dal continuo mutamento del contesto sociale, economico e politico. La scena internazionale, l'economia globale, l'Unione Europea e molti altri fattori influenzano e aggiungono ulteriori elementi di complessità all'azione di governo.

Il tempo gioca ormai un ruolo fondamentale. Occorre decidere prontamente, e sempre più spesso le questioni che si pongono sono nuove, estemporanee, mai emerse in passato. Gli istituti e gli strumenti conoscitivi prodotti per legge (es. gli atti dovuti) hanno una data periodicità, e sono espressamente predisposti per rispondere a quesiti già noti, rappresentano perciò una routine che da sola non è più sufficiente a soddisfare le esigenze informative dell'Amministrazione.

Tra i fattori che influenzano le decisioni gioca un ruolo fondamentale la quantità e la qualità delle informazioni di cui si dispone, il che significa affidabilità, accuratezza, completezza dei dati, ma anche e sempre di più tempestività, freschezza. In assenza di strumenti informativi evoluti, per comporre un quadro di sintesi a fronte di un'esigenza conoscitiva estemporanea, bisogna effettuare una serie di passi: ricerca delle fonti, verifiche, integrazioni, ricicli. Spesso il risultato è comunque insoddisfacente per una bassa qualità dell'informazione e per i lunghi tempi di lavorazione.

Occorre superare la vista settoriale e parcellizzata dell'informazione, le ridondanze e le discrepanze che si possono riscontrare tra diverse fonti che si sovrappongono nel trattare una stessa tipologia di informazione, i ritardi che fatalmente si verificano per la ricerca, il reperimento e l'integrazione di dati spesso non correlati ed incoerenti tra loro.

Bisogna giocare d'anticipo, mettere cioè in comune, per ogni eventuale esigenza conoscitiva che si dovesse presentare, i dati che si originano da diverse fonti, organizzandoli in maniera ordinata e coerente, fissando a priori semantica, proprietà e responsabilità degli stessi, stabilendone regole di qualità e di certificazione. Questa razionalizzazione induce un cambiamento nell'organizzazione e nei comportamenti, ed è una condizione fondamentale per preservare e valorizzare il patrimonio informativo che RGS possiede.

1.2.3 Il sistema informativo per il monitoraggio della spesa pubblica

Un ambito fortemente esemplificativo delle nuove esigenze informative è rappresentato dal monitoraggio della spesa pubblica per il quale, a partire dalla fine dell'anno 2000, RGS e Consip hanno avviato una serie di iniziative volte a favorire il miglioramento della quantità, qualità e tempestività delle informazioni raccolte.

Il rispetto del Patto di stabilità, nella sua espressione sia interna che esterna, la progressiva estensione del federalismo con l'attribuzione agli Enti a finanza decentrata di maggiori funzioni e risorse proprie nonché la prospettiva del superamento della tesoreria unica, hanno rafforzato in RGS la necessità di disporre a livello centrale di tutte quelle informazioni su cui fosse possibile attuare un adeguato monitoraggio dei conti pubblici nel pieno rispetto dell'autonomia riconosciuta dalla legge ai vari Enti.

Sono stati avviati interventi miranti alla realizzazione di un sistema informativo integrato funzionale alle esigenze del controllo della finanza pubblica che, attraverso la razionalizzazione ed omogeneizzazione dei flussi informativi relativi al Bilancio dello Stato e degli Enti della PA e tramite il consolidamento dei flussi informativi relativi alla Tesoreria, integrasse le informazioni in modo coerente ed omogeneo e fornisse gli strumenti conoscitivi adeguati a supportare i decisori in tema di finanza pubblica.

L'architettura concettuale del sistema, sintetizzata in Figura 6, si basa principalmente sull'introduzione di strumenti e metodologie Data Warehouse.

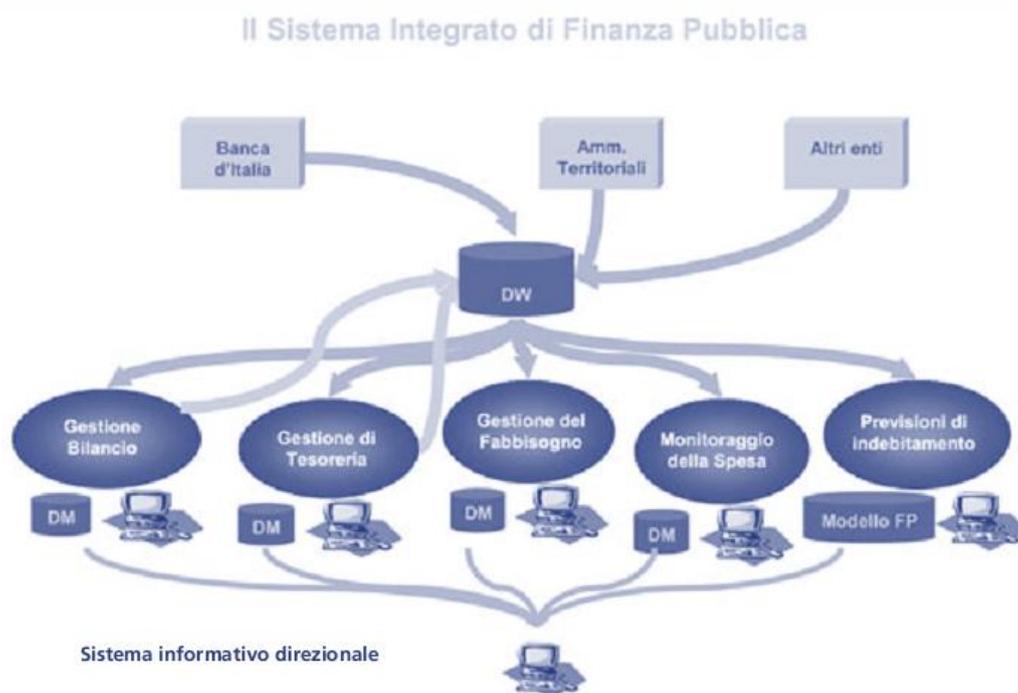


Figura 6

2 La strategia

2.1 Soluzioni di Business Intelligence

2.1.1 Le esigenze “conoscitive” della RGS

L’evoluzione legislativa ed il decentramento della spesa, l’economia globale, l’Unione Europea e molti altri fattori influenzano e aggiungono elementi di complessità sempre crescenti nello svolgimento dei compiti istituzionali della RGS. In questo complesso scenario, teatro di questioni spesso nuove ed estemporanee non si può prescindere dalla disponibilità di informazioni tempestive, affidabili, accurate e complete.

La soluzione informatica unanimemente proposta per il soddisfacimento di tali esigenze conoscitive prende il nome di Business Intelligence (BI) ovvero l’insieme dei processi, dei metodi

e degli strumenti utilizzati per raccogliere, organizzare ed analizzare i dati a supporto dei processi decisionali di carattere operativo, tattico e strategico.

Le soluzioni di BI fondono quindi tecnologia, informatica e statistica con l'obiettivo di sfruttare al massimo il patrimonio informativo di un'organizzazione mettendolo a disposizione delle diverse classi di utenza per fini "conoscitivi" in una forma sintetica, tempestiva, facilmente fruibile fino al massimo dettaglio, condivisa nell'interpretazione e certificata nel contenuto.

Per ottenere tali risultati occorre tuttavia che le organizzazioni pongano particolare attenzione alla formazione degli utenti, all'integrazione della BI nei propri processi e alla creazione di un'apposita struttura organizzativa.

Il mercato della BI è in una situazione di espansione guidata da tre driver fondamentali:

- le regolamentazioni e le normative che obbligano le organizzazioni ad essere più aperte e tempestive nelle attività di reportistica e documentazione dei propri dati verso l'esterno;
- la necessità di ottimizzare il ritorno sui consistenti investimenti sostenuti, negli anni passati, in soluzioni informatiche e l'aumento esponenziale dei dati generati in seno alle organizzazioni;
- la necessità da parte delle organizzazioni di operare un maggior controllo sulla propria attività, a causa della congiuntura economica che ha imposto una razionalizzazione degli investimenti e un contenimento di costi di gestione.

2.1.2 Tipologie di soluzioni di BI

La classificazione delle soluzioni di BI adottata in RGS è basata sui processi decisionali (operativi, tattici o strategici) indirizzati dalle seguenti componenti:

- Data Warehouse (DW): è una piattaforma informativa contenente la visione univoca e certificata dell'informazione dell'organizzazione; esso può essere a ragione considerato come le fondamenta di ogni sistema di Business Intelligence. Il dato aziendale, dopo aver subito un processo di integrazione, pulizia, trasformazione e validazione, viene reso

disponibile all'utente come informazione per analisi predefinita (report) o estemporanea (OLAP) a supporto delle decisioni;

- Sistema Direzionale: è un complesso di strumenti a disposizione dei vertici per il controllo, la misura ed il governo dei processi aziendali. Dalla conoscenza integrata e sintetica della realtà dell'organizzazione (nelle sue risorse materiali ed immateriali) disponibile in un cruscotto è possibile effettuare scelte efficaci ed efficienti in accordo con gli obiettivi prefissati.

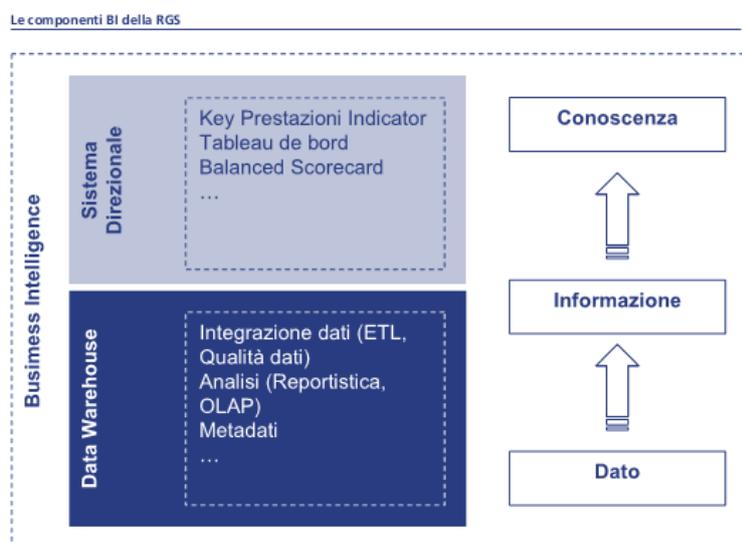


Figura 7

2.1.3 Introduzione al Data Warehouse

Un sistema di data warehouse può essere descritto come un processo di acquisizione, trasformazione e distribuzione di informazioni presenti all'interno o all'esterno dell'organizzazione con finalità di supporto decisionale.

Cuore del sistema è una base dati integrata nella quale confluiscono dati provenienti da più sistemi gestionali e da fonti esterne; è orientata all'oggetto ovvero a temi specifici dell'organizzazione piuttosto che alle applicazioni che la supportano o alle funzioni che svolge; è storicizzata, nel senso che contiene "fotografie" della realtà d'interesse su un orizzonte

temporale molto esteso; ed è permanente, cioè il dato nel DW non è modificabile e può essere acceduto in sola lettura.

Tale base dati, centralizzata ed univoca a livello dell'intera organizzazione, contiene informazioni:

- prodotte dai dati di proprietà dell'organizzazione stessa o a lei resi disponibili, attraverso attività di integrazione, di pulizia (cleaning) e di trasformazione in base a definite regole di business;
- certificate da procedure di analisi della qualità dei dati;
- distribuite attraverso le funzionalità proprie delle applicazioni che realizzano i requisiti di business prefissati;
- interpretate dai responsabili tecnici e dall'utente finale attraverso i metadati, ovvero l'informazione sull'informazione contenuta nel DW.

Possibili architetture (Corporate vs. Federate)

Le finalità "conoscitive" di un DW ed i suoi processi di produzione e distribuzione dell'informazione lo differenziano in modo sostanziale, sia in termini architetturali che organizzativi, dai sistemi gestionali che, per loro natura, hanno essenzialmente il compito di automatizzare i processi secondo un'ottica transazionale.

L'approccio architetturale al DW spazia da un'ottica completamente accentrata (Inmon) ad una completamente distribuita (Kimball).

L'impostazione suggerita da Bill Inmon prevede la creazione di un DW Corporate unico, valido per tutta l'organizzazione costituito da un Enterprise DW (EDW), archivio centralizzato relazionale, normalizzato, di elevato dettaglio e rispondente al modello dati aziendale e da più Data Mart, archivi di maggior aggregazione alimentati dall'EDW, denormalizzati secondo schemi a stella e specializzati per aree di interesse e per classi di utenza.

Viceversa Ralph Kimball propone il DW Federato, uno schema architetturale antitetico all'impostazione Corporate di Inmon, che prevede la creazione di un sistema di vari Data Mart specializzati per aree di interesse, in base alle esigenze di business, strutturati secondo modelli multidimensionali rispondenti ai modelli dimensionali dell'azienda e la gestione di un "DW bus"

condiviso a garanzia della coerenza e della conformità delle dimensioni e delle misure di analisi nei diversi DM.

Ciclo di vita e metodologia

Le due impostazioni, o meglio la loro interpretazione nelle innumerevoli realizzazioni di DW condotte negli ultimi dieci anni nei diversi settori produttivi, hanno in comune un processo di sviluppo che è sia incrementale sia evolutivo:

Incrementale poiché si procede per passi (iterazioni) affrontando delimitati e definiti requisiti funzionali (spesso propri di una singola classe di utenza) su delimitate aree di interesse, realizzando in tempi utili per l'utente alcune funzionalità e affrontando in seguito ulteriori requisiti che daranno luogo a nuove funzionalità integrate con le precedenti.

Evolutivo in quanto deve essere in grado di gestire le naturali modifiche nel tempo degli ambienti sorgente e delle applicazioni gestionali utilizzate dagli utenti finali. Grazie all'adozione di un processo di sviluppo incrementale ed evolutivo i requisiti funzionali dell'utente vengono soddisfatti nel giro di pochi mesi.

Queste caratteristiche del progetto di DW, insieme alla necessità di soddisfare nel più breve tempo possibile i requisiti dell'utente finale, comportano la suddivisione del progetto stesso in più sottoprogetti, detti iterazioni, ciascuna parte delle quali è costituita da fasi di analisi, progettazione, sviluppo, test e collaudo. Ogni iterazione deve essere creata sulle precedenti, in modo da procedere con un effettivo arricchimento del potenziale informativo del DW.

Collante indispensabile per la gestione di un DW, ovvero di tutte le iterazioni di cui si compone il suo ciclo di vita, risulta essere l'applicazione di "best practices" a supporto delle diverse attività di progetto. Obiettivo della metodologia è quello di essere una guida delle attività di progetto, di sincronizzare le diverse attività spesso condotte da team di persone con competenze diverse, di memorizzare e condividere le informazioni proprie delle diverse iterazioni.

Funzionalità a disposizione dell'utente

Le funzionalità di analisi che un DW rende disponibili sono classificabili principalmente come "verification-driven" in quanto la deduzione di informazioni avviene sulla base di ipotesi

precedentemente formulate. E' attraverso tali ipotesi che vengono definiti oggetti di business quali le misure, identificandone la formula di calcolo e le regole di aggregazione, e le gerarchie di analisi, strutturate in attributi dimensionali o descrittivi.

Gli oggetti di business costituiscono lo strato semantico sulla base dati del DW acceduto dalle applicazioni o direttamente dall'utente finale attraverso diverse modalità:

- Report predefiniti prodotti periodicamente e distribuiti a identificate categorie di utenti;
- Analisi estemporanee (OLAP) condotte dall'utente in maniera autonoma e libera attraverso modelli multidimensionali dell'informazione (misure e dimensioni);
- Indicatori sintetici (KPI) in grado di offrire una visione a 360° del contesto di analisi specifico mettendo in evidenza i fenomeni nell'istante in cui si verificano.

Altre tipologie di funzionalità permesse in ambito DW sono quelle di "data mining", che consentono di scoprire nel patrimonio informativo nuova informazione attraverso tecniche di analisi automatica dei dati per la ricerca di relazioni e correlazioni precedentemente sconosciute fra di essi.

In conclusione, fine ed obiettivo di un DW è mettere a disposizione degli utenti di un'organizzazione, attraverso un'interfaccia distribuita, accattivante e di facile utilizzo, un'informazione aggiornata, persistente e storica, univoca e certificata.

2.1.4 Introduzione al Sistema Direzionale

Per Sistema Direzionale si intende lo strumento atto ad integrare le informazioni provenienti da fonti interne ed esterne alla struttura organizzativa per consentire all'alta direzione il monitoraggio delle principali grandezze rappresentative dei processi critici che concorrono allo svolgimento delle attività istituzionali proprie dell'organizzazione.

I Sistemi Direzionali possono essere classificati in tre tipologie sulla base delle attese che l'organizzazione ripone in questi strumenti:

- Management Information System (MIS). Trattano il "Passato" e sono orientati fondamentalmente al controllo interno dell'organizzazione;

- Enterprise Information System (EIS). Sono principalmente orientati all'integrazione, presentazione e monitoraggio dei fattori critici di successo per il controllo sia interno che esterno dell'organizzazione. L'orizzonte temporale trattato è quello "Presente – Passato";
- Decision Support System (DSS). Sono sistemi orientati a supportare l'alta direzione nell'assunzione di decisioni strategiche per il successo dell'organizzazione.

Le necessità odierne delle organizzazioni indirizzano la scelta verso lo sviluppo di soluzioni di tipo EIS. Solo in alcune realtà, dove l'impatto derivante dalle decisioni strategiche ha riflessi economici e/o sociali rilevanti, viene avvertita l'esigenza di dotarsi di strumenti di tipo DSS.

Nell'ambito degli EIS si collocano tutte le applicazioni della moderna BI come, ad esempio, i Cruscotti aziendali o i Tableau de bord e la Balanced ScoreCard. Questi sistemi consentono di rappresentare in modo statico informazioni di sintesi con visualizzazioni grafiche efficaci per fornire informazioni quali allarmi, scostamenti, eccezioni o fenomeni fuori norma. Lo scopo è di attirare l'attenzione dell'alta direzione su quei fenomeni considerati anomali rispetto agli andamenti attesi e per i quali è necessario intervenire con decisioni strategiche.

I DSS hanno l'obiettivo di permettere la modellazione dei fenomeni, la costruzione di scenari e la simulazione degli effetti delle decisioni assunte. Con tali strumenti i decisori hanno a disposizione modelli logico-matematici sui quali fare ipotesi e previsioni su ciò che potrà accadere operando o sull'organizzazione o su altre variabili che influenzano i fattori critici di successo.

I Cruscotti Direzionali

Negli ultimi anni all'interno della Pubblica Amministrazione si è consolidata maggiormente la cultura della conduzione delle attività istituzionali "per obiettivi" legata cioè al raggiungimento di determinati risultati. Si ha pertanto l'esigenza di disporre di pochi ma significativi indicatori per monitorare l'andamento delle attività condotte dall'organizzazione anche al fine di prendere decisioni rapide e tempestive nel caso in cui vengano evidenziati scostamenti rispetto agli andamenti attesi. E' questa la principale funzione di un Cruscotto Direzionale. Uno strumento cioè che:

- consente di avere sotto controllo le principali dimensioni che caratterizzano le attività istituzionali: risultati finanziari e di bilancio, indici di produttività e di qualità, parametri di gestione del personale e che introducendo nel contempo criteri di trasparenza;
- permette di trarre orizzonti temporali, riferiti ad un arco temporale breve per il passato e, in alcune aree strategiche, con proiezioni di medio lungo termine per il futuro;
- presenta tutte le grandezze con modalità ergonomica e visualmente efficace per rendere sinergica l'informazione;
- garantisce un elevato grado di accuratezza dei dati accrescendo la fiducia dell'utente verso l'uso dello strumento e consente di filtrare le informazioni attraverso l'impiego di soglie e allarmi per porre immediatamente l'attenzione su eventi critici nonché di garantire il continuo aggiornamento dell'informazione fornita "in tempo reale" e "al momento giusto".

2.2 Le scelte adottate per la RGS

2.2.1 La soluzione di BI della RGS

La strategia di Business Intelligence adottata dalla RGS per l'assolvimento delle proprie esigenze conoscitive ha condotto alla realizzazione sia di un DW sia di un Sistema Direzionale, con il duplice scopo di:

- creare e distribuire l'informazione, deducibile dall'integrazione dei dati a disposizione dell'organizzazione, alle diverse classi di utenza che necessitano per le loro esigenze operative e tattiche di effettuare analisi multidimensionali, valutare le serie storiche, proiettare gli sviluppi nel futuro e studiare le correlazioni tra fenomeni;
- integrare con altre fonti l'informazione disponibile nel DW ed utilizzarla nella creazione di indicatori (KPI) attraverso cui ottenere un punto di vista immediato e complessivo dell'andamento e delle prestazioni dei processi strategici dell'organizzazione.



Figura 8

La creazione di tale percorso virtuoso dal dato all'informazione e la trasformazione di quest'ultima in strumento di governo e decisione è stata possibile attraverso la realizzazione da parte della RGS dei seguenti due sistemi informatici:

- il DW RGS - un sistema informatico in esercizio dal 2002 nel quale i dati dei sistemi gestionali (es. Bilancio Finanziario, Spese e Contabilità Economica), e di fonte esterna (es. Tesoreria) dopo essere stati omologati ed integrati, sono centralizzati in un Enterprise DW e quindi resi disponibili come informazione certificata su più Data Mart, dedicati ai singoli Ispettorati della RGS, attraverso funzionalità di reportistica predefinita e di analisi multidimensionale;
- il SID - Sistema Informativo Direzionale portale che raccoglie e rende disponibili per l'alta direzione, in un unico punto di osservazione ed in maniera immediata ed integrata, una serie di informazioni ed indicatori per la misura, il controllo ed il governo dei principali processi della RGS.

2.2.2 Architettura delle soluzioni di BI della RGS

La definizione dell'architettura per le soluzioni di BI della RGS adotta il modello di data warehouse Corporate, unico per l'intera organizzazione e specializzato per utenza.

Questo schema architetturale soddisfa requisiti quali la separazione fra l'elaborazione analitica e quella transazionale, l'aggiornabilità del dimensionamento HW e SW a fronte della crescita dei volumi di dati e del numero degli utenti (scalabilità), l'estendibilità a nuove applicazioni e

tecnologie più evolute, il controllo sugli accessi per la natura strategica dell'informazione gestita (sicurezza) , la semplicità delle attività di amministrazione.

Architettura delle soluzioni BI della RGS

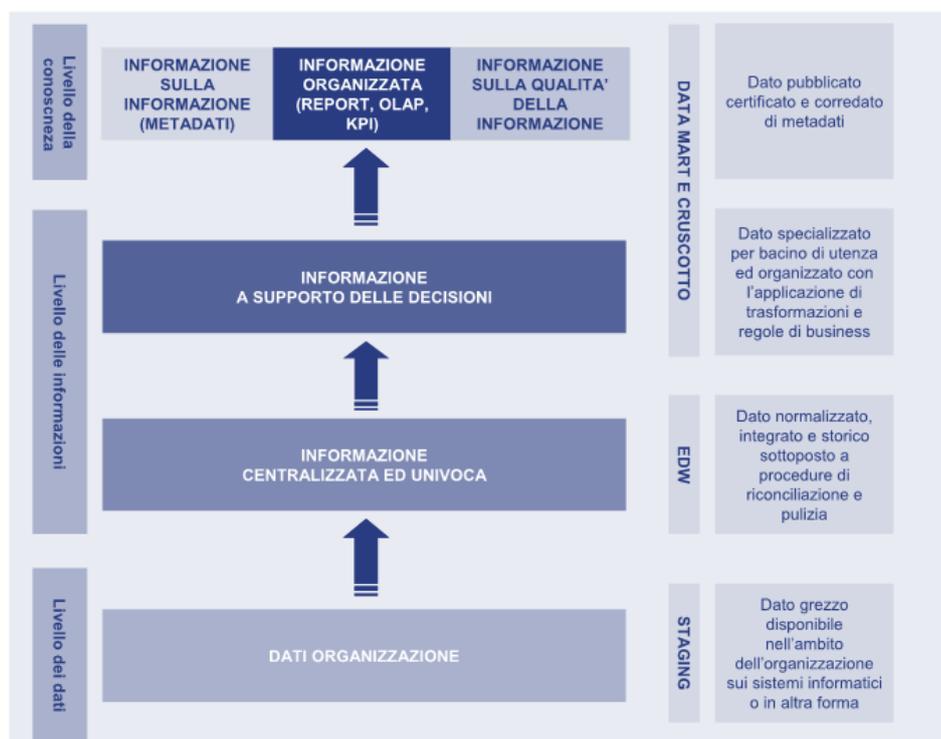


Figura 9

Per la soluzione della BI della RGS sono stati previsti e realizzati tre livelli logici: dei dati, dell'informazione e della conoscenza.

Il primo livello è costituito dall'Area di "staging", che contiene i dati di dettaglio provenienti da sistemi gestionali e da fonti esterne. L'utilizzo di un'area di "staging" è dettato dall'esigenza di rendere l'alimentazione indipendente dai vincoli posti da un accesso diretto ai sorgenti.

Il secondo livello è costituito da una base dati primaria (EDW) e dalle basi dati secondarie ad uso dei DM e del SID, componenti dotate ciascuna del proprio front end di accesso all'informazione.

- L'EDW è un archivio centralizzato, rispondente al modello dati concettuale dell'organizzazione, caratterizzato da dati integrati, di dettaglio e storici giornalmente alimentato dall'area di "staging";

- Le basi dati secondarie per i DM ed il SID sono dedicate alle esigenze operative e tattiche dei singoli Ispettorati della RGS e, per il SID, raccolgono una serie di informazioni e di indicatori inerenti i principali processi della RGS, ad uso del Ragioniere Generale dello Stato e degli Ispettori Generali.

Il terzo livello prevede l'accesso all'informazione contenuta in ciascun DM attraverso:

- il Portale del DM che integra i servizi forniti agli utenti: interrogazioni predefinite, accesso all'ambiente di creazione di interrogazioni personalizzate, metadati relativi all'ambiente ed alle funzionalità predefinite e report sui livelli di qualità del dato pubblicato;
- il Portale del SID dove è possibile valutare l'andamento dei processi da monitorare e l'andamento dei fattori critici di successo.

L'adozione di una metodologia di sviluppo ad hoc basata su un ciclo di sviluppo evolutivo ed incrementale, la definizione di "best practices" calate sulle caratteristiche applicative, tecnologiche ed architetturali del sistema, l'utilizzo di formalismi e metodologie specifici per la modellazione dei dati, per l'integrazione dei metadati e per l'analisi della qualità dei dati costituiscono gli strumenti selezionati per il DW RGS e per il SID al fine di garantire loro una evoluzione razionale, omogenea, coerente e sempre sotto controllo.

A titolo esemplificativo vengono di seguito illustrati gli approcci progettuali e metodologici attinenti i modelli concettuali e i metadati, sottolineando come il carattere "di sistema" che li contraddistingue sia proprio di tutte le tematiche progettuali (ETL, qualità dati, portali, OLAP, riservatezza, ...) del DW RGS e del SID.

2.2.3 Modelli concettuali dell'informazione

L'informazione gestita nei sistemi BI della RGS è descritta a livello concettuale, ovvero al livello più alto e quindi più vicino all'utente finale, attraverso i modelli concettuali Entity/Relationship (E/R) e Dimensional Fact Model (DFM), utilizzati durante la fase di Definizione di ciascuna iterazione per la formalizzazione della realtà oggetto dei sistemi informatici. I due modelli sono integrati attraverso le Regole di business che formalizzano come il requisito utente descritto con

il DFM può essere soddisfatto dai concetti della realtà di interesse definiti e schematizzati con il modello E/R.

Mentre il modello concettuale E/R rappresenta la semantica e la visione integrata dei dati propri dei processi operativi dell'RGS, il DFM è la rappresentazione formale delle modalità di analisi fruibili dall'utente finale. Il modello dimensionale DFM descrive gli eventi caratteristici della realtà di interesse attraverso le misure, attributi con valore numerico, e le gerarchie, percorsi di aggregazione attraverso cui è possibile valorizzare le misure.

Un esempio tratto dall'ambito applicativo trattato è il seguente.

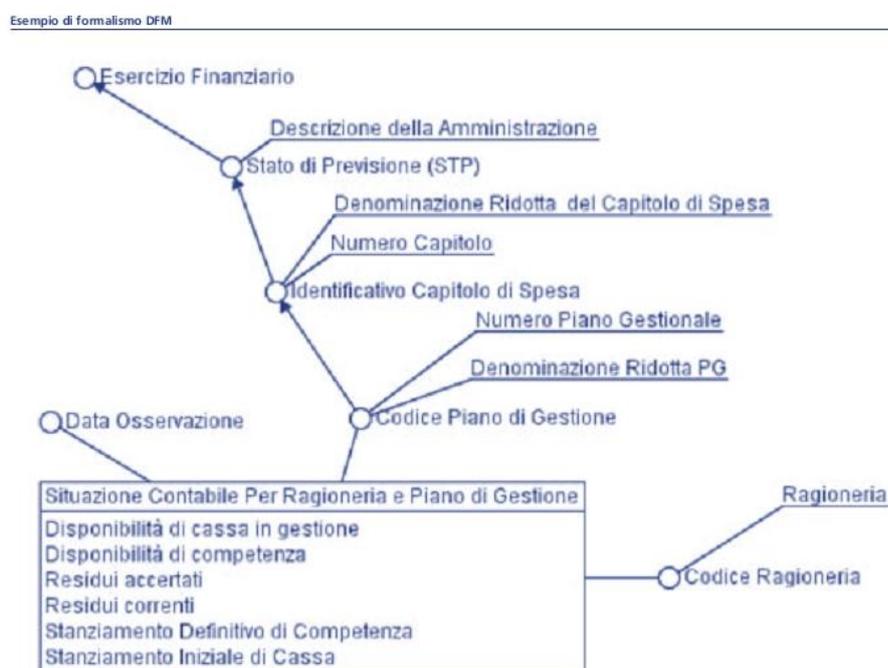


Figura 10

La Figura 10 riporta, secondo il formalismo DFM, la rappresentazione di una funzionalità di analisi del DM realizzato per l'IGICS (Ispettorato Generale per l'Informatizzazione della Contabilità di Stato) che descrive l'esigenza di "osservare la situazione contabile nel tempo di un piano di gestione di ragioneria considerando il dettaglio dei dati di Spesa (Residui accertati, Residui correnti), degli stanziamenti (Stanziamento Definitivo di competenza, Stanziamento Definitivo di cassa) e delle Disponibilità (Disponibilità Competenza, Disponibilità di Cassa in gestione).

Se lo schema Entità/Relazione descrive il patrimonio informativo gestito nel sistema ovvero “cosa” è in esso disponibile, il modello dimensionale DFM descrive “come” l’utente di ciascun Data Mart può analizzare la quota parte di informazione di suo interesse.

Per tali considerazioni mentre lo schema E/R (da cui durante la fase di Progettazione viene derivato il modello logico dell’EDW) è unico per l’intero DW, gli schemi DFM (da cui vengono derivati i modelli logici multidimensionali dei Data Mart e del Sistema direzionale) sono propri e caratteristici dei diversi bacini di utenza.

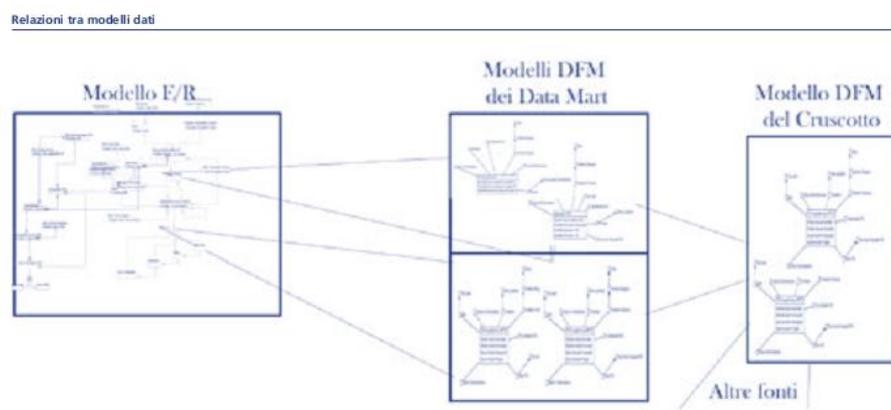


Figura 11

Nella Figura 11 viene mostrato il processo di identificazione e derivazione dell’informazione dal modello E/R ai modelli DFM dei diversi Data Mart e successivamente da questi ultimi al modello DFM del SID. I DFM, descrivendo infatti completamente l’informazione disponibile nei diversi DM, possono essere utilizzati nell’identificazione della fonte informativa degli indicatori del SID.

E’ infatti dall’analisi dei DFM dei Data Mart del DW RGS che si può desumere se un indicatore richiesto sul SID può essere implementato attraverso informazioni già contenute nel DW, oppure si debba accedere anche ad altre base dati dell’organizzazione o a fonti esterne. Il DW viene considerato la fonte da privilegiare nel processo di valorizzazione degli indicatori del SID in quanto garante di maggior efficienza, economicità e qualità: è un dato ufficiale.

2.2.4 Metadati

Nell'informazione presente nel DW RGS e nel SID sono stati associati dei dati correlati (metadati) che ne descrivono l'origine, il significato, la struttura, le versioni, le regole di trasformazione, il formato, le date di validità, ...

I metadati garantiscono una visione coerente, strutturata ed omogenea dell'intero sistema e costituiscono un indispensabile supporto sia al processo di accesso all'informazione contenuta nel DW da parte dell'utente finale, sia al processo iterativo di produzione dell'informazione del DW da parte dei gruppi tecnici.

Il processo di gestione dei metadati del DW RGS e SID prevede la raccolta di gran parte delle informazioni generate durante l'intero processo di sviluppo di ciascuna iterazione:

- nella fase di Definizione vengono raccolte le informazioni del modello concettuale E/R per descrivere la realtà di interesse, del modello dimensionale DFM per la formalizzazione del requisito utente, del modello logico/fisico dei sistemi sorgente per l'analisi delle fonti. Vengono inoltre considerate le regole di business per seguire la trasformazione del dato in informazione;
- nella fase di Progettazione si aggiungono i dettagli del modello logico/fisico dell'EDW, del modello logico/fisico del Data Mart, dello strato semantico degli ambienti di "front end" e del processo di derivazione dei modelli logici dal concettuale;
- nella fase di Realizzazione il quadro dei metadati viene completato con l'integrazione delle informazioni relative alle procedure di alimentazione. Esse costituiscono il legame fra i diversi modelli fisici e la reportistica predefinita, vista come implementazione del requisito funzionale.



Figura 12

La soluzione tecnologica scelta per la gestione dei metadati prevede che essi vengano prodotti nelle diverse fasi del ciclo di sviluppo, dagli strumenti specifici impiegati nella realizzazione del sistema, utilizzando dei moduli software di interfaccia che permettono la loro raccolta all'interno di un unico "repository". Il "repository" è costruito sulla base di un metamodello comune che accoglie i metamodelli delle fonti ("front end", modelli dati, "back end", ...) e ne consente un trattamento omogeneo.

E' quindi possibile condurre sui metadati memorizzati nel "repository" diversi tipi di analisi: Impact Analysis (analisi dell'impatto a fronte di cambiamenti), Where Used (dipendenze funzionali), Data lineage (esito dei processi di ETL), etc.. Ogni necessità di integrazione applicativa e/o pubblicazione del contenuto del "repository" dei metadati su tecnologie di portale viene effettuata secondo lo standard XML.

2.2.5 Riconoscimenti della validità della soluzione

Al fine di verificare la correttezza d'impostazione del Data Warehouse, è stata organizzata e condotta presso la RGS un'attività di "audit" finalizzata alla Certificazione del sistema realizzato. In considerazione della primaria importanza del progetto in questione, la responsabilità del lavoro di monitoraggio e verifica è stata affidata direttamente a Bill Inmon, riconosciuto a livello internazionale come l'ideatore del concetto di data warehouse ed uno dei massimi esperti mondiali in materia.

L'attività di certificazione, attivata nel dicembre 2004 e durata tre settimane, è stata condotta mediante numerose sessioni di presentazione seguite da dibattiti fra gestori e progettisti del DW ed il revisore. A seguire sono stati condotti approfondimenti e discussioni interattive, allo scopo di chiarire e approfondire i temi trattati. Sono stati affrontati argomenti quali gli obiettivi generali e l'organizzazione interna per il governo del sistema, i principi del disegno, gli elementi architettonici, i modelli dei dati, le tecniche per l'analisi della qualità dei dati, la struttura ed il funzionamento delle componenti di "back end" e "front end". Una particolare attenzione è stata impiegata nella valutazione del sistema di gestione dei metadati.

Il giudizio del certificatore, di seguito riportato, evidenzia che il DW RGS si colloca fra i migliori sistemi di warehouse progettati e realizzati negli ultimi anni nel mondo: "... il disegno, la

realizzazione, il sistema di gestione dei metadati e tutta l'infrastruttura a supporto risultano assolutamente eccellenti ...".

Inoltre, il progetto che ha portato alla realizzazione del SID ha avuto un importante riconoscimento dal Computerworld Honors Program, nel corso di una cerimonia che si è tenuta il 5 giugno 2006 a Washington nella prestigiosa sede dell'Andrew W. Mellon Auditorium.

L'assegnazione è avvenuta da parte della Computerworld Information Technology Awards Foundation, associazione che raggruppa presidenti e amministratori delegati delle società leader a livello mondiale nell'IT (in totale circa 100) e che fa capo alla IDG (International Data Group) giunta alla sua 18esima edizione.

Questo riconoscimento rappresenta una sorta di "Oscar" dei progetti informatici più innovativi e di maggiore beneficio verso la società. Nelle dieci categorie in cui è suddiviso il premio, il progetto si è classificato nella categoria "Government & Non-Profit Organizations".

3 La soluzione

3.1 Il Data Warehouse RGS

Il sistema DW RGS raccoglie ed organizza informazioni relative a più ambiti informativi:

- Bilancio Finanziario: informazioni di bilancio e di spesa organizzate per fase di interesse (Formazione, Gestione, Consuntivo) e di riferimento per le attività di costruzione e monitoraggio del Bilancio;
- Analisi Costi e Rendimenti: informazioni strutturate in merito alle rilevazioni economiche analitiche per riqualificare e valutare il costo delle funzioni e dei servizi istituzionali;
- Spesa delle Amministrazioni Centrali: informazioni che rendono possibile la visione completa delle Spese dello Stato, il monitoraggio della Contabilità Ordinaria e Speciale dei Funzionari Delegati;
- Patto di stabilità interno: informazioni per il monitoraggio del rispetto, da parte degli Enti Locali, del "Patto di Stabilità Interno", in applicazione a quanto previsto dall'art. 28

della legge 448/1998, ed agli impegni presi in sede Comunitaria dallo Stato Italiano in ambito Governo Locale (Regioni, Province, Comuni e Comunità Montane);

- Flussi di Cassa Enti: informazioni a supporto della predisposizione della 'Relazione trimestrale di cassa' degli Enti del settore pubblico allargato;
- Flussi di Tesoreria e Flussi SIOPE: informazioni a supporto della predisposizione del Quadro di Costruzione del Settore Statale';
- Spesa per Legge: informazioni per il monitoraggio dell'attuazione del programma economico del Governo e degli stanziamenti di bilancio assegnati con autorizzazioni di interesse per IGAE;
- Spese di Gestione delle Amministrazioni Autonome: informazioni contabili di bilancio e di spesa delle singole Amministrazioni Autonome.

Tali informazioni sono organizzate in data mart indipendenti, alimentati dal DW RGS. Per scelte progettuali, il numero dei data mart non è in corrispondenza uno ad uno con gli ispettorati di riferimento.

Sono rappresentati in Figura 13 i sistemi gestionali alimentanti (fonti) e i DM in esercizio, ognuno dei quali è acceduto da uno specifico segmento d'utenza.

Si riporta qui di seguito la descrizione di alcuni dei data mart presenti in Figura 13.

Data Mart per IGB (Ispettorato Generale per il Bilancio)

Acquisisce i dati previsionali del bilancio finanziario, i dati della gestione delle spese e delle entrate e li organizza per fornire alla dirigenza dell'IGB le informazioni necessarie a supportare le decisioni in materia di Bilancio Finanziario. Le informazioni memorizzate sul Data Mart sono usufruibili a vari livelli (classificazioni economica, amministrativa, funzionale, capitolo, ...) e in base al loro andamento temporale (analisi delle serie storiche e previsioni). L'ambiente è utilizzato nelle singole fasi di bilancio (Formazione, Gestione, Consuntivo) a supporto dei processi trasversali e delle esigenze specifiche dell'Ispettorato.

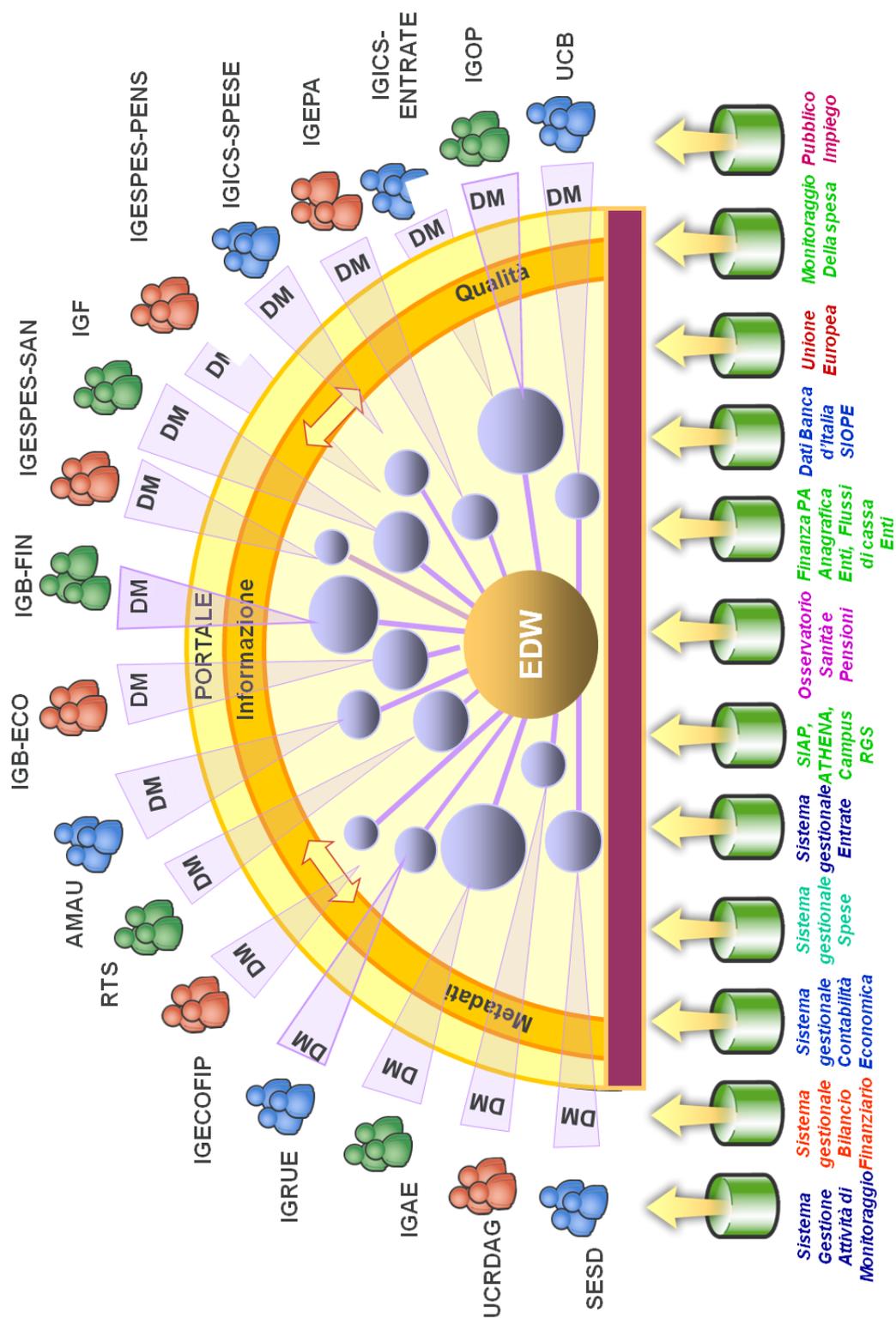


Figura 13

Data Mart per IGEPa (Ispettorato Generale per la Finanza delle Pubbliche Amministrazioni)

A partire dai dati acquisiti dal sistema gestionale circa i prospetti dei 'Flussi di Cassa' e del 'Patto di Stabilità Interno, fornisce gli strumenti per supportare l'IGEPa nei suoi compiti di:

- monitoraggio del rispetto, da parte degli Enti Locali, del "Patto di Stabilità Interno", in applicazione a quanto previsto dall'art. 28 della legge 448/1998, in applicazione agli impegni presi in sede comunitaria dallo Stato Italiano in ambito Governo Locale (Regioni, Province, Comuni e Comunità Montane);
- monitoraggio Flussi di cassa degli Enti per analizzare e monitorare i bilanci inviati degli enti del Settore Pubblico Allargato e supportare i funzionari nella valutazione della relativa gestione di cassa e nella predisposizione delle relazioni trimestrali;
- monitoraggio dei Flussi di Tesoreria e dei Flussi delle operazioni degli Enti (SIOPE) per l'analisi dell'andamento della Spesa Pubblica intesa come integrazione del settore Statale e del Settore Pubblico allargato.

Data Mart per IGICS (Ispettorato Generale per l'Informatizzazione della Contabilità di Stato) e

Data Mart per UCB

Fornisce una visione completa ed esaustiva delle Spese dello Stato sia a livello di singolo titolo di spesa (Ordini di Pagare, Ordini di Accreditamento, Note di Imputazione) che a livello aggregato di Impegno, di Piano Contabile dei Residui e di Piano Gestionale. Permette di monitorare la Contabilità Ordinaria e Speciale dei Funzionari Delegati, figure istituzionali preposte alla distribuzione delle somme messe a disposizione dalle Amministrazioni, attraverso l'aggregazione delle informazioni in base ad un set eterogeneo di dimensioni di analisi .

Fornisce inoltre, solo per il DM per UCB, una serie di indicatori per il monitoraggio mensile di prestazione delle loro attività. La frequenza di aggiornamento dell'informazione è giornaliera.

Data Mart per Amministrazioni Autonome

Supporta le Amministrazioni Autonome, alle quali è stata conferita piena autonomia gestionale in considerazione della peculiare natura delle attività che devono svolgere. Ogni Amministrazione accede esclusivamente alle informazioni di propria pertinenza che riguardano i dati di spesa nella fase di Formazione ed in quella di Gestione/Consuntivo. Le Amministrazioni autonome dello Stato sono (tra parentesi lo stato di previsione al quale sono allegati i relativi bilanci):

- Amministrazione dei Monopoli di Stato (Ministero dell'Economia e delle Finanze);
- Archivi notarili (Ministero della Giustizia);
- Cassa delle Ammende (Ministero della Giustizia);
- Fondo edifici di culto (Ministero dell'Interno);
- Uffici del lavoro portuale (Ministero delle Infrastrutture e dei Trasporti).

Hanno accesso al DM anche la Presidenza del Consiglio dei Ministri ed il Consiglio di Stato e Tribunale Amministrativo Regionale (TAR).

Le informazioni presenti nel DM sono fruibili anche attraverso indicatori calcolabili a differenti livelli di aggregazione e secondo criteri quali: la classificazione economica dei capitoli, le classificazioni, ...

Il sistema ha una frequenza di aggiornamento giornaliera.

3.1.1 Architettura logica e tecnologica

La struttura del DW RGS si compone di una base dati multilivello (Staging area, EDW, DM) e di quattro distinti sottosistemi fra loro integrati che interagiscono nel processo di produzione e pubblicazione dell'informazione:

- Back-End, costituito da tutti i processi relativi alla individuazione ed estrazione dei dati degli ambienti sorgente ed alla loro trasformazione in informazioni all'interno del DW;
- Gestione Qualità, che raggruppa tutti i processi inerenti la valutazione del livello complessivo e l'andamento nel tempo della congruità dei dati del Sistema Conoscitivo;
- Front-End, del quale fanno parte le funzionalità messe a disposizione dell'utente per le sue analisi sull'informazione gestita nel DW e gli strumenti per l'accesso a tali funzionalità (Portale);
- Gestione Metadati che comprende tutte le funzioni relative alla definizione, alimentazione e navigazione del dizionario dei metadati tecnici e di business coinvolti nel ciclo di vita del DW.

Il processo di produzione dell'informazione è finalizzato alla pubblicazione dell'informazione prodotta verso l'utente finale e procede secondo passi successivi che prevedono: l'alimentazione

dell'EDW con dati provenienti dalle diverse fonti (l'informazione viene prelevata dal sistema che ne ha la "proprietà" ossia da dove essa ha origine); l'alimentazione dei DM a partire dall'EDW (filtrando ed aggregando i dati per soddisfare le esigenze peculiari della classe di utenza cui i DM sono destinati); la misurazione della qualità delle informazioni prodotte presenti sia sull'EDW che sui DM (in modo da certificarne la completezza, la validità, l'integrità e la correttezza); la pubblicazione dell'informazione (previa verifica del raggiungimento dei livelli di qualità attesi).

Parallelamente viene gestito l'archivio dei metadati, guida indispensabile dei processi di produzione e di sfruttamento delle informazioni contenute nel Data Warehouse. L'architettura tecnologica del sistema DW RGS è il risultato dell'integrazione di diversi prodotti e di componenti applicative appositamente sviluppate, di cui si fornisce una sintetica presentazione in Figura 14.

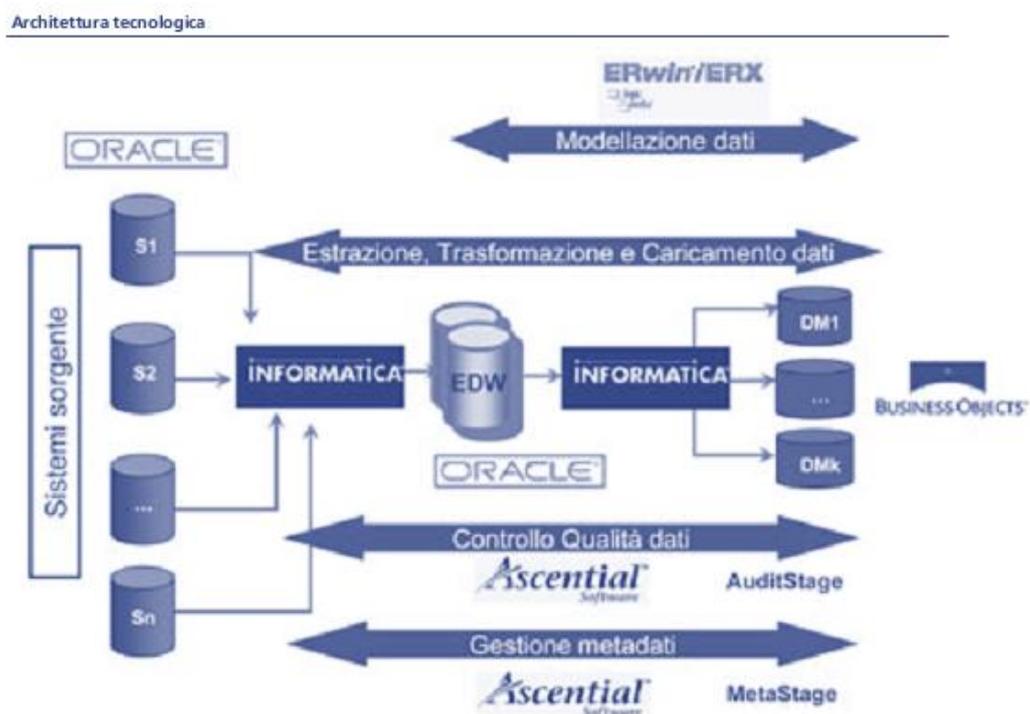


Figura 14

Database: i dati dei tre livelli logici di scomposizione del sistema (Staging Area, EDW e Data Mart) sono ospitati in un unico database Oracle organizzato in differenti utenze, installato su sistema IBM S80 con sistema operativo AIX utilizzato con funzione di database server.

Back end : per l'estrazione dei dati dalle varie fonti della RGS ed esterne e la loro trasformazione e caricamento nei diversi ambienti del DW si utilizza un prodotto di Extraction Transformation Loading (ETL) Informatica – PowerCenter, installato su server IBM S80 con sistema operativo AIX. L'adozione di un tale prodotto per la realizzazione e l'esecuzione delle procedure di alimentazione garantisce migliori prestazioni, affidabilità, manutenibilità e controllo delle procedure. La schedulazione delle procedure è affidata al componente Tivoli Workload Scheduler (TWS).

Data quality management: la qualità delle informazioni prospettate è un fattore cruciale per il DW RGS e ne costituisce l'elemento di misurazione del raggiungimento dei propri obiettivi. Le funzioni base di standardizzazione, correzione e validazione dei dati svolte dai processi ETL sono state affiancate da procedure giornaliere di controllo della qualità dei dati utilizzando procedure realizzate ed eseguite mediante il prodotto IBM/Ascential AuditStage. Audit Stage opera in ambiente Windows e colloquia con le altre componenti del sistema via ODBC.

Front end : la tecnologia BusinessObjects viene utilizzata sia in modalità client/server a due livelli (client – database), sia in modalità web multi-tier (browser – web server/application server/nodi BO – database).

L'accesso dell'utente ai dati è filtrato attraverso uno strato semantico (universo) che fornisce una visione dei dati orientata alle misure e alle dimensioni; su questi universi sono costruiti i report. Universi e report sono memorizzati in un apposito repository.

La soluzione di Portale prevede un accesso web con interfaccia personalizzata, utilizzando le potenzialità di sviluppo del tool Dashboard Manager fornito con il componente Application Foundation. Il controllo degli accessi viene realizzato utilizzando il meccanismo base di controllo accessi della piattaforma BO, che prevede l'uso del repository interno di sicurezza, in cui sono memorizzati gli utenti, le password e i privilegi di accesso. L'accesso vero e proprio avviene attraverso un'unica fase di identificazione iniziale a cui fa seguito la propagazione automatica e trasparente delle credenziali a tutti i sistemi acceduti (single sign on).

Front end

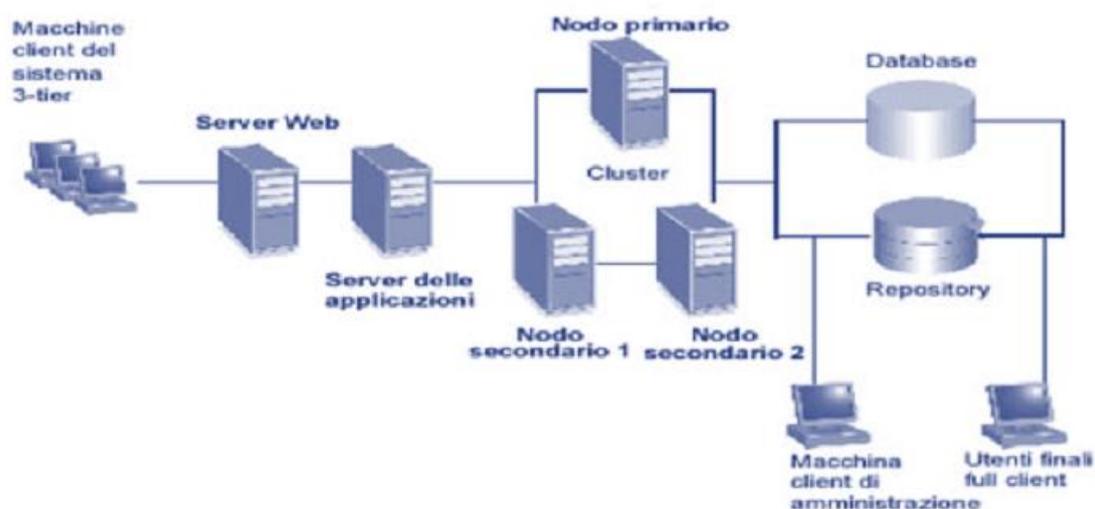


Figura 15

Metadata management: per l'acquisizione, integrazione e pubblicazione dei metadati, è stata realizzata una soluzione custom che include il prodotto IBM/Ascential MetaStage attraverso componenti realizzate ad-hoc. L'architettura del sistema consente di fatto l'integrazione dei metadati prodotti dai tool Erwin, PowerCenter, BusinessObjects e DFMCASE e la loro pubblicazione sul portale per la navigazione da parte degli utenti del sistema conoscitivo.

Nelle diverse fasi del ciclo di realizzazione di ciascuna iterazione di sviluppo sono inoltre utilizzati strumenti "case" di supporto alle attività di analisi, progettazione e realizzazione, sia proprietari che realizzati come soluzioni "trasversali" a supporto del processo, quali: ERWin, DFMCASE, BO Integrator, ecc. Tali strumenti costituiscono anche l'ambiente di produzione dei metadati successivamente integrati applicativamente e pubblicati sul portale.

3.2 Il Sistema Informativo Direzionale

Il SID RGS (Sistema Informativo Direzionale RGS) è un portale che raccoglie e rende disponibili in un unico punto di osservazione ed in maniera immediata ed integrata una serie di informazioni ed indicatori sui principali processi della RGS.

Obiettivi primari supportati dal SID RGS, sono: seguire l'andamento delle Risorse Pubbliche, quale missione della RGS (Finanza Pubblica, Bilancio dello Stato, Monitoraggio della spesa pubblica, Analisi dei costi); seguire il Dipartimento (Controllo di gestione, Attività Ispettiva, Attività Prelegislativa, Personale, Incarichi ecc.); potenziare, con l'ausilio di strumenti evoluti, il colloquio tra il centro e la periferia ampliando la condivisione delle informazioni tra Ragioniere Generale dello Stato, Ispettori Generali, RPS e UCB.

Proprio per cogliere quest'ultimo obiettivo relativo all'ottimizzazione del flusso comunicativo interno alla RGS, nell'ambito del SID sono state realizzate le applicazioni:

- Weekly Report - strumento di condivisione delle informazioni che presenta ai responsabili delle diverse strutture della RGS, per la parte di propria competenza, la sintesi settimanale degli indicatori di processo di rilievo per l'attività operativa;
- Sistema Flussi RGS - strumento di acquisizione e condivisione degli indicatori di prestazione relativi a RPS e UCB e delle tematiche di discussione ai briefing periodici (agende); tali informazioni non sono presenti in altre basi dati e parte di esse contribuisce ad alimentare il Weekly Report.

3.2.1 Architettura tecnologica

La piattaforma adottata si basa sul prodotto Microstrategy, che produce output in formato HTML puro, senza utilizzare applet Java o controlli ActiveX. Tutte le funzionalità di interfaccia ed interazione con l'utente operano in ambiente web.

L'architettura si articola su tre livelli (classica architettura 3-tier):

1. la componente "Web", installata su due macchine configurate in condivisione del carico utente (load balancing), alla quale è assegnato il compito di ospitare le pagine web che permettono agli utenti l'accesso e la navigazione sulla reportistica realizzata con il prodotto Microstrategy;
2. la componente "Intelligence Server" di elaborazione dei report, ospitata su due macchine configurate in cluster, che opera le interrogazioni dei dati sui Database sorgenti e la successiva predisposizione dei grafici e delle tabelle contenuti nei report;

3. la componente “Database Server”, costituita dai sistemi che ospitano i dati utilizzati per la produzione dei report.

La suddivisione delle componenti consente numerosi vantaggi operativi, principalmente un’ottimale distribuzione del carico elaborativo sui sistemi. Infatti la differenziazione tra componente Web e componente di Intelligence Server consente di operare la preparazione delle pagine web per l’accesso e la navigazione tra i report disponibili, mantenendo inalterata tutta la capacità di calcolo necessaria alla predisposizione degli elaborati grafici.

Inoltre la completa separazione logica tra “contenitore” e “contenuto” della reportistica consente di avere utenti che, pur operando su tematiche, archivi e livelli di sicurezza differenti, utilizzano il medesimo motore di reporting attraverso interfacce di accesso indipendenti. Tale configurazione è ad oggi utilizzata nella produzione del Weekly Report, che ha utenze, accessi e report totalmente isolati ed indipendenti dagli elementi disponibili agli utenti del SID.

La componente “Narrowcast Server”, su un’ulteriore macchina, completa l’architettura della piattaforma di SID RGS, ed è dedicata ad effettuare l’invio dei report tramite e-mail o SMS.

Infine, un’altra macchina ospita l’ambiente di validazione di SID RGS dedicato ai collaudi di nuovi report. Tale ambiente, completamente distinto negli archivi e nel funzionamento dall’ambiente di esercizio, consente la verifica delle nuove realizzazioni evitando il rischio di ingenerare anomalie sull’ambiente di esercizio.

Le applicazioni del Sistema Flussi, sviluppate in tecnologia ASP su piattaforma Microsoft, sono disponibili su due server configurati in condivisione del carico utente. I dati raccolti da questi sistemi vengono depositati nelle basi dati presenti sulle macchine di Database Server, per poter essere successivamente interrogati ed eventualmente stampati dagli utenti dell’applicazione stessa.

3.2.2 Fonti dati

Per l’efficacia del SID, la qualità delle fonti dati assume un ruolo fondamentale. L’utilità strategica dello strumento dipende infatti dal grado di affidabilità, coerenza e chiarezza di significato dei dati. Il SID opera su una propria base dati alimentata con diverse modalità. La

fonte principale è il DW RGS, al quale attinge direttamente tramite DB Link. Questa modalità fa sì che le tabelle di DW vengono utilizzate come se fossero lo strato semantico iniziale del SID.

3.2.3 Architettura applicativa

Il SID è suddiviso logicamente in due macroaree informative: l'area Risorse Pubbliche e l'Area del Dipartimento.

L'Area Risorse Pubbliche, che persegue l'obiettivo del Monitoraggio della Gestione delle Risorse dello Stato, è composta dalle seguenti aree tematiche:

- Finanza Pubblica. Sintetizza attraverso indicatori e aggregazioni a vari livelli i principali aggregati finanziari del conto economico della P.A. e fornisce un insieme di informazioni aggiuntive di supporto alla analisi dei report;
- Bilancio dello Stato - Risultati Differenziali. Visualizza indicatori e aggregazioni di dati contabili relativi all'esercizio in formazione, in gestione e in consuntivazione delle entrate e delle spese delle Amministrazioni Centrali.
- Monitoraggio della Spesa Pubblica. Fornisce una visione d'insieme e un dettaglio per Ministero delle segnalazioni effettuate dai process owners della RGS, in base a quanto disposto dalla legge n. 246 del 2001, a fronte di criticità emerse sulle autorizzazioni di fattore legislativo e sui capitoli di spesa che manifestano un anomalo andamento degli impegni o stanno per raggiungere il tetto di spesa;
- Analisi dei Costi dello Stato. Sintetizza, attraverso indicatori e report con diversi livelli di aggregazione, i dati relativi al bilancio economico a livello Stato e ne fornisce l'approfondimento a livello di Macronatura e Finalità.

L'Area del Dipartimento, che supporta gli uffici nell'attività di monitoraggio dei processi della RGS, è composta dalle seguenti aree tematiche:

- Attività Prelegislativa. E' suddivisa nelle sezioni Pareri, con informazioni relative alla situazione dell'ultima settimana e all'andamento mensile delle criticità, e Oneri, con informazioni relative alla portata finanziaria dei provvedimenti in discussione nella settimana;

- Attività Ispettiva. Consente il monitoraggio delle attività ispettive (in prevalenza verifiche amministrativo-contabili nei confronti di altre Amministrazioni statali) svolte dai Servizi Ispettivi di Finanza Pubblica in termini di attività pianificate e concluse nel biennio;
- Personale. Si avvale dei dati forniti dal sistema di Banca Dati del Personale del Ministero dell'Economia e delle Finanze, limitatamente a quanto di competenza del Dipartimento della RGS. Costituisce lo strumento conoscitivo in materia di rapporto di servizio del personale, finalizzato a fornire informazioni di analisi dati e statistiche;
- Controllo di Gestione. Si avvale dei dati forniti dal sistema di Controllo di Gestione del Ministero dell'Economia e delle Finanze, strumento atto alla verifica dell'efficacia, efficienza ed economicità dell'azione amministrativa. La navigazione attraverso l'area tematica consente l'analisi delle prestazioni delle Strutture Organizzative della RGS, del loro andamento tendenziale attraverso l'esame dei dati di budget, consuntivo e del loro scostamento, permettendone l'approfondimento a livello di Macronatura.

Per supportare l'utente finale nella comprensione degli oggetti visualizzati è stata realizzata per ciascun indicatore una "carta d'identità" che fornisce tutte le informazioni rilevanti in termini di descrizione, unità di misura, formula, scala, soglia, ecc..

3.3 L'Organizzazione

Per le caratteristiche del contesto RGS e ancor più per la scelta metodologica di un approccio evolutivo incrementale per il DW, l'implementazione del DW RGS non si può ricondurre tout court ad un progetto, con un inizio e una fine, dopo il quale subentra una gestione del sistema realizzato.

Viceversa per assicurare coerenza e successo all'iniziativa occorrono, oltre alle tradizionali attività di sviluppo e gestione, attività permanenti di governo e di presidio del sistema.

Per tali motivi nasce il "Business Intelligence Competency Center" (BICC) (Figura 16), struttura operativa che combina:

- risorse e mezzi della RGS per il governo e il presidio del DW RGS (parte superiore del rettangolo in Figura 16);

- risorse e mezzi della Consip per lo sviluppo e la gestione del DW RGS (parte inferiore del rettangolo in Figura 16);

che operano all'interno di regole prefissate.

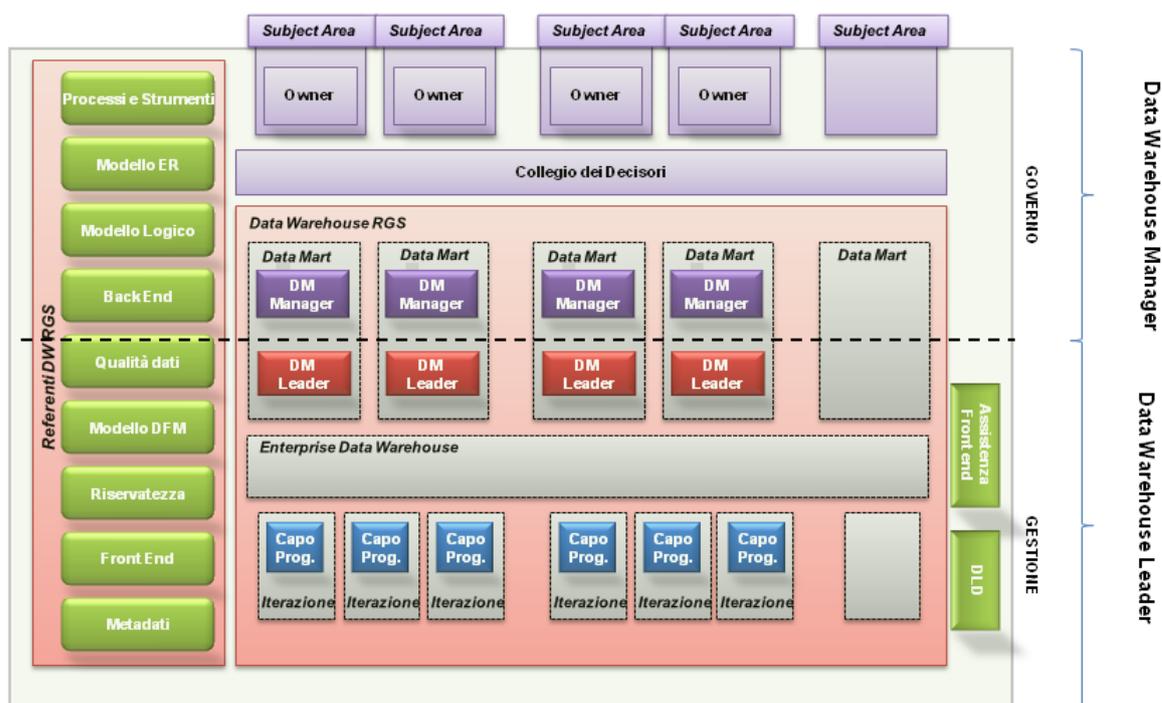


Figura 16

La struttura RGS preposta al governo del DW comprende:

Il **Collegio dei Decisori**, responsabile dell'intero sistema, che ha le seguenti responsabilità:

- project management del Data Warehouse RGS ovvero della pianificazione delle iterazioni, delle classificazione delle priorità, del rispetto del budget e della concessione delle autorizzazioni necessarie;
- gestione degli albi degli Owner e dei Data Mart Manager e presidio della policy e dei processi di riservatezza;
- verifica delle esigenze o della disponibilità delle funzionalità d'analisi richieste dal Data Mart Manager;
- monitoraggio della qualità del dato/informazione;

L'**Owner** che assume la responsabilità della semantica, della disponibilità e dei livelli di qualità/riservatezza propri di ciascun dato proveniente dai sistemi operazionali ed infine delle fonti esterne di dati che rappresenta presso RGS.

Il **Data Mart Manager** che ha la responsabilità:

- di un DM e del bacino di utenza per la soddisfazione delle cui esigenze conoscitive il DM è stato predisposto;
- della identificazione e del soddisfacimento delle esigenze conoscitive;
- della definizione delle politiche di accesso;
- della verifica dei livelli di qualità e di servizio forniti.

La struttura Consip deputata allo sviluppo ed alla gestione tecnica ed applicativa del DW prevede:

Il **Responsabile EDW**, incaricato della pianificazione, realizzazione e gestione dell'EDW e dei DM;

Il **Data Mart Leader**, incaricato della raccolta e formalizzazione delle esigenze della RGS per la creazione o evoluzione dei DM, focal point delle necessità operative del DM Manager e garante del suo livello di soddisfazione;

I **Capi Progetto**, responsabili del rilascio, nei tempi e modalità prefissati, delle componenti di DM proprie di una iterazione progettuale.

I **Referenti del Data Warehouse RGS**, che hanno gli obiettivi di essere di ausilio al processo di produzione del Data Warehouse RGS recependo, integrando ed omogeneizzando le problematiche proprie di un Data Warehouse (metodologia, "front end", "back end", metadati, qualità dati, ...) e di essere di supporto alla struttura RGS ed in particolare al DW Manager e in generale del Collegio dei Decisori, per il governo e il presidio del DW. I Referenti governano i seguenti ambiti:

- **Modello concettuale:** verifica la correttezza, la completezza ed i collegamento fra i modelli concettuali delle subject area e degli EDW (modelli E/R) ed i modelli concettuali dei DM (Dimensional Fact Model - DFM);
- **Qualità dati:** supporta lo sviluppo per l'attività di pubblicazione dell'informazione propria di ciascun DM, attraverso la gestione dei livelli di qualità prefissati e delle eventuali segnalazioni di non conformità; all'interno di ciascuna iterazione del DW RGS; è inoltre il

responsabile della raccolta, della definizione e del monitoraggio delle “best practices” da adottare per la corretta applicazione della tecnologia di analisi della qualità dati;

- Riservatezza dati: supporta la redazione dell’albo degli Owner e dei Data Mart Manager; assicura inoltre la gestione dei livelli di riservatezza dei dati propri delle Subject area, delle informazioni proprie dei DM e quindi dei criteri di accesso a queste ultime;
- Processi e Strumenti DW: supporta l’evoluzione del DW nelle metodologie di processo e degli standard documentali (linee guida progettuali) selezionati per la corretta conduzione dell’intero ciclo di vita di ciascuna iterazione del DW RGS; è altresì il riferimento per le attività di gestione degli interventi di adeguamento evidenziati per il DW RGS dovuti all’introduzione di nuove tecnologie o di cambiamenti dell’environment;
- Front End: supporta lo sviluppo delle funzionalità d’analisi proprie dei processi decisionali identificate come requisito utente ed implementate attraverso reportistica predefinita o estemporanea nell’ambito di una iterazione del DW RGS;
- Back End: presidia il processo di alimentazione proprio di ciascuna iterazione del DW RGS, ovvero delle problematiche inerenti le metodologie, gli ambienti, i processi e le tecnologie proprie delle procedure di estrazione, trasformazione e caricamento (ETL) dei dati dell’EDW e dei DM;
- Metadati: garantisce la completezza e correttezza delle informazioni “guida” del DW e supporta ciascun DM per tutte le tematiche inerenti la raccolta e la gestione dell’insieme dei metadati tecnici e di business in forma accurata ed attuale; assicura che gli utenti siano in grado di comprendere e consultare i metadati all’interno del portale, verificando il livello di integrazione applicativa e logica con le funzionalità di analisi;
- Modello Logico/Fisico: supporta disegno, scelte progettuali, gestione degli schemi dati e configurazione degli ambienti, assicurando la coerenza delle diverse basi dati (“staging” Area, EDW, DM) costituenti nel loro complesso le strutture dati del DW.

Il gruppo **Daily Loading Data**: struttura dedicata preposta alle attività di caricamento dei dati, avente come obiettivo primario quello di massimizzare la disponibilità e la qualità dell’informazione verso l’utente finale nel rispetto delle finestre temporali assegnate, che ha un compito molto delicato per l’ampiezza rilevante (per informazione gestita e moli di dati trattate) raggiunta dal DW e per il numero e la complessità dei componenti che quotidianamente aggiornano e mettono in linea l’informazione.

3.4 I processi di governo e di sviluppo

La gestione dei processi di governo e di sviluppo è strutturata sulla base di BI Actions (BIA), ovvero delle macro attività di amministrazione di cui il BICC si serve per governare il DW.



Figura 17

Le BIA sono inserite all'interno di un processo generale di governo del DW. Tale processo prevede che, sulla base di input quali una nuova esigenza utente, una modifica all'ambiente di sistema o come conseguenza dell'ordinario monitoraggio della qualità dei dati gestiti dal sistema, vengano eseguite una serie di attività tendenti all'evoluzione del DW in termini di patrimonio informativo e funzionalità, alla gestione quotidiana dell'esercizio ed alla diffusione/promozione dell'utilizzo del DW da parte del parco utenti.

Ognuna delle BIA ha un evento attivatore (trigger,) uno svolgimento (workflow), cui partecipano i vari attori coinvolti nel processo, ed un output. Il diagramma di contesto nella Figura 18 riporta le relazioni che ciascuna BIA ha con le altre, i trigger e gli output di ciascuna.

Diagramma di contesto

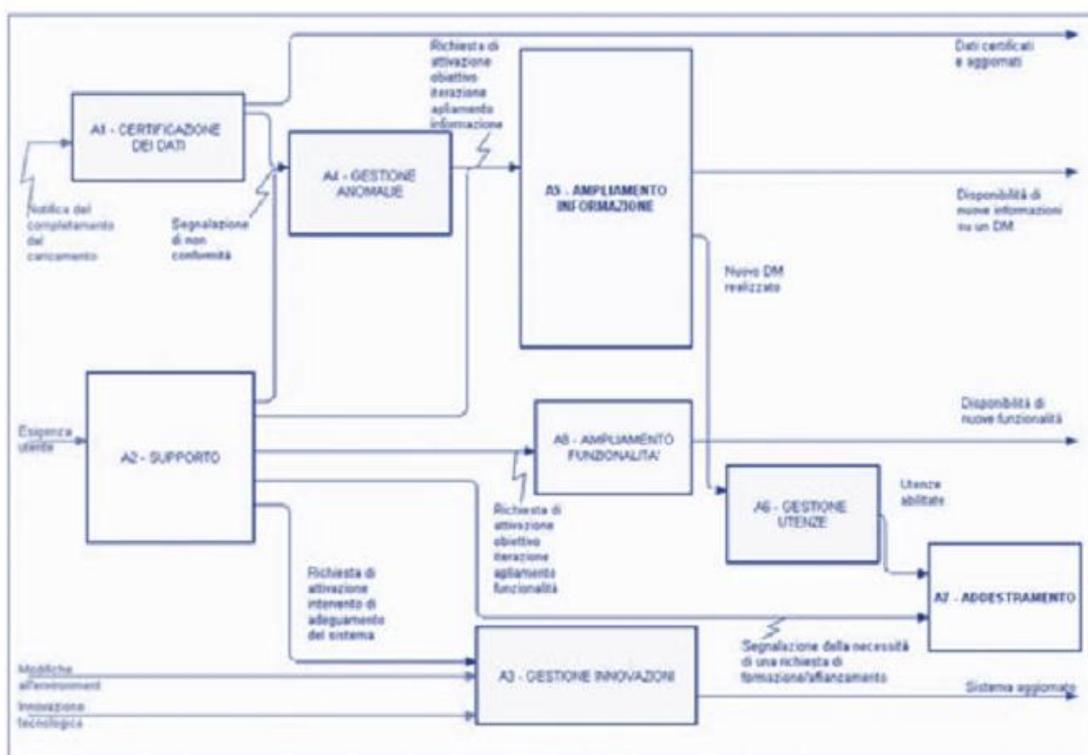


Figura 18

- **Certificazione dei dati:** è la BIA finalizzata a garantire la certificazione dei dati dei diversi DM. Attivata ad ogni aggiornamento dell'informazione nel DW, mediante l'attività di verifica della qualità dei dati, conduce:
 - in caso di esito positivo della verifica, all'aggiornamento dell'informazione che viene resa disponibile agli utenti del DM,
 - in caso di esito negativo, all'apertura di una non conformità ed all'attivazione della BIA Gestione Anomalie per la risoluzione della non conformità stessa;
- **Supporto:** è la BIA finalizzata all'individuazione delle azioni più adatte al soddisfacimento di un'esigenza manifestata da un DMM per sua iniziativa o su indicazione del parco utenti. La BIA, attraverso le attività di raccolta ed analisi dell'esigenza (il cui risultato è una classificazione dell'esigenza stessa), conduce all'assolvimento estemporaneo dell'esigenza o all'attivazione di una BIA fra le seguenti: Gestione Innovazioni, Gestione Anomalie, Ampliamento Informazione, Ampliamento Funzionalità, Addestramento;
- **Gestione Innovazioni:** è la BIA mediante la quale il gruppo preposto al governo del DW autorizza e gestisce l'introduzione di innovazioni tecnologiche e garantisce l'adeguamento del sistema ai cambiamenti dell'environment in cui è collocato.

- Gestione Anomalie: un'apertura di non conformità, scaturita dall'analisi di una segnalazione da parte dell'utente o dall'esito negativo dell'attività di monitoraggio della qualità svolta mediante la BIA Certificazione dei dati, porta alla sua correzione oppure all'attivazione della BIA Ampliamento Informazione;
- Ampliamento Informazione: è la BIA finalizzata alla realizzazione di un'iterazione progettuale che porterà all'estensione del patrimonio informativo del DW; la BIA Ampliamento Informazione può essere attivata da una nuova esigenza del parco utenti di un DM o, nel corso della Gestione Anomalie, nel caso in cui per la risoluzione della non conformità, sia necessario ampliare il patrimonio informativo dell'EDW;
- Ampliamento Funzionalità: è la BIA finalizzata, a partire da una nuova esigenza utente, alla realizzazione di un'iterazione progettuale che porterà all'estensione dello strato funzionale del DW;
- Gestione Utenze: è la BIA finalizzata, a seguito di un rilascio di una nuova iterazione o per iniziativa diretta del DMM di un DM, alla creazione e gestione delle utenze dei DM;
- Addestramento: nei confronti degli utenti di ogni DM vengono condotte le attività di addestramento, mediante l'erogazione di corsi, o affiancamento all'utente volte a diffondere, semplificare, illustrare l'utilizzo delle funzionalità del sistema.

4 I risultati

4.1 L'informazione nel processo decisionale

L'adozione di metodologie e strumenti di Business Intelligence da parte della RGS è scaturita direttamente dalla necessità di rispondere ad esigenze concrete di disponibilità di dati e dell'informazione da essi derivabile a supporto dei differenti processi decisionali. Pertanto, l'introduzione di sistemi evoluti e completamente dedicati alla funzione di supporto alle decisioni quali il DW RGS e il SID ha permesso alla RGS di mettere a frutto l'enorme ricchezza offerta dal patrimonio di dati direttamente prodotti dagli uffici nell'attuazione dei propri compiti istituzionali o raccolti come parte attiva nei processi e negli adempimenti propri di altri soggetti istituzionali.

L'introduzione progressiva degli strumenti di BI nell'ambito dei vari processi operativi è stata ben accolta da parte del personale dell'Amministrazione, a fronte della facilità di trattamento dei dati e della velocità di conduzione dell'analisi abbinata alla efficacia della prospettazione dei risultati che ne facilita l'interpretazione. L'aumento della quantità e della qualità delle informazioni a disposizione, le garanzie di affidabilità, accuratezza e completezza dei dati, la maggiore tempestività e freschezza dei risultati distribuibili, la possibilità di comporre semplicemente quadri di sintesi o di segnalare puntualmente eventi significativi attraverso "misuratori di fenomeni specifici" (indicatori) hanno favorito la naturale evoluzione delle prassi lavorative di tipo conoscitivo.

Tali strumenti sono oggi utilizzati sia come fonte per la produzione di informazione di scambio fra i vari uffici secondo una data periodicità ed a fronte di iter di lavorazioni consolidate (produzione di prospetti a scadenza, atti dovuti, ecc.), sia come mezzi per operare valutazioni a fronte delle quali decidere interventi propri di ciascun ambito di competenza.

Attualmente sono a disposizione ed in uso nella pratica quotidiana presso tutti gli uffici della RGS abilitati all'utilizzo dei sistemi conoscitivi una varietà di report predefiniti e l'accesso agli ambienti per le analisi estemporanee. La disponibilità di indicatori di sintesi per la conduzione di analisi di trend ed il monitoraggio di particolari fenomeni di interesse consente, inoltre, agli utenti che svolgono tali attività, di avere un ulteriore sofisticato livello di supporto alle proprie valutazioni. Le funzionalità offerte coadiuvano gli utenti nella maggior parte delle attività istituzionali dell'Amministrazione, dalla programmazione delle risorse pubbliche alla verifica puntuale della loro gestione, dalla valutazione dell'impatto finanziario dei provvedimenti di Governo e Parlamento alla verifica di coerenza con gli obiettivi di finanza pubblica.

I diversi prospetti e servizi forniti mettono a disposizione su questi temi punti di vista differenti, specificatamente disegnati per la singola classe di utenza. Così, ad esempio, le informazioni relative alla programmazione delle risorse pubbliche sono riorganizzate e prospettate ad uso di IGB ed IGAE secondo un taglio orientato per il primo a supportare le diverse operazioni di costruzione del bilancio, per il secondo ad operare il monitoraggio della portata finanziaria e degli oneri correnti delle autorizzazioni. Analogamente il tema dell'analisi dell'andamento delle spese relativo alla gestione del bilancio è affrontato diversamente per gli UCB, ai quali viene offerta una vista puntuale sui differenti titoli di spesa e per l'utenza del SID, per la quale sono stati predisposti indicatori di sintesi. Questi ultimi consentono anche un monitoraggio dell'andamento complessivo degli impegni e dei pagamenti delle amministrazioni raffrontati con

quello delle previsioni e la misurazione di fattori quali la capacità di impiego della massa impegnabile o la capacità di erogazione della spesa.

L'informazione a disposizione viene fruita con modalità differenti e per le diverse finalità, seguendo le tipicità dei vari processi e svolgendo così un ruolo di primo piano nel miglioramento della qualità del risultato finale.

I principali benefici indotti da tale approccio riguardano il contributo alla ottimizzazione dell'organizzazione delle attività con particolari risultati in termini di:

- tempestività di attivazione dei processi;
- velocità nell'espletamento delle attività;
- qualità e precisione dei risultati;
- ottimizzazione delle procedure;
- raggiungimento di risultati di maggiore complessità.

Risulta chiaro come riducendo la complessità e l'onere degli sforzi necessari per procurarsi gli strumenti (disponibilità di informazioni di base, mezzi di elaborazione e strumenti per la previsione) e garantendo la precisione e la correttezza di metodi e risultati (qualità dell'informazione), si consente al decisore di concentrarsi sulla decisione stessa ed al contempo si forniscono elementi di maggiore utilità e più semplice interpretazione per il raggiungimento della finalità.

Come ulteriore ma non secondaria conseguenza, inoltre, la disponibilità di sistemi dedicati al supporto delle decisioni, costituiti non soltanto dai sistemi informatici realizzati e dalle loro funzionalità, ma dalle strutture e risorse ad essi affiancate per attività quali la formazione all'uso ed ai contenuti, il supporto di consulenza, l'assistenza di processo, si rileva la possibilità per l'utente di raffinare i propri processi lavorativi integrando in un unico processo: la definizione dei propri obiettivi, il reperimento delle informazioni necessarie, l'analisi, la decisione, l'intervento, mantenendo il governo del processo ed automatizzandone i passi più operativi.

4.2 Il governo dell'informazione

All'interno della RGS i sistemi gestionali e le basi dati sono stati utilizzati storicamente a supporto dei processi operazionali dell'Amministrazione. Le funzionalità fino ad oggi sviluppate hanno riguardato principalmente attività di funzionamento della macchina pubblica quali la gestione del bilancio, l'erogazione della spesa, l'accertamento delle entrate. Gli investimenti si sono concentrati sulla automatizzazione dei processi per ottenerne vantaggi sia in termini di velocizzazione delle attività che di accuratezza dei dati. Gli effetti si sono principalmente concretizzati in un reale abbattimento dei costi di processo a fronte degli investimenti effettuati.

Successivamente l'attenzione dell'IT della RGS è passata dalla focalizzazione sull'ottimizzazione dei processi ad una forte attenzione al supporto strategico, con l'idea di fornire strumenti efficaci ai processi decisionali dell'Amministrazione in materia di finanza pubblica. Le attività, a questo punto si sono indirizzate verso la gestione della grande quantità di dati disponibili con l'obiettivo di trasformarli in informazioni strategicamente utili.

La creazione della conoscenza all'interno della RGS ha comportato di conseguenza la ricerca sia di:

- metodologie per l'individuazione di indicatori di prestazione dei fattori critici di successo dell'Amministrazione;
- fonti informative sia interne che esterne da utilizzare per le analisi;
- efficaci modalità di rappresentazioni delle informazioni.

4.2.1 La metodologia per l'individuazione degli indicatori

La metodologia messa a punto si basa su un processo definito che prevede l'impiego di una modulistica atta a facilitare la raccolta e la registrazione di tutte le informazioni necessarie ad individuare le esigenze nonché gli indicatori di sintesi rappresentativi dei processi critici.

I passi da compiere, nel rispetto della metodologia, per l'individuazione dei requisiti di alto livello sono:

1. Identificazione e classificazione dei processi e dei fattori critici di successo CSF (I fattori critici di successo sono quegli elementi ritenuti critici dalle organizzazioni per il raggiungimento della propria missione. La loro individuazione permette di circoscrivere il dominio delle informazioni da raccogliere). I processi” vengono classificati secondo le seguenti tipologie:

- processi “core” che creano valore riconosciuto all’organizzazione ed alla pubblica amministrazione;
- processi di supporto necessari allo svolgimento delle attività istituzionali e generalmente orientati all’interno dell’organizzazione.

2. Individuazione delle grandezze e degli indicatori di prestazione KPI distinguendoli in:

- KPI di efficienza;
- KPI di qualità;
- KPI di servizio.

L’efficacia di un processo, in particolare, rappresenta il valore espresso da un’attività rispetto alla qualità immessa nel processo per produrre il risultato.

L’efficienza di un processo misura invece il rapporto tra volumi di risorse necessarie a produrre il risultato ed il valore del risultato ottenuto.

L’efficacia di un processo può risultare in conflitto con la sua efficienza. Ad esempio l’aumento di risorse assegnate ad una attività, ne peggiora l’efficienza a vantaggio dell’efficacia ottenuta diminuendo i tempi di svolgimento dell’attività.

3. Profilatura degli indicatori. Per ogni tipo indicatore vengono poi definite le seguenti caratteristiche:

- l’unità di misura;
- l’algoritmo di calcolo;
- le fonti dalle quali rilevare le grandezze utilizzate nell’algoritmo
- l’importanza (alta, media, bassa);

4. Verifica di copertura tra indicatori di prestazione (KPI) e fattori critici (CSF). L’attività ha lo scopo di valutare il grado di copertura/strategicità degli indicatori individuati rispetto ai fattori critici enunciati nel corso delle precedenti fasi di analisi. I risultati di tale verifica possono dar luogo ad un’ ulteriore fase di riciclo dei passi precedenti per:

- individuare nuovi KPI
- rivedere la lista dei CSF.

5. Analisi dimensionale. Viene effettuata la ricerca delle dimensioni tramite le quali rappresentare i KPI effettivi e la scomposizione gerarchica di ciascuna dimensione.
6. Rifinitura. Attraverso uno o più cicli vengono perfezionati:
 - i KPI che saranno realizzati;
 - la reportistica direzionale necessaria;
 - la griglia di accesso alle informazioni (profilatura utenti).

Al termine viene prodotto e condiviso con gli utenti un documento di sintesi contenente:

- tutti i moduli compilati durante le fasi precedenti;
- la descrizione di sintesi degli indicatori e delle dimensioni di analisi di cui si richiede la realizzazione sul SID;
- la lista dei gruppi di accesso alle informazioni.

4.2.2 Le fonti informative

Nell'ambito dei processi interni della RGS vengono prodotti e gestiti quotidianamente una enorme quantità di dati sia di natura strutturata – archiviati all'interno dei vari sistemi informativi - che destrutturata – archiviati su carta. I sistemi informativi che presiedono alla gestione di questi dati, sono di diversa tipologia e autonomi in quanto supportano processi verticali. Il management della RGS sempre più spesso manifesta la necessità di informazioni affidabili, sintetiche e possibilmente fruibili all'interno di un unico ambiente di consultazione che fosse anche di facile utilizzo e, soprattutto, che garantisse la stessa affidabilità dei sistemi gestionali.

Questo ha posto nella fase realizzativa dei sistemi decisionali l'esigenza di effettuare alcune scelte strategiche.

In primo luogo la necessità di dotarsi di un DW, cioè di un magazzino di dati opportunamente strutturati ai fini di reporting ed analisi, riferibili ai processi maggiormente consolidati all'interno della RGS.

In secondo luogo l'eventualità di ricorrere a fonti alternative di dati per rispondere ad esigenze "tattiche" non prevedibili a priori.

Infatti mentre il DW svolge le funzioni tipiche di un deposito di dati, fondamentale per la costruzione di informazioni che possano diventare fruibili nel tempo da parte dei livelli superiori del management, non può essere trascurato l'approccio più pragmatico all'informazione e quindi il ricorso a flussi informativi predisposti ad hoc per acquisire nuovi dati.

Una delle esigenze più avvertite dalla RGS è la possibilità di disporre di soluzioni tempestive ed efficaci per far fronte alla normativa che annualmente la legge finanziaria introduce per il controllo dell'andamento della spesa pubblica. Ad esempio la legge finanziaria del 2006 ha introdotto due nuove modalità di controllo della spesa: il comma 7 che dispone l'obbligo da parte delle Amministrazione Centrali dello Stato di erogare la spesa per alcuni comparti frazionandola in parti uguali nell'anno; il comma 21 che ha disposto la possibilità di blocco dei capitoli di spesa qualora venissero rilevati eccessivi impegni rapportati all'andamento atteso.

Per poter fornire agli organi preposti al controllo della spesa uno strumento efficace e facilmente fruibile, si è operato immediatamente con un approccio tattico andando ad utilizzare all'interno del SID sia informazioni presenti nel DW che fonti esterne; successivamente perseguendo l'approccio strategico andando ad integrare nel DW anche i dati provenienti dalle fonti esterne.

Questo approccio ha permesso di fornire in brevissimo tempo e con la garanzia della qualità dei dati, lo strumento necessario al controllo, da parte degli organi preposti, del rispetto di quanto introdotto dalla normativa.

In generale l'approccio strategico che passa per l'integrazione con il DW è preferito all'approccio tattico laddove i tempi di realizzazione della soluzione individuata, la disponibilità dei dati e le necessità espresse lo consentono.

4.2.3 La rappresentazioni delle informazioni

La necessità di contenere all'interno di poche pagine tutte le grandezze rappresentative dei complessi fenomeni presenti in ciascuna delle aree tematiche di interesse della RGS, ha imposto l'impiego di metodologie di rappresentazione che, sfruttando la percezione visiva, fossero in grado di comunicare una grande quantità di informazioni con eccezionale chiarezza (visual design).

Si è pertanto fatto ricorso alla “Progettazione Visuale”.

Nella pratica, partendo dagli indicatori di prestazione individuati nella fase di analisi, ricorrendo a rappresentazioni grafiche, iconografiche e tabulari, si procede con la predisposizione di bozzetti che vengono resi sempre più raffinati collaborando strettamente con gli utenti.

La pagina del SID può essere costruita tenendo presente i seguenti fattori:

- impiego dell’area di “massima attenzione”, posizionata nel primo quadrante in alto a sinistra della pagina, per prospettare le informazioni più rilevanti;
- uso omogeneo di colori a rappresentare analoghe condizioni dei fenomeni rappresentati. Ad esempio l’uso del colore rosso per indicare livelli elevati di allarme, l’uso del colore verde per situazioni che rientrano nella normalità, uso di colori intermedi per evidenziare il raggiungimento di soglie di attenzione;
- uso di simboli iconografici per rappresentare tendenze o stati di allarme. Triangoli orientati verso l’alto rappresentano trend positivi, verso il basso negativi. Un semaforo rappresenta ad esempio fenomeni che possono assumere tre diversi livelli di soglia;
- largo uso delle soglie. La soglia rappresenta la conoscenza condivisa tacitamente dagli esperti e alla quale riferirsi per giudicare lo stato di criticità dei fenomeni. Le soglie non sono generalmente definite nelle fonti dati ma provengono dall’esperienza di esperti della tematica;
- impiego di rappresentazioni grafiche che diano una indicazione qualitativa dei fenomeni. L’uso di tachimetri per la rappresentazione di indicatori critici (KPI) evidenzia lo stato di un fenomeno “a colpo d’occhio” privilegiando l’effetto visuale a scapito della puntuale rappresentazione del suo valore;
- il posizionamento dei singoli elementi sulla pagina viene studiato per massimizzare la possibilità di confronti immediati tra fenomeni correlati. L’accostamento tra grandezze relative alla disponibilità di cassa e di competenza nell’ambito della spesa ad esempio, evidenzia con immediatezza eventuali scostamenti tra gli andamenti attesi;
- la tassonomia delle informazioni posizionate sulle pagine componenti il SID è fondamentale per consentirne una fruizione ottimale. Ad esempio la scelta di suddividere i contenuti delle pagine per “processi” operativi, permette una lettura più immediata dei KPI che presiedono al loro monitoraggio.

4.3 I benefici

Per evidenziare i vantaggi ed i benefici indotti dai sistemi di BI della RGS è opportuno considerare la natura “strumentale” che è propria della tecnologia informatica e il valore di fattore “abilitante” che essa assume nell’ambito delle più generali strategie dell’Amministrazione. Nel caso in esame occorre verificare come la BI abbia effettivamente contribuito ai risultati assicurando l’efficienza e la qualità dei processi informativi e l’acquisizione di effettivi vantaggi attraverso soluzioni anche di tipo innovativo.

Nella valutazione dei benefici della soluzione di BI della RGS sono stati presi in considerazione sia fattori economici quali il ritorno degli investimenti ed il contenimento dei costi, sia fattori qualitativi che evidenziano la creazione di “valore” in termini di benefici organizzativi, procedurali, strutturali e culturali. Proprio questi ultimi rappresentano i veri fattori cui sono connessi l’effettiva evoluzione ed il reale sviluppo dell’Amministrazione.

Gli indicatori, riconosciuti sia dai responsabili IT sia dall’Amministrazione, che hanno guidato la valutazione del DW RGS e del SID sono i seguenti :

- Il ritorno sull’investimento (ROI): il parametro è costantemente migliorato dall’aumento di produttività generato dalla possibilità offerta agli utenti di dedicare un tempo sempre maggiore all’analisi dei dati piuttosto che alla loro raccolta, collezione e preanalisi; attraverso il DW si ha inoltre la possibilità di gestire contemporaneamente sia una macro che una micro-prospettiva dell’organizzazione lavorando con dati di dettaglio e con dati aggregati risparmiando ore di lavoro per soddisfare le richieste ed evitando costosi errori risultato di infondate assunzioni;
- Il livello di utilizzo del sistema: il numero di utenti interni ed esterni è crescente nel tempo ed il numero di interrogazioni del sistema presenta una considerevole regolarità nell’erogazione temporale;
- L’impatto organizzativo del sistema: il DW RGS ed il SID hanno indotto un cambiamento nell’operatività degli utenti ed un miglioramento della qualità delle loro decisioni. Si è infatti attivato un processo di avvicinamento dell’utente al dato come risultato di un processo di chiarificazione dei contenuti e di condivisione degli stessi che, facilitando la comunicazione interna, ha reso trasparenti le logiche ed i significati favorendo un processo di integrazione procedurale ed organizzativa;

- Il rispetto della pianificazione: i progetti che portano all'integrazione della nuova informazione richiesta dall'utente nell'EDW ed alla sua distribuzione attraverso i DM ed il SID, vengono rilasciati nel pieno rispetto dei tempi non solo grazie al processo di sviluppo incrementale ed evolutivo adottato ma, soprattutto, alla struttura organizzativa (BICC) dedicata alla gestione ed al governo del DW;
- Il rispetto del budget: il totale rispetto del budget è garantito dalla centralità del "Piano IT" concordato tra l'Amministrazione e la Consip. E' infatti in accordo a tale piano che sono state stimate le richieste "conoscitive" dell'utente ed assegnata la loro realizzazione a fornitori nelle more dei contratti affidati e degli standard in uso;
- Il livello di soddisfazione degli utenti: la possibilità di sviluppare in autonomia e in perfetta aderenza alle proprie esigenze conoscitive, analisi sofisticate che possono rilevare aspetti non conosciuti ed utili per l'evoluzione dei processi decisionali, rappresenta per l'utente un beneficio appurato in termini di maturazione professionale e di affinamento di metodiche;
- La richiesta di nuove funzionalità e di nuove informazioni: nel corso degli anni gli utenti del sistema hanno costantemente richiesto la disponibilità di nuova informazione come pure l'evoluzione delle funzionalità d'analisi su informazioni già disponibili; nello stesso tempo si è verificato un costante aumento dei bacini di utenza (Ispettorati) che hanno visto il loro ingresso come fruitori del sistema;
- Il raggiungimento degli obiettivi: possono dirsi raggiunti tutti i principali obiettivi posti, quali la generazione di un archivio univoco e certificato dell'informazione (EDW), la disponibilità di evolute e condivise funzionalità a supporto delle decisioni (DM), e l'integrazione di informazioni in un SID per il controllo, la misura ed il governo degli obiettivi propri della RGS;

E' sulla base di tali indicatori che risulta evidente il successo dei sistemi di BI della RGS sia dal punto di vista "economico", visti i risparmi derivanti dall'ottimizzazione del processo di acquisizione e trattamento delle informazioni, sia dal punto di vista "culturale" in quanto tali sistemi contribuiscono ad un reale sviluppo delle professionalità e della sinergia fra le strutture aziendali a fronte della condivisione del patrimonio informativo.

La partecipazione e il coinvolgimento di tutta l'utenza sono state la chiave di successo dell'iniziativa avviata in un processo di progressiva valorizzazione del patrimonio informativo aziendale inteso quale risorsa strategica per l'efficacia dei processi operativi e, nei casi più evoluti di utilizzo dei DW, quale "prodotto" fruibile dall'utente .

5 Le prospettive

Quali sono le sfide e gli obiettivi che si pongono al DW RGS nel prossimo futuro?

Una delle direzioni di sviluppo del DW RGS è quella del “Knowledge Management”, per rendere sempre più esplicita e maggiormente fruibile la conoscenza maturata, in ambito professionale, da ciascun utente e per metterla a disposizione della RGS al fine di poterne trarre il massimo vantaggio.

Il numero sempre più elevato di sorgenti informative e di informazioni da trattare, strutturate e non (banche dati e documenti), rende particolarmente sfidante la fase di gestione della conoscenza che, in mancanza di tecniche di ricerca evolute, sarebbe impossibile sintetizzare in maniera utile. Per questa ragione sono previste grosse evoluzioni in ambito Metadata Management. Occorrerà un “repository” di metadati e tools sempre più articolato e complesso. Solo con l’utilizzo di metodiche innovative, quali le ontologie, sarà forse possibile fruire tale complesso di informazioni in modo chiaro e significativo abbreviando i tempi di lavoro.

Un altro ambito di evoluzione potrà riguardare l’introduzione di nuove metodologie di ricerca semantica. Sarà sempre meno importante “dove” e “come” i dati sono “catalogati”, ma sarà sempre più importante associare ad essi delle definizioni significative ed accurate, leggibili e riconoscibili da parte delle funzioni automatiche di ricerca. Sarà anche possibile la costruzione automatica di percorsi in grado di guidare l’utente verso l’informazione cercata.

Sarà fondamentale passare dalle forme attuali di portale, dove le ricerche sono eseguite per stringhe in strutture definite, ad un sito web di tipo semantico per il quale vanno adottate nuove strutture, linguaggi e metodologie di integrazione tra i metadati e i dati coinvolti.

Ciò renderà possibile la pubblicazione non più solo di documenti in formato html, immagini, file multimediali, file in formato “Office”, ma anche di informazioni come la descrizione e i singoli dati atomici (numeri, date, nomi) che all’atto della creazione saranno stati definiti ed inseriti secondo precise regole semantiche, seguendo i principi della logica dei predicati (soggetto, predicato, valore) sfruttando così la possibilità di estendere l’elaborazione automatica con costrutti e aumentando la quantità di informazioni interrogabili.

Un altro aspetto di prospettiva futura sarà l’adozione del “data mining”, cioè di tecniche automatizzate di analisi e ricerca, di nuove proprietà e di collegamenti tra dati o gruppi di dati o

tra sorgenti diverse con l'obiettivo di fornire visioni all'interno dei processi di business supportando le attività decisionali.

I sistemi di data warehousing del futuro, infine, risponderanno alle richieste di evoluzione del business avvicinandosi al "real-time-DW". Attualmente tali sistemi sono spesso costruiti come entità separate dai sistemi transazionali, che forniscono loro i dati, e li connotano con obiettivi sbilanciati verso il reporting. Questo fattore li colloca in spazi differenti per i processi di valutazione dell'informazione, ma dovranno presto far parte dei sistemi "mission-critical".

Una delle prospettive, che l'IT consente di prevedere per la RGS, è quella che delinea un aumento del numero dei sistemi gestionali coinvolti nel processo di DW con una conseguente garanzia di informazioni sempre più complete, integrate e certificate. Ciò consentirà alla RGS di utilizzare questo complesso sistema come un concreto sostegno per un'integrazione informativa strutturata con tutti quegli organismi esterni che sono suoi interlocutori nell'ambito di processi istituzionali fondamentali per un sempre più efficace controllo dei conti pubblici.

Il Dimensional Fact Model

La *rapid warehousing methodology* è una metodologia iterativa ed evolutiva per la gestione di progetti di data warehousing, pensata da una delle aziende leader del settore utilizzata da Consip nell'ambito del DW RGS. Essa si basa sulla suddivisione di progetti vasti e potenzialmente rischiosi in sottoprogetti, più piccoli e molto meno rischiosi, chiamati *build*. Ciascun sottoprogetto riprende l'ambiente sviluppato dal sottoprogetto precedente, estendendolo per conseguire nuove funzionalità e facendolo evolvere per far sì che continui a soddisfare i mutevoli bisogni degli utenti. Questo approccio mantiene vivo nel tempo il coinvolgimento e l'interesse degli utenti, gettando le basi per il successo a lungo termine del progetto. La successione di fasi prevista da questa metodologia è illustrata in Figura 19 e brevemente discussa nel seguito.

1. *Accertamento*. L'obiettivo di questa fase è duplice: verificare che l'azienda sia effettivamente pronta ad affrontare il progetto da un lato, determinare gli scopi, i rischi e i benefici dall'altro.
2. *Requisiti*. Consiste nella raccolta delle specifiche di analisi, di progetto e di architettura per l'intero sistema.
3. *Progettazione*. L'attenzione è ora concentrata su un solo sottoprogetto per volta. Le specifiche di analisi vengono raffinate per generare il progetto logico e fisico dei dati e il progetto dell'alimentazione; vengono inoltre selezionati gli strumenti di implementazione.
4. *Costruzione*. Il DW viene implementato e popolato con i dati estratti dalle sorgenti, le applicazioni di front-end vengono sviluppate e collaudate.
5. *Attuazione*. Finalmente il sistema viene consegnato e avviato, dopo che gli utenti sono stati adeguatamente addestrati.
6. *Amministrazione e manutenzione*. Questa fase continua durante tutta la vita del sistema e prevede l'estensione delle sue funzionalità, il ridimensionamento dell'architettura per venire incontro ai nuovi fabbisogni, il controllo della qualità dei dati.
7. *Riesame*. Ciascun sottoprogetto si accompagna a tre processi di riesame: uno a verifica dell'implementazione, uno a seguito dell'attuazione per accertarsi che il sistema sia

stato ben accettato dall'organizzazione, uno a valle dell'intero processo per misurarne i benefici effettivi.

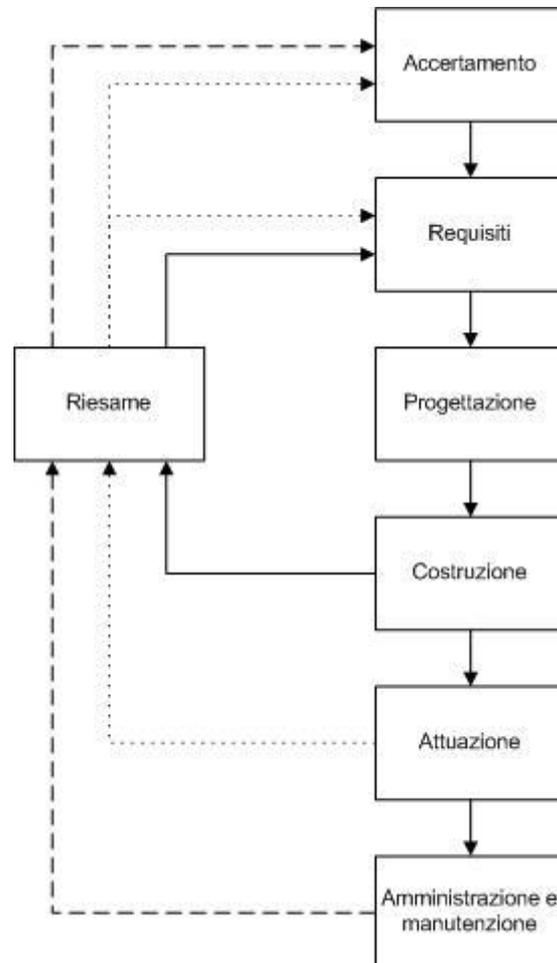


Figura 19

La fase di progettazione su esposta, relativa ad ogni build (si faccia l'astrazione che ogni build corrisponde alla progettazione di un data mart) è a sua volta suddivisa nel seguente modo:

- *Analisi e riconciliazione delle fonti dati.* Questa fase della progettazione richiede di definire lo schema del livello dei dati operazionali a partire dal quale verrà alimentato il data mart, ossia il livello dei dati riconciliato.
- *Analisi dei requisiti.* In questa sessione il progettista raccoglie, filtra e documenta i requisiti degli utenti finali, con l'obiettivo di delineare quali sono le informazioni di interesse da rappresentare in sintonia con gli obiettivi strategici perseguiti. In uscita da

questa fase vengono prodotte specifiche sui fatti da modellare e indicazioni preliminari sul carico di lavoro.

- *Progettazione concettuale*. La progettazione concettuale comporta l'utilizzo dei requisiti utente catturati durante la fase precedente per disegnare, a partire dallo schema riconciliato, uno schema concettuale per il data mart.
- *Raffinamento del carico di lavoro, validazione dello schema concettuale*. Occorre raffinare il carico di lavoro, già espresso in forma preliminare, formulando ciascuna interrogazione direttamente sullo schema concettuale. Ciò permette di verificare che tutte le interrogazioni previste siano effettivamente esprimibili, e quindi in ultima analisi di validare lo schema concettuale prodotto al passo precedente.
- *Progettazione logica*. Requisito essenziale per affrontare la progettazione logica è la scelta del modello logico di riferimento: si tratta cioè di scegliere tra le due seguenti implementazioni:
 - *ROLAP (Relational OLAP)*. L'implementazione su DBMS relazionale memorizza i dati tramite tabelle. Dal punto di vista architetturale, l'adozione della soluzione ROLAP richiede di predisporre un *middleware* specializzato intermedio tra un server back-end relazionale e il lato front-end. Il middleware riceve le interrogazioni OLAP formulate dall'utente sullo strumento di front-end e le traduce in istruzioni SQL per il back-end relazionale, consultando i meta-dati.
 - *MOLAP (Multidimensional OLAP)*. Un sistema MOLAP si basa su un modello logico *ad hoc* sul quale i dati e le operazioni multidimensionali possono essere direttamente rappresentati. Nelle basi di dati multidimensionali, infatti, i dati vengono fisicamente memorizzati in vettore (*array*) e l'accesso è di tipo posizionale.
- *Progettazione dell'alimentazione*. Durante questa fase vengono prese, di concerto con gli utenti e gli amministratori del sistema informatico, tutte le decisioni che riguardano il processo di alimentazione del livello riconciliato. Tra queste, particolarmente importante è la scelta dell'intervallo di aggiornamento periodico del data mart a partire dalle sue sorgenti.
- *Progettazione fisica*. Il problema di maggior rilievo durante la progettazione fisica è la scelta degli indici da costruire per ottimizzare le prestazioni. In questa fase non è più sufficiente aver scelto il modello relazionale come piattaforma a livello logico, occorre anche riferirsi a uno specifico DBMS.

Il modello concettuale adottato da Consip per la progettazione concettuale, del quale si parlerà dettagliatamente in questo capitolo, è il *Dimensional Fact Model*.

Il modello, prevedendo la creazione di uno schema di fatto per ciascun fatto di interesse evidenziato dall'utente, fornirà agli utenti di business una chiara visione delle misure reperibili nel DW RGS. Ciò risulta possibile in quanto la sorgente di informazione *dwrgrs.xml* comprende l'insieme di schemi di fatto dei diversi data mart. Il valore aggiunto ulteriore che dunque si ottiene da tale progetto, deriva dalla conoscenza che ne consegue dalla visione unificata degli schemi di ognuno dei data mart. Si amplia dunque il contesto di conoscenza, cercando di pervenire ad una chiara ottica strategica di fondo, non vincolata ai silos funzionali che potrebbero essere rappresentati, in questo caso, dai diversi dipartimenti della RGS.

1 Concetti di base

Il DFM è un modello concettuale specificamente concepito per fungere da supporto alla progettazione di data mart; è essenzialmente di tipo grafico, e può essere considerato come una specializzazione del modello multidimensionale per applicazioni di data warehousing. Il DFM si pone i seguenti obiettivi:

- Dare supporto efficace alla progettazione concettuale;
- Creare un ambiente in cui le interrogazioni utente possano essere formulate in modo intuitivo;
- Rendere possibile la comunicazione tra il progettista e l'utente finale con l'obiettivo di raffinare le specifiche dei requisiti;
- Costituire una piattaforma stabile per la progettazione logica;
- Fornire una documentazione di progetto espressiva e non ambigua.

La rappresentazione concettuale generata dal DFM consiste in un insieme di *schemi di fatto*. Gli elementi di base modellati dagli schemi di fatto sono i *fatti*, le *misure*, le *dimensioni* e le *gerarchie*. Si riporta nella Figura 20 un esempio di uno schema di fatto.

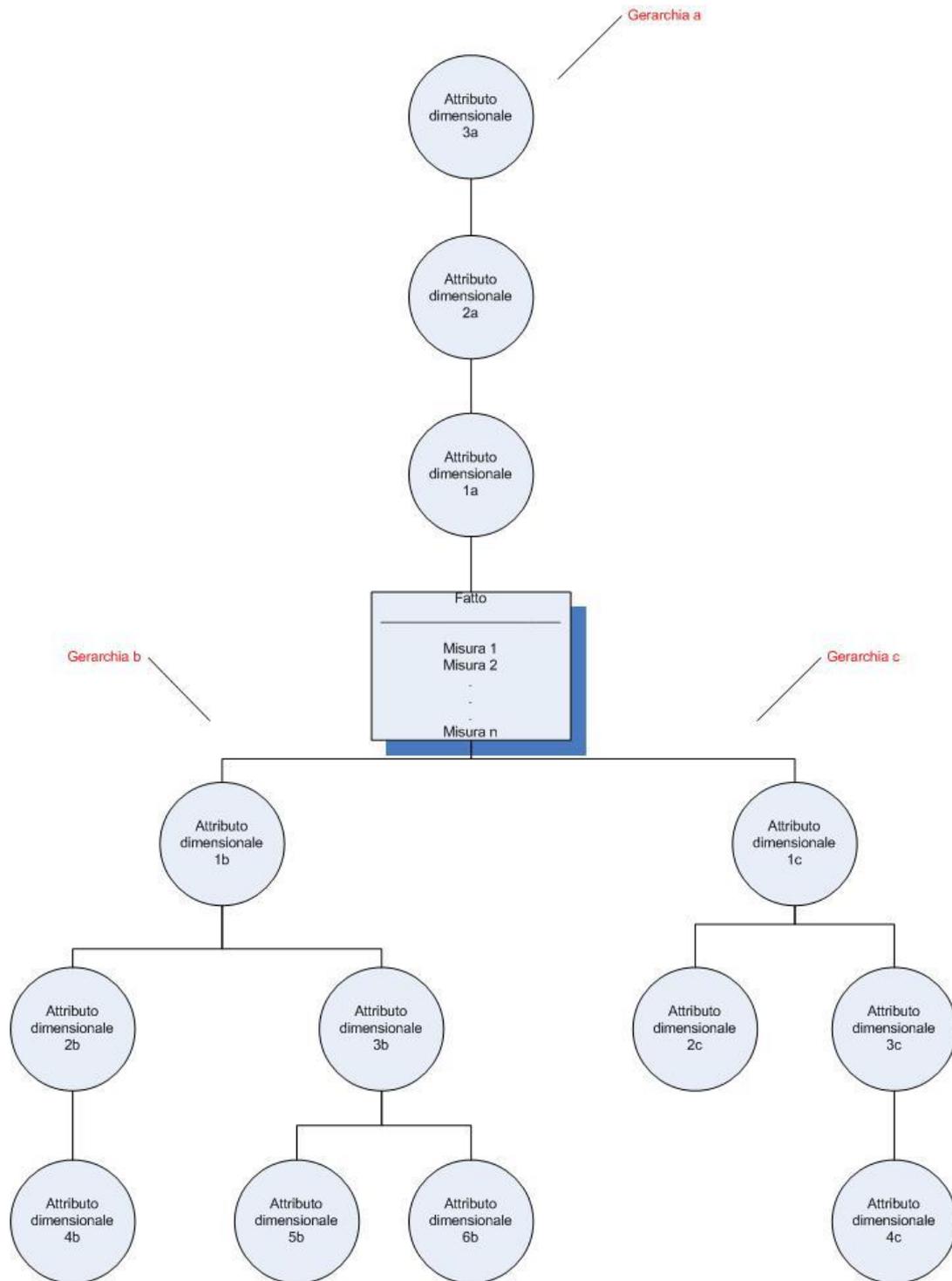


Figura 20

1.1 Fatto

Un *fatto* è un concetto di interesse per il processo decisionale; tipicamente modella un insieme di eventi che accadono nell'impresa.

È essenziale che un fatto abbia aspetti dinamici, ovvero evolva in qualche modo nel tempo. In realtà, pochi tra i concetti rappresentati in una base di dati sono completamente statici. Perciò, la distinzione tra fatti e altri concetti deve essere basata o sulla periodicità media dei cambiamenti, o sugli specifici interessi dell'impresa.

1.2 Misura

Una *misura* è una proprietà numerica di un fatto e ne descrive un aspetto quantitativo di interesse.

Il motivo per cui le misure devono essere preferibilmente numeriche è che esse vengono in genere usate per effettuare calcoli. Un fatto può anche non avere misure, nel caso in cui sia interessante registrare solo il verificarsi di un evento; in questo caso lo schema di fatto viene detto vuoto.

1.3 Dimensione

Una *dimensione* è una proprietà con dominio finito di un fatto e ne descrive una coordinata di analisi.

Un fatto ha in genere più dimensioni, che ne determinano la granularità minima di rappresentazione. Il legame tra misure e dimensioni è espresso, a livello estensionale (ossia a livello di dati piuttosto che di schemi) dal concetto di evento, del quale si darà una definizione successivamente.

1.4 Evento primario

Un *evento primario* è una particolare occorrenza di un fatto, individuata da una ennupla costituita da un valore per ciascuna dimensione. A ciascun evento primario è associato un valore per ciascuna misura.

1.5 Attributo dimensionale

Con il termine generale *attributi dimensionali* si intendono le dimensioni e gli eventuali altri attributi, sempre a valori discreti, che le descrivono.

Le relazioni tra attributi dimensionali sono espresse dalle gerarchie.

1.6 Gerarchia

Una *gerarchia* è un albero direzionato i cui nodi sono attributi dimensionali e i cui archi modellano associazioni molti-a-uno tra coppie di attributi dimensionali. Essa racchiude una dimensione, posta alla radice dell'albero, e tutti gli attributi dimensionali che la descrivono.

Ogni cammino direzionato all'interno di una gerarchia rappresenta a sua volta una dipendenza funzionale tra l'attributo di partenza e quello di arrivo.

Gli attributi dimensionali sono rappresentati da circoletti, e sono collegati da linee che denotano gli archi delle gerarchie esprimendo le regioni di appartenenza. Le gerarchie sono strutturate come alberi con radice nelle dimensioni; pertanto, non è necessario mostrare esplicitamente le direzioni degli archi: ciascuno di essi è implicitamente orientato nella direzione che si allontana dalla radice.

La presenza delle dipendenze funzionali stabilisce un'associazione molti-a-uno tra i valori assunti da una dimensione e quelli assunti da ogni attributo dimensionale che compone la

corrispondente gerarchia. Di conseguenza, ciascuna ennupla di valori di un insieme qualsiasi di attributi dimensionali è associata a un insieme di ennuple di valori delle dimensioni, ossia a un insieme di eventi primari. Ciò rende possibile utilizzare le gerarchie per definire il modo in cui gli eventi primari possono essere aggregati e selezionati significativamente per il processo decisionale; mentre la dimensione in cui una gerarchia ha radice ne definisce la granularità più fine di aggregazione, agli altri attributi dimensionali corrispondono granularità via via crescenti. Questo concetto viene catturato dalla definizione di evento secondario.

1.7 Evento secondario

Dato un insieme di attributi dimensionali (in genere appartenenti a gerarchie distinte), ciascuna ennupla di loro valori individua un *evento secondario* che aggrega tutti gli eventi primari corrispondenti. A ciascun evento secondario è associato un valore per ciascuna misura, che riassume in sé tutti i valori della stessa misura negli eventi primari corrispondenti.

2 Modellazione avanzata

Si introdurranno in questa sezione i concetti di attributo descrittivo; gerarchie condivise e incomplete; archi multipli; dinamicità; additività. Essi descrivono sfumature concettuali che caratterizzano lo scenario del DW RGS.

Altri concetti della modellazione avanzata come gli attributi cross-dimensionali, la convergenza, gli archi opzionali, le gerarchie ricorsive, non vengono riportati in quanto non sfruttati nell'ambito trattato.

Si riporta nella Figura 21 la rappresentazione grafica dei concetti che verranno esposti nel seguito.

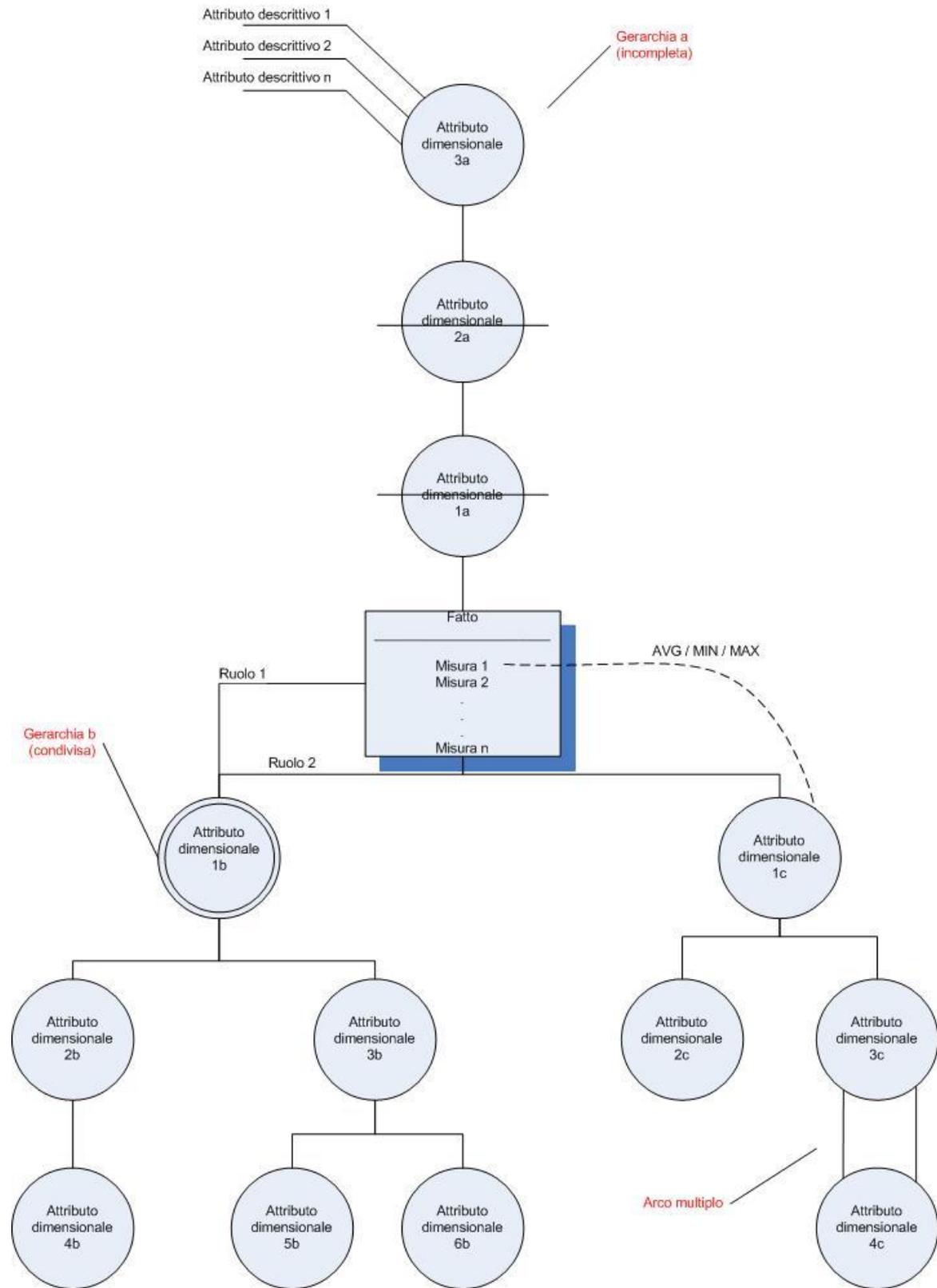


Figura 21

2.1 Attributi descrittivi

In molti casi è importante rappresentare informazioni aggiuntive su un attributo dimensionale presente in una gerarchia, sebbene sia poco interessante utilizzare queste informazioni come criterio di aggregazione.

Un *attributo descrittivo* specifica una proprietà di un attributo dimensionale di una gerarchia, ed è pertanto da quest'ultimo determinato tramite una dipendenza funzionale. Gli attributi descrittivi hanno spesso domini con valori continui; per questo motivo, a differenza degli attributi dimensionali, non possono essere usati per l'aggregazione.

Gli attributi descrittivi sono sempre foglie delle gerarchie e sono rappresentati graficamente nel DFM da linee orizzontali.

Un attributo descrittivo può anche essere connesso direttamente al fatto, se descrive un evento primario ma non è possibile o interessante utilizzarlo né per identificare i singoli eventi né per effettuare calcoli (altrimenti sarebbe, rispettivamente, una dimensione o una misura). Si tenga presente che, affinché sia corretto collegare un attributo descrittivo al fatto, occorre necessariamente che esso assuma, in corrispondenza di ciascun evento primario, un valore univoco.

2.2 Gerarchie condivise

Accade frequentemente negli schemi di fatto che intere porzioni di gerarchie vengano replicate due o più volte. Un esempio classico è costituito dalla gerarchia temporale: un fatto può facilmente avere come dimensioni più attributi di tipo data con significati differenti, e può essere necessario costruire su ciascuno di essi una gerarchia mese-anno; un altro esempio sono le gerarchie geografiche, che possono essere costruite a partire da tutti gli attributi città presenti nello schema.

Si ritiene pertanto consigliabile introdurre una notazione grafica abbreviata che enfatizzi la condivisione delle gerarchie, anche nell'ottica di permettere l'adozione di soluzioni ad hoc di progettazione logica. L'attributo di partenza per la condivisione viene evidenziato raddoppiando

il circolo che lo rappresenta. Si intende sempre, implicitamente, che tutti i discendenti dell'attributo di partenza siano anch'essi condivisi; qualora uno o più discendenti del tratto condiviso non fossero a loro volta condivisi, occorrerà pertanto rappresentare le gerarchie separatamente.

Quando la condivisione ha inizio a livello di dimensione, è indispensabile aggiungere a ciascun arco entrante un *ruolo* che ne specifichi il significato.

2.3 Archi multipli

Nella maggior parte dei casi, interessa inserire all'interno delle gerarchie solo attributi collegati ai loro padri da associazioni multi-a-uno. Esistono però situazioni in cui è utile, o addirittura necessario, includere anche attributi che, in corrispondenza di un singolo valore assunto dall'attributo padre, possano assumere valori multipli.

Si utilizza la notazione illustrata in Figura 21, che dal punto di vista grafico comporta il raddoppio dell'arco entrante.

In generale, il significato di un arco multiplo che va da un attributo a ad un attributo b è che tra a e b esiste un'associazione multi-a-molti, in altre parole che a un singolo valore assunto da a possono corrispondere più valori di b .

2.4 Gerarchie incomplete

Sia a_1, \dots, a_n una sequenza di attributi che forma un percorso all'interno di una gerarchia in uno schema di fatto. Si è dato per scontato che, per ogni valore di a_1 , esista esattamente un valore per ciascuno degli altri attributi del percorso. Ci sono però situazioni, non infrequenti, in cui questa assunzione cade.

Si definisce *gerarchia incompleta* una gerarchia in cui, per alcune istanze, risultano assenti (in quanto non noti oppure non definiti) uno o più livelli di aggregazione. Dal punto di vista grafico, come mostrato in Figura 21, contraddistinguiamo una gerarchia incompleta marcando con un trattino tutti gli attributi i cui valori possono mancare. In letteratura, per designare questo tipo di gerarchia si usa spesso il termine *ragged hierarchy* – letteralmente, gerarchia “sfilacciata”.

2.5 Dinamicità

Il tempo è un fattore chiave nei sistemi di data warehousing, poiché il processo decisionale si basa spesso sulla valutazione di serie storiche e sul confronto tra “fotografie” dell’azienda scattate a istanti diversi.

Fino a questo punto si è ipotizzato che l’unica componente dinamica descritta in uno schema di fatto siano il fatto stesso e gli eventi che lo istanziano; alle gerarchie si è invece attribuita una natura esclusivamente statica. Ciò evidentemente non è del tutto vero.

Si farà riferimento nel seguito di questo paragrafo solo alla dinamicità a livello estensionale, ovvero dei valori: non ci si occuperà quindi delle possibili modifiche che alterano la struttura delle gerarchie (per esempio, l’aggiunta di un nuovo attributo a una gerarchia o l’aggiunta di una nuova dimensione), che non sono da considerarsi fenomeni di routine bensì eventi straordinari legati alla manutenzione del data mart: si parlerebbe in tal caso di dinamicità a livello intensionale.

Tuttavia, tale aspetto svolge un ruolo fondamentale nel progetto in questione. Infatti, una funzionalità che sarà sviluppata nelle successive releases dell’applicazione GAD permetterà agli utenti tecnici di modificare l’architettura del DW RGS. Risulterà più semplice ed immediato adeguare la struttura del DW RGS alle esigenze degli utenti di business che, in un’ottica di mutamento continuo dettato dalle esigenze strategiche della RGS, sono in continua evoluzione, seppur con una periodicità di medio-lungo termine.

La rappresentazione a livello concettuale della dinamicità delle gerarchie è strettamente legata al suo impatto sulle interrogazioni. In presenza di una gerarchia dinamica è infatti possibile ipotizzare quattro differenti scenari temporali di analisi degli eventi.

- *Oggi per ieri*. Tutti gli eventi vengono riportati all'attuale configurazione delle gerarchie.
- *Ieri per oggi*. Tutti gli eventi vengono riportati alla configurazione precedentemente assunta dalle gerarchie.
- *Oggi o ieri*. Ciascun evento viene riportato alla configurazione assunta dalle gerarchie nell'istante di tempo in cui l'evento si è verificato.
- *Oggi e ieri*. Vengono considerati solo gli eventi riferiti alla parte di gerarchia rimasta immutata.

La dinamicità utilizzata nel caso del DW RGS è *oggi per ieri*.

2.6 Additività

L'aggregazione richiede di definire un operatore adatto per comporre i valori delle misure che caratterizzano gli eventi primari in valori da abbinare a ciascun evento secondario. Da questo punto di vista è possibile distinguere tre categorie di misure:

- *Misure di flusso*. Si riferiscono a un periodo, al cui termine vengono valutate in modo cumulativo.
- *Misure di livello*. Vengono valutate in particolari istanti di tempo.
- *Misure unitarie*. Vengono valutate in particolari istanti di tempo, ma sono espresse in termini relativi.

Gli operatori di aggregazione utilizzabili sui diversi tipi di misure, durante l'aggregazione su gerarchie temporali e non, sono riassunti in Tabella 1.

Tabella 1 - operatori di aggregazione

	Gerarchie temporali	Gerarchie non temporali
Misure di flusso	SUM, AVG, MIN, MAX	SUM, AVG, MIN, MAX
Misure di livello	AVG, MIN, MAX	SUM, AVG, MIN, MAX
Misure unitarie	AVG, MIN, MAX	AVG, MIN, MAX

Una misura è detta *additiva* su una dimensione se i suoi valori possono essere aggregati lungo la corrispondente gerarchia tramite l'operatore di somma, altrimenti è detta *non-additiva*. Una misura non-additiva è *non-aggregabile* se nessun operatore di aggregazione può essere usato su di essa.

Dalla Tabella 1 appare evidente che, come regola di massima, le misure di flusso sono additive su tutte le dimensioni, quelle di livello sono non-additive sulle dimensioni temporali, quelle unitarie sono non-additive su tutte le dimensioni.

Essendo l'additività il caso più frequente, al fine di semplificare la notazione grafica nel DFM si rappresentano esplicitamente solo le eccezioni. In particolare, come mostrato in Figura 21, una misura viene collegata alle dimensioni su cui è non-additiva tramite una linea tratteggiata etichettata con gli eventuali operatori di aggregazione utilizzabili. Se una misura presenta lo stesso tipo di additività su tutte le dimensioni, l'operatore di aggregazione può esserle riportato a lato. Qualora il numero complessivo di non-additività fosse tale da ridurre la leggibilità dello schema, è consigliabile adottare una rappresentazione tabellare.

3 DFMCASE

Il DFMCASE è il software tramite il quale sono stati realizzati gli schemi, secondo il DFM, dei data mart costituenti il DW RGS.

Esso permette la visualizzazione grafica di dettaglio, nonché la generazione di report di analisi, in riferimento alle seguenti componenti di uno schema DFM:

- Gerarchie;
- Misure;
- Schemi di fatto;
- Attributi descrittivi.

Si riporta in Figura 22, a titolo di esempio, la visualizzazione tramite il DFMCASE dello schema di fatto *Impegno dati annuali*.

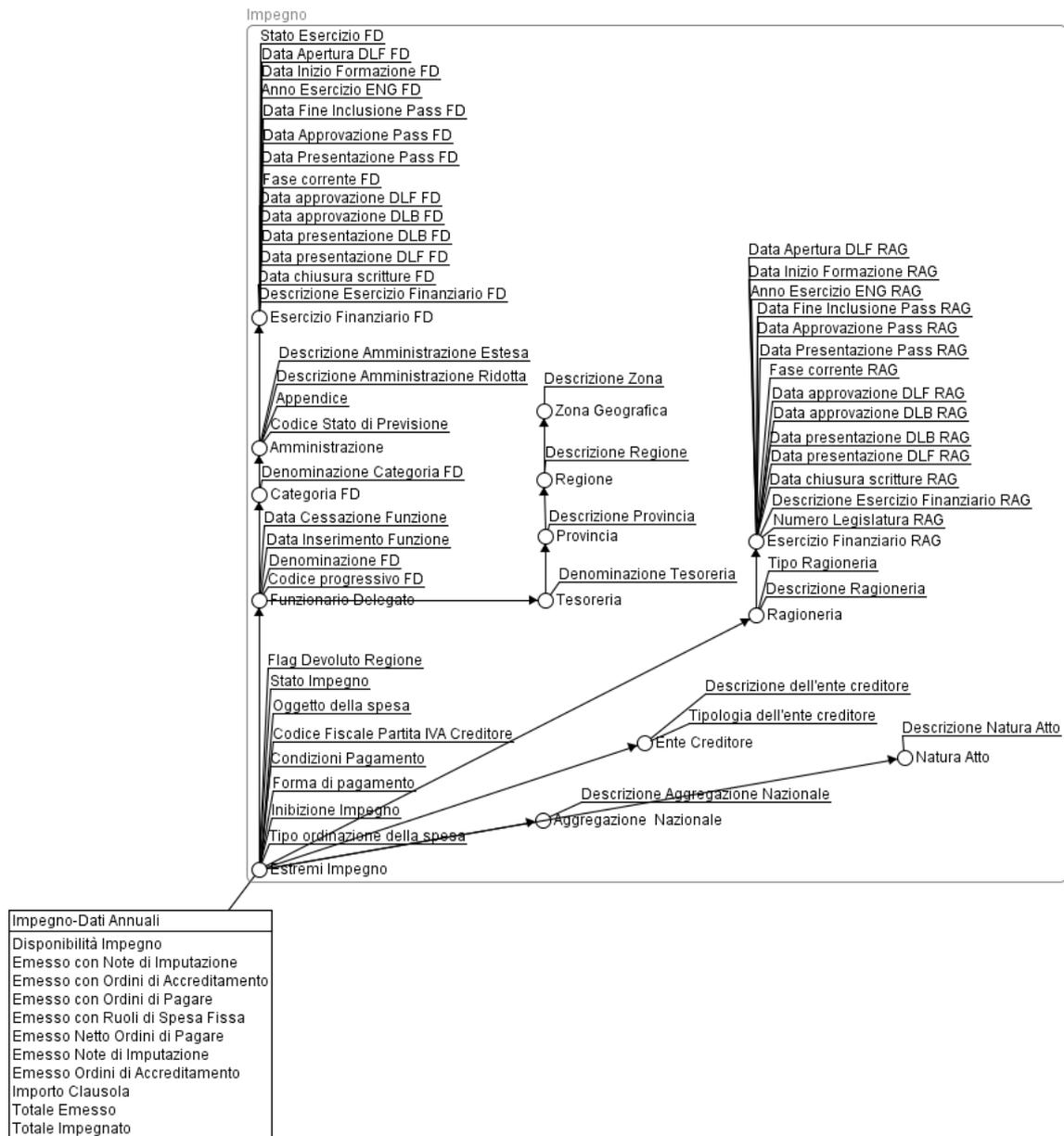


Figura 22

Il rettangolo in basso a sinistra riporta l'insieme delle misure associate allo schema di fatto, valutabili secondo gli attributi dimensionali contenuti nella gerarchia *Impegno*, rettangolo in alto a destra.

Il DW RGS è costituito di un gran numero di schemi di fatto, la maggior parte dei quali avente dimensioni ben più rilevanti rispetto allo schema di fatto su riportato.

Inoltre, il DFM non permette una visione di più alto livello rispetto alla semplice consultazione della struttura di una qualunque gerarchia. In altre parole, non è possibile osservare le correlazioni esistenti fra schemi di fatto appartenenti allo stesso schema concettuale associato ad un determinato data mart.

Risulta dunque ora più chiaro come si renda necessaria l'integrazione dell'informazione dei DFM dei vari data mart in modo da permettere una visione d'insieme dell'intero contenuto informativo, permettendo allo stesso tempo di focalizzare l'attenzione su particolari componenti dell'intero data warehouse.

Nel DFM, fatti diversi vengono rappresentati da schemi di fatto distinti. D'altronde, parte delle interrogazioni formulate dall'utente possono richiedere la comparazione di misure prese da più fatti tra loro correlati; nella terminologia OLAP, queste interrogazioni vengono dette di drill-across. Tali interrogazioni diventano possibili combinando due o più schemi di fatto in un nuovo schema. Questa operazione, detta *sovrapposizione di schemi di fatto*, può essere operata anche su schemi di fatto appartenenti a modelli concettuali differenti. Ecco il punto chiave del progetto: applicando la sovrapposizione di schemi di fatto alle entità presenti negli schemi dei data mart costituenti il DW RGS si ottiene un unico schema in grado di rappresentare in un unicum i metadati relativi al data warehouse tutto.

4 Sovrapposizione di schemi di fatto

Gli schemi di fatto possono essere sovrapposti quando sono confrontabili (hanno lo stesso significato semantico) e comparabili (sintatticamente equivalente).

In uno schema sovrapposto:

- le misure sono l'unione delle misure presenti nei vari schemi;
- le gerarchie includono solo gli attributi presenti in entrambe le gerarchie;

- il dominio di ogni attributo è l'intersezione del dominio dei corrispondenti attributi sugli schemi di origine.

Tuttavia, lo scopo finale che l'applicazione si propone è quello di fornire una rappresentazione unica dei metadati contenuti del data warehouse. Poiché l'applicazione non deve occuparsi dell'ottimizzazione delle interrogazioni drill-across, si è scelto dunque di modificare il concetto di sovrapposizione degli schemi di fatto. Il comportamento sarà differente relativamente alla gerarchia che si ottiene dalla sovrapposizione, ovvero sia, l'insieme degli attributi della gerarchia sovrapposta sarà l'unione degli attributi di tutte le gerarchie concorrenti alla sovrapposizione, non l'intersezione. Questo comporta, a livello estensionale (istanze che popolano il data warehouse) la creazione di attributi dimensionali incompleti, ossia, la possibilità che un attributo dimensionale risulti non valorizzato. Tale problema nel caso del progetto in esame non sussiste in quanto il dominio dei dati trattati non riguarda il livello estensionale ma solamente quello intensionale, ossia l'insieme dei metadati relativi alla struttura del DW RGS. L'operazione di sovrapposizione implica la necessità di effettuare la riduzione aciclica del grafo originato per via della presenza di archi ridondanti ottenuti dalla fusione delle gerarchie.

Si riporta in Figura 23 un esempio specifico per una migliore comprensione del procedimento su esposto.

Schema di fatto 1 e Schema di fatto 2 possono essere sovrapposti poiché le gerarchie collegate hanno il medesimo significato e non ci sono conflitti fra le dipendenze funzionali che rappresentano.

La riduzione aciclica ha eliminato l'arco ridondante che collegava Città con Stato (Stato è comunque raggiungibile attraverso gli archi che collegano Città con Provincia e Provincia con Regione).

Quando i ruoli delle dimensioni coinvolte nella sovrapposizione sono differenti, lo schema risultante dalla sovrapposizione contiene tanti archi quanti sono i ruoli della dimensione, quindi quest'ultima diventa condivisa (doppio circoletto).

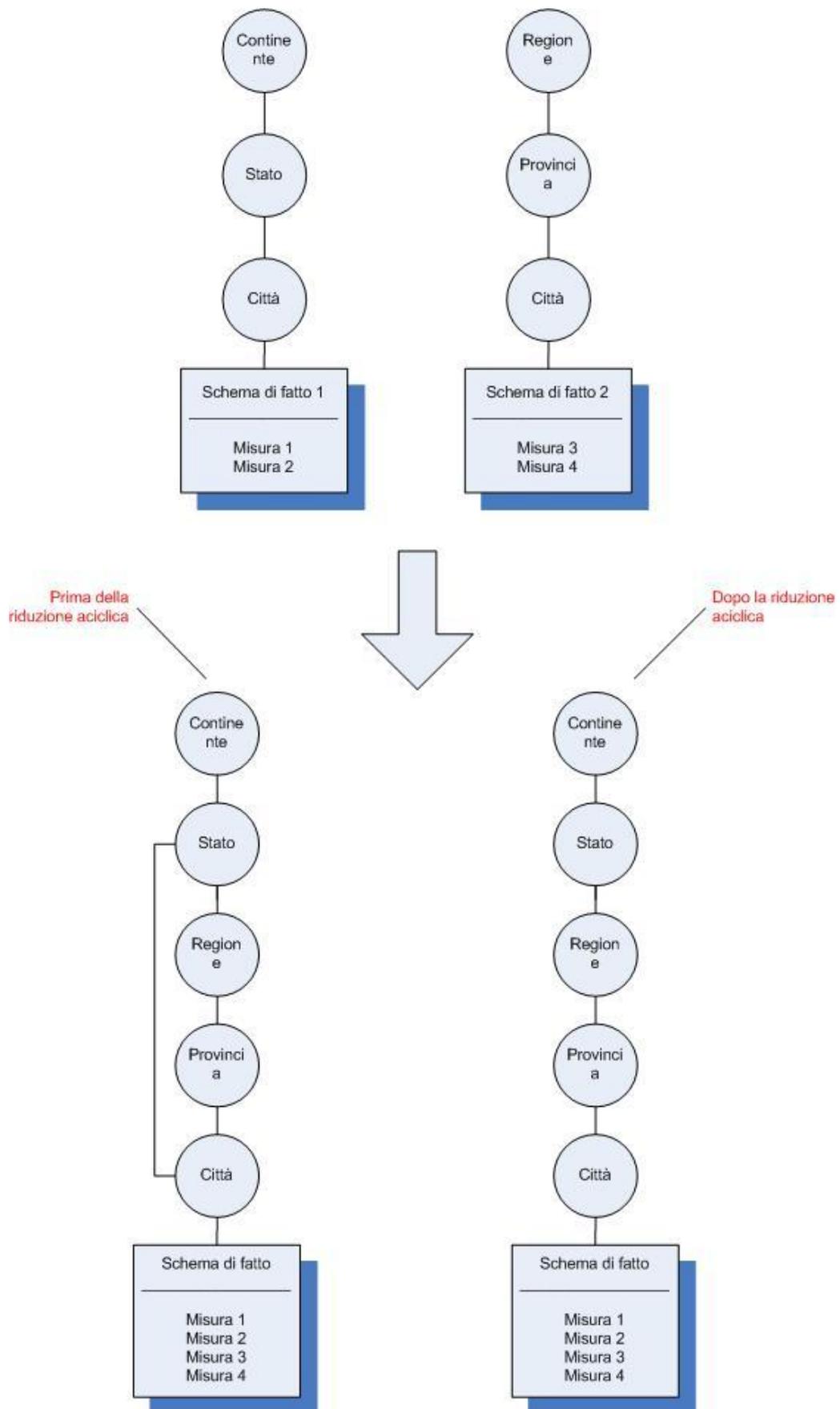


Figura 23

5 Dominio dei dati trattati

Per motivi di semplificazione della rappresentazione dell'informazione, l'applicazione GAD non riporterà alcune informazioni che fanno parte del formalismo del DFM.

In particolare, i concetti utilizzati dall'applicazione saranno:

- Attributo dimensionale;
- Attributo dimensionale condiviso;
- Misura;
- Gruppo di Misure: esso corrisponde ad uno schema di fatto ottenuto da sovrapposizione;
- Attributi descrittivi;
- Funzioni di aggregazione.

Gli altri concetti su esposti non vengono implementati in quanto non necessari o comunque superflui poiché esulano la finalità dell'applicazione oggetto di realizzazione.

Si perde ad esempio il concetto di evento. Un evento è associato ad uno schema di fatto, e dopo la sovrapposizione lo schema di fatto risultante conterrà le misure degli schemi che hanno concorso alla sovrapposizione. Ciò non risulta essere un problema in quanto non si devono effettuare interrogazioni sui dati ma si vogliono mostrare le correlazioni esistenti fra le misure utilizzate nel DW RGS.

L'applicazione: GAD

Il presente Capitolo è il risultato concreto delle attività di analisi e progettazione relativamente al progetto in questione. Si riporta dunque la documentazione ufficiale redatta fino alla data 27 aprile 2012, costituita dalla Specifica Funzionale completa e dalla parte del Manuale Utente riferita alle funzionalità implementate dal prototipo del software realizzato.

La Specifica Funzionale comprende:

- Scopo del progetto;
- Scelta della soluzione, con particolare riferimento all'architettura generale del sistema;
- Descrizione dei dati;
- Specifica dei requisiti, la quale contiene:
 - Requisiti funzionali;
 - Requisiti non funzionale;
 - Diagramma generale dei casi d'uso;
 - Casi d'uso;
 - Diagramma di sequenza;
 - Diagramma delle classi;

Il Manuale Utente riporterà esempi di utilizzo del software limitatamente alle schermate *Elaborazioni e Visualizzazione query*.

Un ulteriore paragrafo riporterà quelli che potrebbero essere gli sviluppi futuri dell'applicazione nelle nuove release che saranno rilasciate.

1 Specifica funzionale

1.1 Scopo del progetto

1.1.1 Descrizione del progetto

Il presente scritto è il risultato dell'attività di analisi e progettazione nell'ambito del progetto GAD (Grafo Attributi Dimensionali) svolto dalla Consip S.p.A per il committente rappresentato dal Ministero dell'Economia e delle Finanze.

Il presente documento si prefigge lo scopo di dettagliare i contenuti dell'intervento per la realizzazione di un componente applicativa che permetta l'esportazione di dati relativi a un grafo in uno dei formati prestabiliti.

Il dettaglio delle funzionalità, formalizzate nel presente documento, è definito sulla base delle analisi svolte dal Gruppo di Lavoro ed a fronte degli incontri avuti con i referenti della Consip S.p.A.. Il committente è rappresentato dal Ministero dell'Economia e delle Finanze.

1.1.2 Tipologie di utenti

Le tipologie di utenti utilizzatrici del sistema in esame sono le seguenti:

- Utente di business: rappresentato da dirigenti e funzionari della Pubblica Amministrazione e della Ragioneria Generale dello Stato (RGS), avrà la possibilità di navigare (sfruttando il paradigma del modello dimensionale) i metadati relativi alle misure e alle dimensioni di analisi oggetto di interesse per l'amministrazione. La visione integrata ed omogenea dei concetti di interesse e delle relazioni esistenti tra esse sarà di ausilio all'utente nella comprensione della tematica di interesse e nella individuazione delle informazioni necessarie a supportare i processi decisionali;
- Utente tecnico: rappresentato dai capi progetto Consip e i fornitori responsabili dello sviluppo nelle diverse fasi del ciclo di vita del software, adopererà il sistema come strumento di ausilio per:

- progettazione concettuale dei Data Mart del DW RGS;
- progettazione logica dei Data Mart del DW RGS;
- progettazione dello strato semantico a supporto dell'utente di business.

Tuttavia, alla differenziazione della tipologia di utenza non corrisponde una differenziazione in termini di accessibilità alle funzioni messe a disposizione dal software. In altre parole, entrambi gli utenti potranno accedere alla totalità delle funzioni messe a disposizione.

L'unico requisito necessario è costituito dal possesso delle credenziali per l'accesso al sistema tramite il Portale SAP BO. Tuttavia, ciò esula l'ambito del presente progetto, dunque il gruppo di progetto GAD non dovrà gestire tale aspetto.

Ove necessario nel seguito si farà la semplificazione di riferirsi ad entrambe le categorie di utenti come se essi afferissero ad un'unica categoria.

1.1.3 Glossario definizioni ed acronimi

Si elencano e si definiscono, utilizzando la Tabella 2, tutti i termini e gli acronimi utilizzati nel documento, al fine di fornirne una definizione rigorosa e non ambigua.

Tabella 2 - definizioni ed acronimi

Termine / Acronimo	Descrizione
GAD	Grafo Attributo Dimensionali
GADWeb	Parte Web del progetto
GraphXMLGenerator	Convertitore di protocollo dati
DFMCase	Tool alimentante GraphXMLGenerator
DFMX file	Protocollo proprietario dello strumento DFMCase utilizzato in ambito PA per la modellazione del grafo mediante il formalismo DFM
GAD files	Protocollo proprietario dello strumento GADWeb utilizzato in ambito PA per la visualizzazione tramite browser del grafo
GraphML	E' un formato file basato su XML per grafi
Login	Prima schermata di GADWeb
Contesti	Seconda schermata di GADWeb
Elaborazioni	Terza schermata di GADWeb
Visualizzazione query	Quarta schermata di GADWeb
Vista generale/Vista di dettaglio	Pannello della schermata Elaborazioni
Ricerca avanzata	Pannello della schermata Elaborazioni
Dettaglio elemento	Pannello della schermata Elaborazioni
Grafo principale	Pannello della schermata Elaborazioni

1.2 Scelta della soluzione

1.2.1 Premessa

La soluzione adottata, GADWeb, sarà fruibile mediante browser Web. Ciò rende possibile l'utilizzo dell'applicazione da qualsiasi postazione multimediale avente caratteristiche architettoniche e applicative comuni a tutti i Personal Computer. Diviene a tal proposito necessario introdurre dei sistemi autenticazione dell'utenza finale.

L'applicazione attinge i contenuti informativi necessari in modo indiretto dal tool DFMCASE. Tali contenuti non sono direttamente fruibili in quanto la loro struttura (DFMX file) non può rappresentare, se non attraverso una elaborazione specifica, il modello dei dati utilizzato dall'applicazione (GAD files). Si rende necessaria a tal proposito la realizzazione di un'ulteriore componente applicativa denominata GraphXMLGenerator, trasparente all'utente finale, la quale si occuperà di trasformare il formato dei dati in un modello coerente con lo scopo dell'applicativo finale.

Tale impostazione soddisfa i requisiti funzionali REF001 e REF020 ed i requisiti non funzionali RNF001, RNF002 e RNF006.

1.2.2 Architettura del sistema

Si riportano in Figura 24 gli aspetti architettonici della soluzione realizzata.

Si descriveranno nei due sottoparagrafi seguenti le due macrocomponenti .

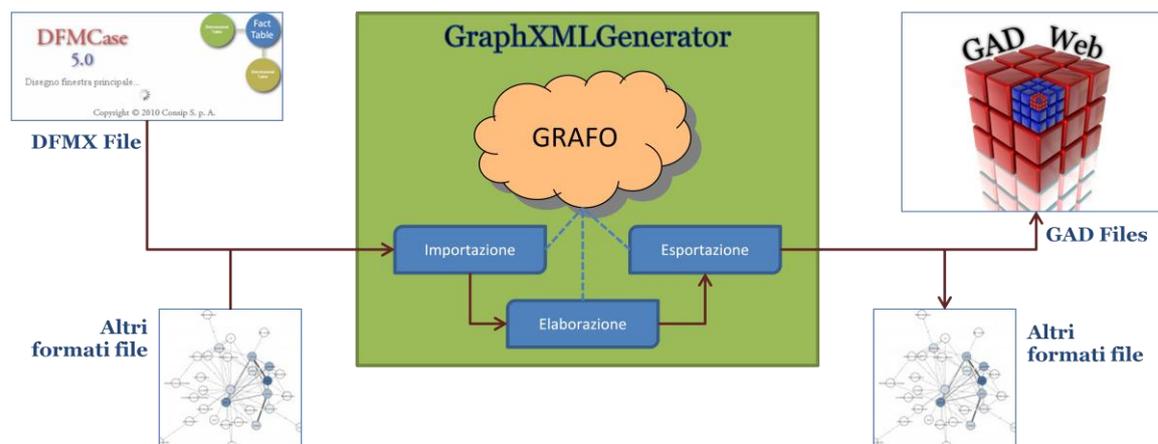


Figura 24

1.2.2.1 GraphXMLGenerator

Per rendere la soluzione più generale ed adattabile a sviluppi futuri, la componente GraphXMLGenerator sarà in grado di recepire in ingresso altri formati di file in aggiunta al formato esportato dal DFMCase (DFMX file), e potrà generare in uscita altri formati di file in aggiunta al formato esportato (GAD files). In particolare, un altro formato che potrà essere manipolato è il GraphML.

L'esecuzione della componente, ad opera di un utente tecnico, avviene da riga di comando secondo la sintassi prestabilita:

```
java -jar GraphXMLGenerator.jar [action] [filein] [fileout] [options]
```

I parametri di configurazione corrispondono a:

- Action: specificare l'azione di conversione da eseguire. Da formato file X a formato file Y. I valori possibili sono: DFMCase2GRAPHML, DFMCase2GAD, GRAPHML2GAD.
- Filein: specificare il file di lettura su cui recuperare i dati del grafo.
- Fileout: specificare il file di scrittura su cui generare il grafo.
- Options: specificare gli algoritmi o funzionalità da applicare al grafo. I valori possibili sono:

- -rt: Riduzione transitiva
- -ct: Chiusura transitiva
- -e: Abilita la gestione dei parametri di estensione.

Le operazioni eseguite dal batch ad ogni ciclo (inteso come lavorazione della singola richiesta) sono le seguenti:

- Verifica configurazione e predisposizione alla lavorazione;
- Validazione del file d'ingresso;
- Importazione del grafo dal file specificato;
- Elaborazione del grafo;
- Esportazione del grafo nel file specificato;
- Validazione del file d'uscita.

Il ciclo di lavorazione di un grafo è composto dalle tre macro-fasi rappresentate in Figura 25:



Figura 25

La fase di Importazione consiste nell'acquisizione del grafo in uno dei formati file messi a disposizione dalla componente applicativa. La struttura modulare del software permette facilmente l'implementazione di nuovi formati.

La fase di Elaborazione consiste nell'eventuale attuazione sul grafo di uno o più algoritmi (o servizi) messi a disposizione dalla componente applicativa. La struttura modulare del software permette facilmente l'implementazione di nuovi formati.

La fase di Esportazione consiste nella generazione del grafo in uno dei formati file messi a disposizione dalla componente applicativa. La struttura modulare del software permette facilmente l'implementazione di nuovi formati.

Tale schematizzazione è generale e applicabile a tutti i formati che i grafi potranno assumere: dfmx, Graphml, etc.

Tale operazione si renderà necessaria ogni qual volta la struttura del grafo subirà dei cambiamenti.

1.2.2.2 GADWeb

L'applicazione Web si compone delle seguenti schermate:

- *Login*. È la prima schermata che l'utente visualizza. In tale contesto dovranno essere immesse le credenziali necessarie all'autenticazione per accedere alla schermata successiva, *Contesti*.
- *Contesti*. L'utente visualizza e seleziona i contesti sui quali effettuare lo studio, il quale potrà essere sviluppato grazie alle funzioni messe a disposizione nella schermata *Elaborazioni*.
- *Elaborazioni*. Tale schermata è graficamente suddivisa nei seguenti pannelli:
 - *Vista generale*: visualizza una panoramica del grafo;
 - *Ricerca avanzata*: permette la ricerca per nome di Attributi Dimensionali e Misure;
 - *Grafo Principale*: visualizza il grafo con un livello di dettaglio personalizzato dall'utente in termini di zoom e layout;
 - *Dettaglio elemento*: in seguito alla selezione di un elemento, nel grafo o nella lista testuale ottenuta tramite ricerca per nome, visualizza informazioni aggiuntive dell'elemento selezionato. Permette inoltre di redigere una query personalizzata che sarà visibile nella schermata *Visualizzazione query*.
- *Visualizzazione query*. Mostra il risultato della query sottoposta dall'utente mediante forma tabellare paginata.

1.2.3 Vincoli

I vincoli individuati fanno riferimento ai requisiti funzionali e non, emersi nel corso delle riunioni preliminari. Di seguito si riportano brevemente per punti tali vincoli:

- Il sistema operativo a disposizione sull'ambiente di collaudo/esercizio ha JDK 1.6;
- Sarà possibile sull'ambiente di collaudo/esercizio avere uno spazio su file system dove salvare file di properties, file di log e quant'altro ritenuto necessario;
- La libreria grafica per la renderizzazione del grafo è MxGraph;
- La componente applicativa deve prevedere l'esecuzione sul grafo della riduzione transitiva;
- La componente applicativa deve poter importare il grafo dallo strumento di modellazione DFMCASE (REF020);
- I dati del grafo devono poter supportare eventuali estensioni in termini di informazioni descrittive.

1.3 Specifica dei requisiti

1.3.1 Requisiti funzionali

1.3.1.1 REF001

L'utente deve poter effettuare il login.

1.3.1.2 REF002

L'utente deve poter scegliere i contesti sui quali compiere l'analisi di interesse.

1.3.1.3 REF003

L'utente deve poter navigare il grafo in forma grafica.

1.3.1.4 REF004

L'utente deve poter visualizzare una panoramica del grafo completo.

1.3.1.5 REF005

L'utente deve poter visualizzare il grafo o un suo sottoinsieme attraverso una vista di dettaglio mediante il livello di zoom desiderato.

1.3.1.6 REF006

L'utente deve poter visualizzare le componenti del grafo contenute nella vista di dettaglio attraverso i layout disponibili.

1.3.1.7 REF007

L'utente deve poter stampare il contenuto della vista di dettaglio.

1.3.1.8 REF008

L'utente deve poter selezionare dalla vista di dettaglio gli Attributi Dimensionali e i Gruppi di Misure.

1.3.1.9 REF009

L'utente in seguito alla selezione dalla vista di dettaglio di un Attributo Dimensionale deve poterne visualizzare: in forma testuale, gli Attributi Descrittivi e le Misure con le loro relative Funzioni di Aggregazione; in forma grafica gli Attributi Dimensionali Successori e gli Attributi Dimensionali Predecessori.

1.3.1.10 REF010

L'utente in seguito alla selezione dalla vista di dettaglio di un Attributo Dimensionale deve poterne visualizzare e selezionare l'elenco delle colonne della tabella associata, quindi eseguire la query personalizzata.

1.3.1.11 REF011

L'utente in seguito alla selezione dalla vista di dettaglio di un Gruppo di Misure deve poterne visualizzare: in forma grafica, i Successori; in forma testuale la lista delle Misure.

1.3.1.12 REF012

L'utente in seguito all'invio del comando di esecuzione della query deve poter visualizzare il risultato mediante forma tabellare paginata.

1.3.1.13 REF013

L'utente deve poter effettuare ricerche per nome di Attributi Dimensionali e Misure appartenenti al grafo.

1.3.1.14 REF014

L'utente deve poter selezionare un Attributo Dimensionale, risultato della ricerca, per ottenerne: in forma testuale, gli Attributi Descrittivi e le Misure con le loro relative Funzioni di Aggregazione; in forma grafica, gli Attributi Dimensionali Successori e gli Attributi Dimensionali Predecessori.

1.3.1.15 REF015

L'utente deve poter selezionare un Attributo Dimensionale, risultato della ricerca, per poterne visualizzare e selezionare l'elenco delle colonne della tabella associata, quindi eseguire la query personalizzata.

1.3.1.16 REF016

L'utente deve poter selezionare una Misura, risultato della ricerca, per ottenerne: in forma grafica gli Attributi Dimensionali e gli Attributi Dimensionali Successori; in forma testuale gli Attributi Descrittivi, gli Attributi Dimensionali con le Funzioni di Aggregazione e la lista delle altre Misure facenti parte del Gruppo di Misure della Misura stessa.

1.3.1.17 REF017

L'utente deve poter stampare la visualizzazione grafica degli Attributi Dimensionali Successori e degli Attributi Dimensionali Predecessori di cui ai requisiti funzionali REF014 e REF016.

1.3.1.18 REF018

L'utente deve poter applicare il livello di zoom desiderato alla visualizzazione grafica degli Attributi Dimensionali Successori e degli Attributi Dimensionali Predecessori.

1.3.1.19 REF019

L'utente deve avere a disposizione, per la visualizzazione grafica del grafo o di un suo sottoinsieme, di cui ai requisiti funzionali REF006, REF009, REF011, REF014, REF016, i seguenti layout: verticale gerarchico, orizzontale gerarchico, verticale ad albero, orizzontale ad albero, organico e circolare.

1.3.1.20 REF020

L'applicazione deve attingere le informazioni di interesse dal tool DFMCASE.

1.3.2 Requisiti non funzionali

1.3.2.1 RNF001

Il Sistema deve poter essere accessibile da un personal computer avente le seguenti caratteristiche hardware minime:

- Processore: dual core 2.0 GHz;
- Ram: 4 GB.

1.3.2.2 RNF002

Il sistema deve poter essere accessibile dal browser Internet Explorer 9 e deve essere compatibile con i browser più diffusi quali Mozilla Firefox, Google Chrome.

1.3.2.3 RNF003

Manutenibilità: deve essere possibile apportare facilmente modifiche al sistema realizzato in previsione di miglioramento della qualità del software, eliminazione di eventuali errori, introduzione di nuove funzionalità.

1.3.2.4 RNF004

Usabilità: il sistema deve essere di facile comprensione e utilizzo anche a utenti che non hanno un'elevata conoscenza delle tecnologie informatiche.

1.3.2.5 RNF005

Robustezza: il sistema deve tornare in uno stato consistente nel caso in cui si verificano situazioni non previste dalle specifiche (dati di input non corretti, carico di lavoro elevato, etc.).

1.3.2.6 RNF006

Sicurezza: bisogna assicurarsi che l'utente abbia accesso a tutti e soli i servizi a lui concessi.

1.3.2.7 RNF007

Confidenzialità: i dati e le informazioni scambiate tra client e server devono essere protetti in modo da garantire la privacy degli utenti e/o evitare che terze parti possano entrarne in possesso.

1.3.2.8 RNF008

Integrità: è necessario ridurre il rischio di alterazione dei dati durante la loro trasmissione o memorizzazione, sia essa accidentale o effettuata da terze parti.

1.3.3 Diagramma generale dei casi d'uso

Il modello di business dei casi d'uso individuati con eventuali flussi secondari attivati è descritto in Figura 26.

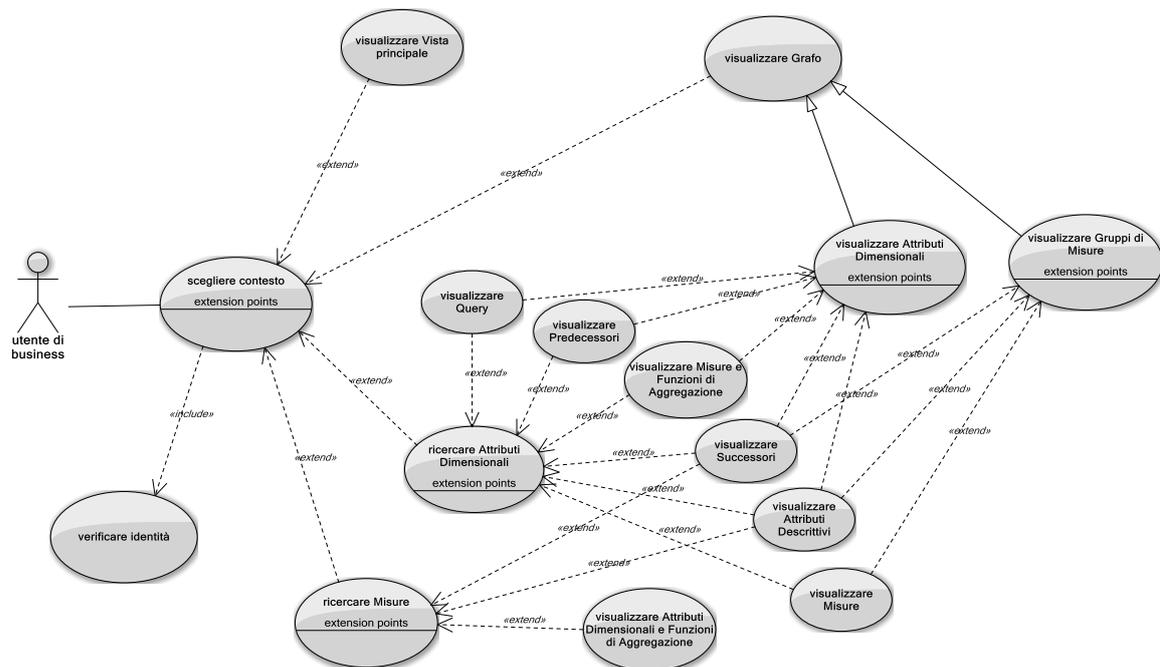


Figura 26

Nel seguito si fornirà una descrizione dettagliata di ogni caso d'uso.

1.3.4 Casi d'uso

1.3.4.1 CDU001 - verificare identità

L'utente per poter accedere all'applicazione deve effettuare il login; dovrà accedere alla pagina di benvenuto del Portale BO, inserendo le proprie credenziali (username e password), quindi accedere all'applicazione "Grafo Attributi Dimensionali" attraverso il link omonimo.

Questa operazione si rende necessaria per il soddisfacimento dei requisiti non funzionali RNF006 (Sicurezza) e RNF007 (Confidenzialità).

Tabella 3 - CDU001: verificare identità

Caso d'uso: CDU001	verificare identità	Data: 17/04/2012
		Versione: 1.0
Requisito funzionale	REF001	
Descrizione	L'utente di business deve poter effettuare il login	
Priorità	Normale	
Attori	Utenti di business e utenti tecnici	
Flusso principale		
1 - L'utente accede al Portale BO		
2 - L'applicazione mostra due caselle di testo che andranno compilate con username e password dell'utente		
3 - L'utente inserisce le proprie credenziali		
4 - Il Portale BO dirige l'utente alla schermata Contesti		

1.3.4.2 CDU002 - scegliere contesto

L'utente, una volta acceduto all'applicazione, ha a disposizione una schermata dalla quale potranno essere scelti i contesti su cui effettuare lo studio. Ad ogni contesto è associato un sottografo: l'insieme di tutti i sottografi compone il grafo completo del DW RGS.

Ogni contesto è rappresentato da un box, la cui dimensione è proporzionale al numero di attributi in esso contenuti.

Ogni box è indipendentemente selezionabile e deselegonabile. Per aiutare l'utente nella conoscenza del dominio dei dati trattati, la selezione di un box provoca l'evidenziazione dello stesso e di tutti quegli altri contesti raggiungibili da un qualsiasi elemento del contesto selezionato inizialmente.

Definiti gli ambiti sui quali proseguire lo studio, cliccando sul bottone "Inizia", l'utente accede alla schermata successiva, la quale permette le operazioni di cui ai casi d'uso CDU003, CDU004, CDU005 e CDU006.

Tabella 4 - CDU002: scegliere contesto

Caso d'uso: CDU002	scegliere contesto	Data: 17/04/2012
		Versione: 1.0
Requisito funzionale	REF002	
Descrizione	L'utente di business deve poter scegliere i contesti sui quali compiere l'analisi di interesse	
Priorità	Normale	
Attori	Utenti di business e utenti tecnici	
Precondizioni	Login dell'utente (CDU001)	
Flusso principale		
1 - L'utente visualizza i contesti		
2 - L'utente seleziona i contesti su cui proseguire l'analisi		
3 - L'applicazione aiuta l'utente nella selezione dei contesti evidenziando le correlazioni fra i contesti selezionati		
4 - L'utente sottomete al sistema il comando di generazione del grafo ed accede alla schermata successiva		
5 - L'applicazione reindirige l'utente alla schermata di Elaborazione		

1.3.4.3 CDU003 - visualizzare Vista generale

L'utente deve poter visualizzare una panoramica del grafo completo. Sarà possibile sfruttare una finestra di visualizzazione la quale conterrà, grazie ad un basso livello di zoom, l'intero contenuto informativo del grafo. Ciò faciliterà l'osservazione nel caso in cui gli elementi costituenti il grafo siano in numero tale da non permetterne, allo stesso tempo, una visualizzazione di alto livello e una visualizzazione di dettaglio.

Una finestra di selezione potrà essere posizionata e modificata nelle sue dimensioni allo scopo di visualizzare gli elementi che essa contiene nella vista principale di cui al caso d'uso CDU004.

Tabella 5 - CDU003: visualizzare Vista generale

Caso d'uso: CDU003	visualizzare Vista generale	Data: 17/04/2012
		Versione: 1.0
Requisito funzionale	REF004	
Descrizione	L'utente di business deve poter visualizzare una panoramica del grafo completo	
Priorità	Normale	
Attori	Utenti di business e utenti tecnici	
Precondizioni	Selezione dei contesti di analisi (CDU002)	
Flusso principale		
1 - L'utente di business visualizza una panoramica del grafo		
2 - L'applicazione mette a disposizione dell'utente la finestra di selezione		
3 - L'utente interagisce con la finestra di selezione		

1.3.4.4 CDU004 - visualizzare grafo

L'utente deve poter navigare il grafo in forma grafica. Per mezzo della finestra di selezione facente parte della Vista principale di cui al caso d'uso CDU003, un'ulteriore vista permetterà la visualizzazione del grafo o di un suo sottoinsieme secondo il grado di dettaglio specificato dall'utente. Sarà inoltre possibile:

- Stampare il contenuto della finestra in un numero di pagine personalizzabile (stampa formato poster);
- Applicare diversi tipi di layout.

Ogni elemento del grafo, di cui ai casi d'uso CDU007 e CDU008, deve essere selezionabile e riposizionabile attraverso un'operazione di drag&drop per permettere all'utente una personalizzazione della vista.

Tabella 6 - CDU004: visualizzare Grafo

Caso d'uso: CDU004	visualizzare Grafo	Data: 17/04/2012
		Versione: 1.0
Requisito funzionale	REF003, REF005, REF006, REF007, REF008, REF019	
Descrizione	L'utente di business deve poter navigare il grafo in forma grafica.	
Priorità	Normale	
Attori	Utenti di business e utenti tecnici	
Precondizioni	Selezione dei contesti di analisi (CDU002)	
Flusso principale		
1 - L'utente visualizza il grafo		
2 - L'utente seleziona gli elementi del grafo e l'applicazione ne visualizza le informazioni aggiuntive associate		
3 - L'utente posiziona gli elementi del grafo		
3 - L'utente applica i layout		
4 - L'utente stampa il contenuto della vista		
5 - L'utente applica lo zoom		

1.3.4.5 CDU005 - ricercare Attributi Dimensionali

L'utente deve poter effettuare ricerche per nome di Attributi Dimensionali. L'applicazione deve reperire tutti gli Attributi Dimensionali il cui nome contiene la sequenza di caratteri specificata dall'utente (tale stringa può essere vuota).

Ogni Attributo Dimensionale del risultato della ricerca deve essere selezionabile in modo da permettere il caricamento delle informazioni associate (Attributi Descrittivi, Misure con relative Funzioni di Aggregazione, Attributi dimensionali Successori, Attributi dimensionali Predecessori e Query) di cui ai casi d'uso, rispettivamente, CDU009, CDU011, CDU012, CDU013 e CDU014.

La selezione dell'Attributo Dimensionale comporta la sua localizzazione ed evidenziazione all'interno della Vista generale di cui ai casi d'uso CDU003 e CDU004.

Tabella 7 - CDU005: ricercare Attributi Dimensionali

Caso d'uso: CDU005	ricercare Attributi Dimensionali	Data: 17/04/2012
		Versione: 1.0
Requisito funzionale	REF013, REF014, REF015	
Descrizione	L'utente di business deve poter effettuare ricerche per nome di Attributi Dimensionali	
Priorità	Normale	
Attori	Utenti di business e utenti tecnici	
Precondizioni	Selezione dei contesti di analisi (CDU002)	
Flusso principale		
1 - L'utente inserisce la stringa di caratteri		
2 - L'applicazione restituisce la lista degli Attributi Dimensionali trovati		
3 - L'utente seleziona un Attributo Dimensionale		

1.3.4.6 CDU006 - ricercare Misure

L'utente deve poter effettuare ricerche per nome di Misure. L'applicazione deve reperire tutte le Misure il cui nome contiene la sequenza di caratteri specificata dall'utente (tale stringa può essere vuota).

Ogni Misura del risultato della ricerca deve essere selezionabile in modo da permettere il caricamento delle informazioni associate (Attributi Dimensionali con relative Funzioni di Aggregazione, Attributi dimensionali Successori e Lista delle Misure del Gruppo di Misure id cui fa parte) di cui ai casi d'uso, rispettivamente, CDU010, CDU012 e CDU015.

La selezione di una Misura comporta la localizzazione ed evidenziazione del Gruppo di Misure cui appartiene all'interno della Vista generale di cui al caso d'uso CDU003.

Tabella 8 - CDU006: ricercare Misure

Caso d'uso: CDU006	ricercare Misure	Data: 17/04/2012
		Versione: 1.0
Requisito funzionale	REF013, REF016	
Descrizione	L'utente di business deve poter effettuare ricerche per nome di Misure	
Priorità	Normale	
Attori	Utenti di business e utenti tecnici	
Precondizioni	Selezione dei contesti di analisi (CDU002)	
Flusso principale		
1 - L'utente inserisce la stringa di caratteri		
2 - L'applicazione restituisce la lista delle Misure trovate		
3 - L'utente seleziona una Misura		

1.3.4.7 CDU007 - visualizzare Attributi Dimensionali

Ogni Attributo Dimensionale del grafo deve essere selezionabile in modo da permettere il caricamento delle informazioni associate (Attributi Descrittivi, Misure con relative Funzioni di Aggregazione, Attributi dimensionali Successori, Attributi dimensionali Predecessori e Query) di cui ai casi d'uso, rispettivamente, CDU009, CDU011, CDU012, CDU013 e CDU014.

Tabella 9 - CDU007: visualizzare Attributi Dimensionali

Caso d'uso: CDU007	visualizzare Attributi Dimensionali	Data: 17/04/2012
		Versione: 1.0
Requisito funzionale	REF008	
Descrizione	L'utente di business deve poter selezionare dalla vista di dettaglio gli Attributi Dimensionali	
Priorità	Normale	
Attori	Utenti di business e utenti tecnici	
Precondizioni	Selezione dei contesti di analisi (CDU002)	
Flusso principale		
1 - L'utente visualizza gli Attributi Dimensionali		
2 - L'utente posiziona gli Attributi Dimensionali		
3 - L'utente seleziona un Attributo Dimensionale		

1.3.4.8 CDU008 - visualizzare Gruppi di Misure

Ogni Gruppo di Misure del grafo deve essere selezionabile in modo da permettere il caricamento delle informazioni associate (Attributi Dimensionali con relative Funzioni di Aggregazione, Attributi dimensionali Successori e Lista delle Misure) di cui ai casi d'uso, rispettivamente, CDU010 e CDU012 CDU015.

Tabella 10 - CDU008: visualizzare Gruppi di Misure

Caso d'uso: CDU008	visualizzare Gruppi di Misure	Data: 17/04/2012
		Versione: 1.0
Requisito funzionale	REF008	
Descrizione	L'utente di business deve poter selezionare dalla vista di dettaglio i Gruppi di Misure	
Priorità	Normale	
Attori	Utenti di business e utenti tecnici	
Precondizioni	Selezione dei contesti di analisi (CDU002)	
Flusso principale		
1 - L'utente visualizza i Gruppi di Misure		
2 - L'utente posiziona i Gruppi di Misure		
3 - L'utente seleziona i Gruppi di Misure		

1.3.4.9 CDU009 - visualizzare Attributi Descrittivi

In seguito alla selezione di un Attributo Dimensionale devono venire visualizzati, in forma testuale, nome e descrizione degli Attributi Descrittivi associati.

Tabella 11 - CDU009: visualizzare Attributi Descrittivi

Caso d'uso: CDU009	visualizzare Attributi Descrittivi	Data: 17/04/2012
		Versione: 1.0
Requisito funzionale	REF009, REF014, REF016	
Descrizione	L'utente deve visualizzare la lista degli Attributi Descrittivi di un Attributo Dimensionale o di una Misura	
Priorità	Normale	
Attori	Utenti di business e utenti tecnici	
Precondizioni	Selezione di un Attributo Dimensionale o di una Misura	
Flusso principale		
1 - L'utente visualizza nome e descrizione degli Attributi descrittivi associati all'Attributo Dimensionale o alla Misura selezionata		

1.3.4.10 CDU010 - visualizzare Attributi Dimensionali e Funzioni di Aggregazione

In seguito alla selezione di una Misura devono essere visualizzati, in forma testuale, gli Attributi Dimensionali associati insieme alle relative Funzioni di Aggregazione.

Dovranno essere visualizzati anche gli Attributi Dimensionali non direttamente collegati al Gruppo di Misure di cui la Misura stessa ne fa parte. In altre parole, dovranno comparire tutti gli Attributi Dimensionali raggiungibili dal Gruppo di Misure.

Tabella 12 - CDU010: visualizzare Attributi Dimensionali e Funzioni di Aggregazione

Caso d'uso: CDU010	visualizzare Attributi Dimensionali e Funzioni di Aggregazione	Data: 17/04/2012
		Versione: 1.0
Requisito funzionale	REF016	
Descrizione	L'utente visualizza la lista degli Attributi Dimensionali e delle relative Funzioni di Aggregazione associati alla Misura	
Priorità	Normale	
Attori	Utenti di business e utenti tecnici	
Precondizioni	Selezione di una Misura	
Post Condizioni		
Flusso principale		
1 - L'utente visualizza il nome degli Attributi Dimensionali e delle relative Funzioni di Aggregazione associati alla Misura precedentemente selezionata		

1.3.4.11 CDU011 - visualizzare Misure e Funzioni di Aggregazione

In seguito alla selezione di un Attributo Dimensionale devono essere visualizzate, in forma testuale, le Misure associate insieme alle relative Funzioni di Aggregazione.

Dovranno essere visualizzate anche le Misure non direttamente collegate all'Attributo Dimensionale selezionato. In altre parole, dovranno comparire tutte le Misure di ogni Gruppo di Misure raggiungibili dall'Attributo Dimensionale selezionato.

Tabella 13 - CDU011: visualizzare Misure e Funzioni di Aggregazione

Caso d'uso: CDU011	visualizzare Misure e Funzioni di Aggregazione	Data: 17/04/2012
		Versione: 1.0
Requisito funzionale	REF010, REF014	
Descrizione	L'utente visualizza la lista delle Misure e delle relative Funzioni di Aggregazione associate all'Attributo Dimensionale	
Priorità	Normale	
Attori	Utenti di business e utenti tecnici	
Precondizioni	Selezione di un Attributo Dimensionale	
Flusso principale		
1 - L'utente visualizza il nome delle Misure e delle relative Funzioni di Aggregazione associate all'Attributo Dimensionale precedentemente selezionato		

1.3.4.12 CDU012 - visualizzare Successori

In seguito alla selezione di un Attributo Dimensionale o di una Misura o di un Gruppo di Misure devono essere visualizzati graficamente, secondo il layout di default verticale, gli Attributi Dimensionali Successori.

Sarà inoltre possibile:

- Stampare il contenuto della finestra in un numero di pagine personalizzabile (stampa formato poster);
- Applicare diversi tipi di layout.

Tabella 14 - CDU012: visualizzare Successori

Caso d'uso: CDU012	visualizzare Successori	Data: 17/04/2012
		Versione: 1.0
Requisito funzionale	REF009, REF011, REF014, REF016, REF019	
Descrizione	L'utente visualizza i Successori dell'elemento selezionato	
Priorità	Normale	
Attori	Utenti di business e utenti tecnici	
Precondizioni	Selezione di un Attributo Dimensionale o di una Misura o di un Gruppo di Misure	
Flusso principale		
1 - L'utente visualizza i Successori dell'elemento selezionato		
2 - L'utente posiziona gli elementi del sottografo originato		
3 - L'utente stampa il contenuto della vista		
4 - L'utente applica i layout disponibili		
5 - L'utente applica lo zoom		

1.3.4.13 CDU013 - visualizzare Predecessori

In seguito alla selezione di un Attributo Dimensionale devono essere visualizzati graficamente, secondo il layout di default organico, gli Attributi Dimensionali Successori. Il layout può essere modificato dall'utente.

Sarà inoltre possibile:

- Stampare il contenuto della finestra in un numero di pagine personalizzabile (stampa formato poster);
- Applicare diversi tipi di layout.

Tabella 15 - CDU013: visualizzare Predecessori

Caso d'uso: CDU013	visualizzare Predecessori	Data: 17/04/2012
		Versione: 1.0
Requisito funzionale	REF010, REF014, REF019	
Descrizione	L'utente visualizza i Predecessori dell'elemento selezionato	
Priorità	Normale	
Attori	Utenti di business e utenti tecnici	
Precondizioni	Selezione di un Attributo Dimensionale	
Flusso principale		
1 - L'utente visualizza i Successori dell'elemento selezionato		
2 - L'utente posiziona gli elementi del sottografo originato		
3 - L'utente stampa il contenuto della vista		
4 - L'utente applica i layout disponibili		
5 - L'utente applica lo zoom		

1.3.4.14 CDU014 - visualizzare Query

Ad ogni attributo dimensionale è associata una query personalizzabile nella selezione dei campi che essa dovrà contenere. Dunque, in seguito alla selezione di un Attributo Dimensionale devono essere visualizzati, in forma testuale, i campi della tabella su cui effettuare la query.

L'utente seleziona i campi desiderati e sottopone al sistema l'interrogazione desiderata. Il risultato deve apparire in una nuova finestra, in forma tabellare paginata.

Tabella 16 - CDU014: visualizzare Query

Caso d'uso: CDU0014	visualizzare Query	Data: 17/04/2012
		Versione: 1.0
Requisito funzionale	REF010, REF012, REF015	
Descrizione	L'utente visualizza la query sottoposta al sistema	
Priorità	Normale	
Attori	Utenti di business e utenti tecnici	
Precondizioni	Selezione di un Attributo Dimensionale	
Flusso principale		
1 - L'utente visualizza la lista dei campi della tabella su cui effettuare la query		
2 - L'utente seleziona i campi della tabella su cui effettuare la query		
3 - L'utente sottomette il comando al sistema		
4 - L'applicazione restituisce all'utente il risultato della query		

1.3.4.15 CDU015 - visualizzare Misure

In seguito alla selezione di una Misura deve essere visualizzata, in forma testuale, la lista delle Misure facenti parte del Gruppo di Misure di cui fa parte la Misura stessa.

In seguito alla selezione di un Gruppo di Misure deve essere visualizzata, in forma testuale, la lista delle Misure facenti parte del Gruppo di Misure.

Il risultato è dunque il medesimo sia che venga selezionata una Misura, sia che venga selezionato un Gruppo di Misure.

Tabella 17 - CDU015: visualizzare Misure

Caso d'uso: CDU015	visualizzare Misure	Data: 17/04/2012
		Versione: 1.0
Requisito funzionale	REF011, REF016	
Descrizione	L'utente visualizza la lista delle Misure associate all'elemento selezionato	
Priorità	Normale	
Attori	Utenti di business e utenti tecnici	
Precondizioni	Selezione di una Misura o di un Gruppo di Misure	
Flusso principale		
1 - L'utente visualizza la lista delle Misure associate alla Misura o al Gruppo di Misure selezionati		

1.3.5 Diagramma di sequenza

Si riporta in Figura 27 il diagramma di sequenza della componente applicativa GraphXMLGenerator.

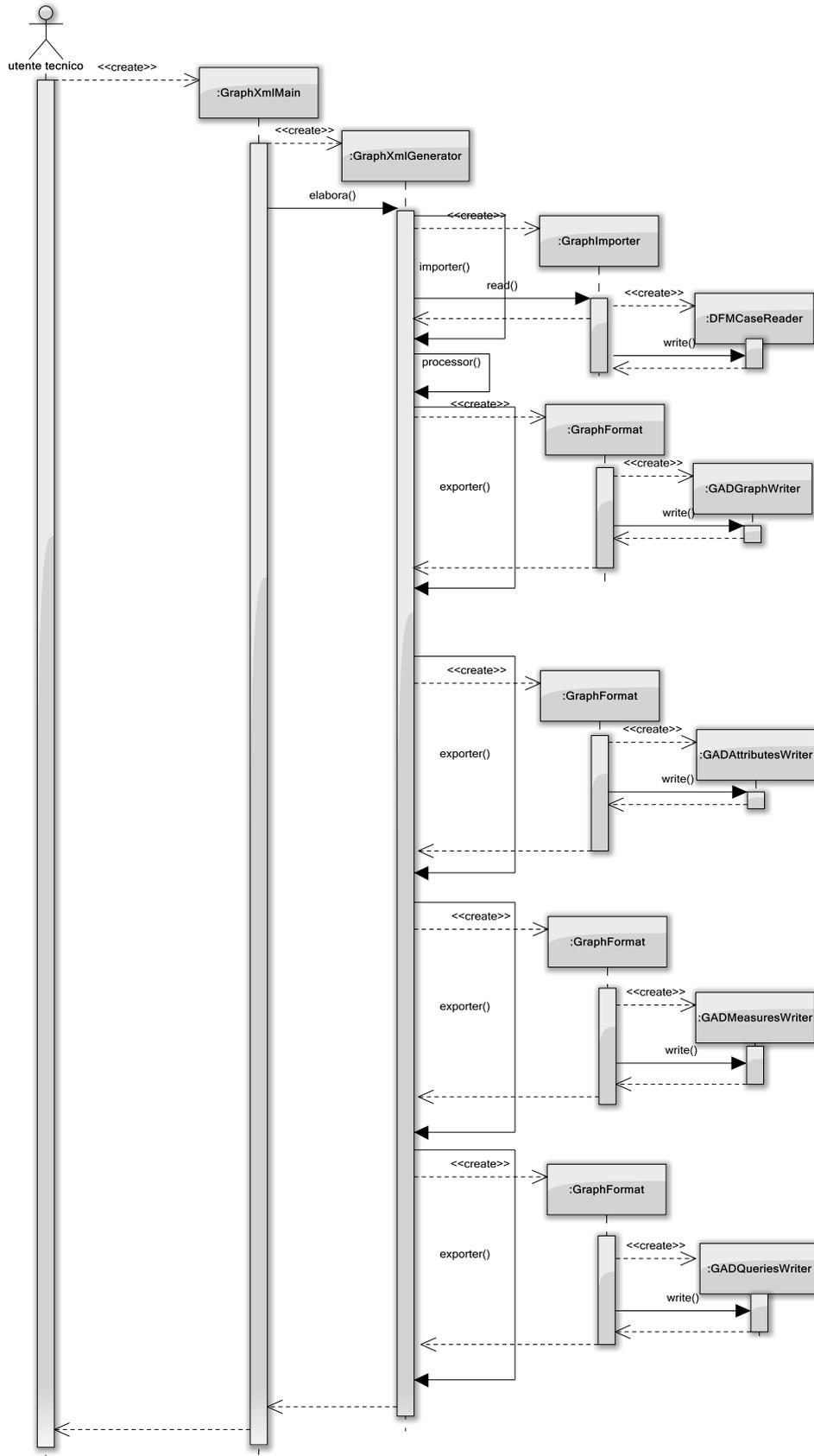


Figura 27

1.3.6 Diagramma delle classi

Si riporta in Figura 28 il diagramma delle classi della componente applicativa GraphXMLGenerator.

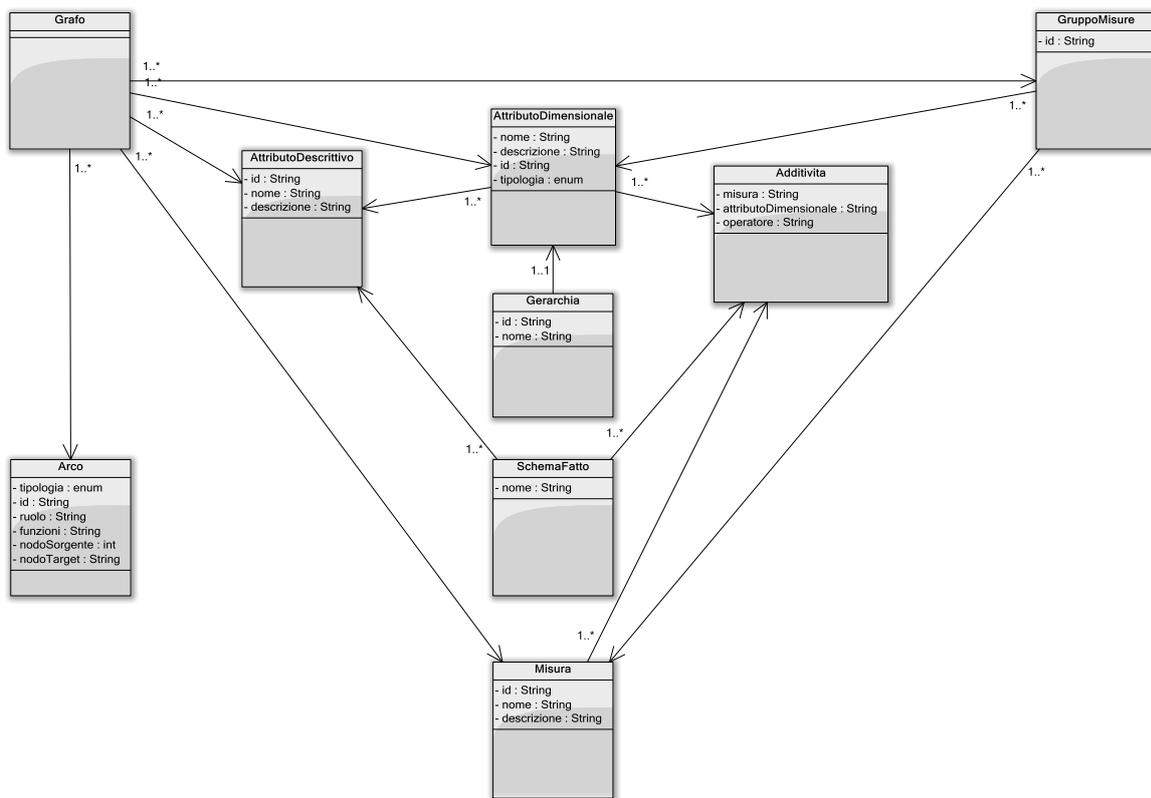


Figura 28

1.3.7 Diagramma di Gantt

In Figura 29 viene riportato il Diagramma di Gantt, il quale descrive con quali tempistiche si sono svolte le attività progettuali.

ID	Task Name	Start	Finish	Duration	mar 2012	apr 2012	mag 2012	giu 2012	lug 2012	ago 2012	set 2012																	
					4/3	11/3	18/3	25/3	1/4	8/4	15/4	22/4	29/4	6/5	13/5	20/5	27/5	3/6	10/6	17/6	24/6	1/7	8/7	15/7	22/7	29/7	5/8	12/8
1	Raccolta dei requisiti	27/02/2012	20/04/2012	8w																								
2	Realizzazione prototipo	27/02/2012	27/04/2012	9w																								
3	Realizzazione software	27/04/2012	07/09/2012	19,2w																								
4	Test software	07/09/2012	21/09/2012	2,2w																								
5	Redazione documentazione	27/04/2012	29/11/2012	31w																								

Figura 29

Raccolta dei requisiti. Definita la necessità di realizzare il software in questione, la prima fase con cui è il progetto è cominciato è stata la raccolta dei requisiti. Il capo progetto Consip periodicamente ha incontrato i rappresentanti della committenza ed insieme hanno definito gli obiettivi che il progetto si proponeva.

Realizzazione prototipo. Parallelamente alla definizione dei requisiti, la realizzazione del prototipo avanzava. Esso ha consentito di illustrare alla committenza le funzionalità principali del sistema con mesi di anticipo rispetto alla data di consegna fissata per il primo ottobre 2012. Coinvolgere attivamente gli utenti ed ottenere la loro indispensabile collaborazione aumenta enormemente la possibilità di successo, definisce con maggiore precisione tempi e costi di realizzazione del progetto, consente di intervenire tempestivamente correggendo errate interpretazioni.

Realizzazione software. Successivamente alla definizione finale dei requisiti è stato possibile ultimare il prototipo, completo di tutte le funzionalità messe a disposizione dal prodotto finale. Tuttavia, problematiche di correttezza dei dati trattati, performance del prodotto e realizzazione della sorgente dati alimentante l'applicazione (il file *dfmcase.dfm* definitivo è tutt'ora in fase di realizzazione), fanno sì che il prototipo necessiti di diverse settimane di sviluppo per la sua evoluzione.

Test software. In tale fase sarà possibile definire i casi di test associati ai casi d'uso del prodotto, quindi valutare le performance dello stesso. Sarà successivamente possibile fare una valutazione dello sforzo, quindi determinare attraverso il calcolo dei function point il costo finale del progetto.

Redazione documentazione. In seguito alla definizione dei requisiti è stato possibile redigere la documentazione ufficiale del progetto, parte della quale è stata riportata nel presente Capitolo. Ulteriore documentazione sarà prodotta nell'ambito del test del prodotto e della redazione del manuale utente finale.

2 Manuale utente

Si riporta in tale paragrafo la redazione del Manuale Utente in riferimento alla parte dell'applicazione implementata dal prototipo. Esso comprende la definizione delle funzionalità cardine del prodotto; saranno dunque riportate le descrizioni delle schermate *Elaborazioni* e *Visualizzazione query*. Le schermate di *Login* e *Contesti* non sono state ad oggi incluse nel prototipo.

2.1 Schermata Elaborazioni

Un'istanza di esempio della schermata *Elaborazioni* è riportata in Figura 30.

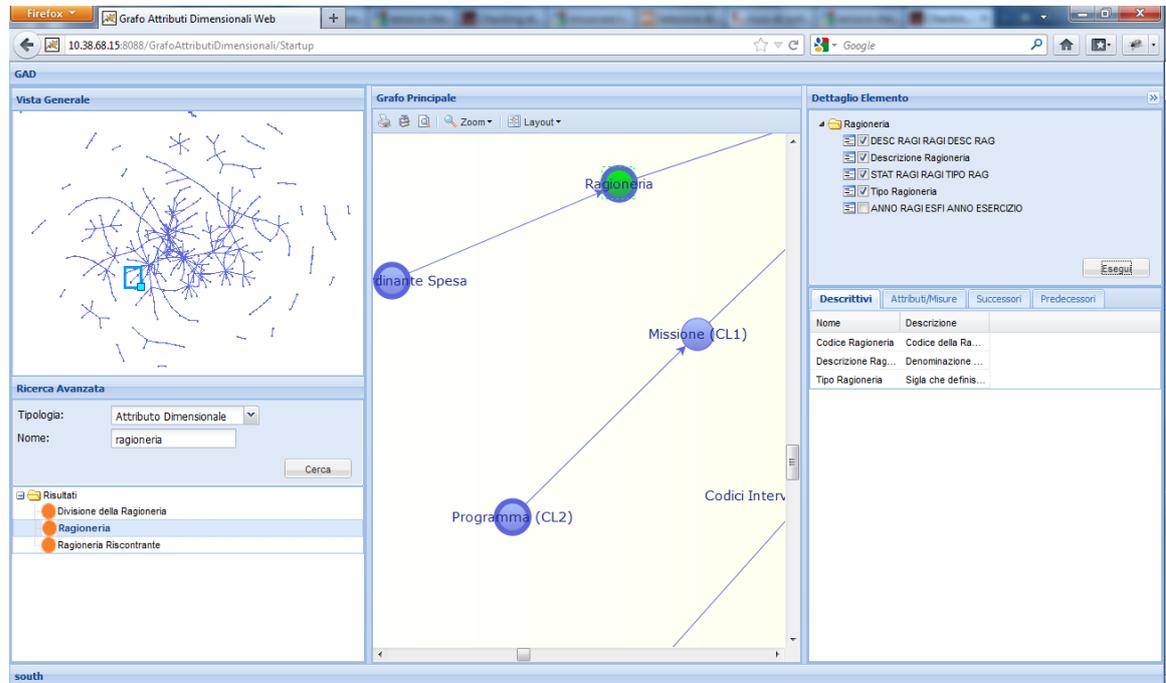


Figura 30

2.1.1 Schermata Vista generale

La funzione principale del pannello "Vista generale" è offrire una panoramica dell'intero contenuto informativo del grafo.

Esso mostra il contenuto del pannello "Grafo principale" con un livello di zoom tale da rendere non necessario lo scroll della finestra. In altre parole, è possibile visualizzare una panoramica del contenuto informativo del grafo tramite un vista globale.

Una finestra di selezione, personalizzabile nelle sue dimensioni, può essere posizionata in modo da contenere una porzione del grafo, la quale verrà automaticamente mostrata nel pannello "Grafo principale".

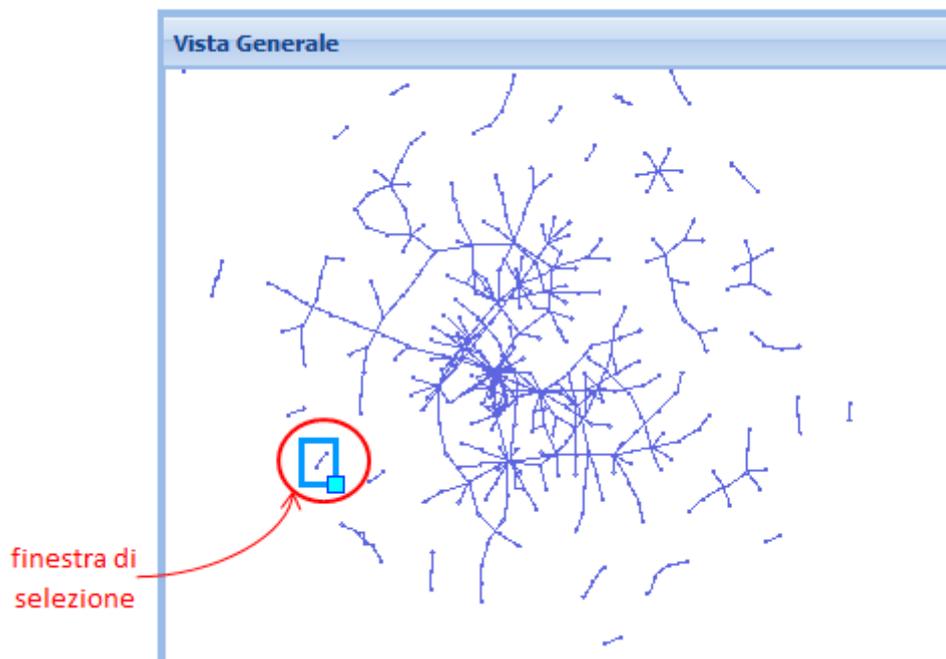


Figura 31

2.1.2 Schermata Ricerca avanzata

Tramite il pannello "Ricerca avanzata" è possibile effettuare una ricerca per nome di Attributi dimensionali e Misure.

La ricerca viene avviata attraverso il click del bottone "Cerca". Verranno reperiti tutti gli Attributi Dimensionali o Misure (la ricerca è esclusiva: nello stesso istante possono essere ricercati Attributi dimensionali oppure Misure, non entrambi, selezionando la voce desiderata dal menù "Tipologia") il cui nome contiene la sequenza di caratteri, specificata dall'utente, contenuta nella casella di testo "Nome" (se quest'ultima non contiene alcun carattere, la ricerca restituirà l'intera lista di Attributi dimensionali o Misure).

Se la ricerca produce una lista non vuota di elementi, sarà possibile visualizzare il risultato tramite il click sul simbolo "+" accanto alla scritta "Risultati". Successivamente, facendo click su un elemento di tale lista, l'applicazione provvederà automaticamente:

- alla localizzazione dell'elemento nel grafo, visualizzandolo nel pannello "Grafo principale";
- alla visualizzazione della query associata all'elemento, nonché al caricamento dei suoi Attributi descrittivi, Attributi Dimensionali Successori, Attributi Dimensionali Predecessori, Attributi Dimensionali (se l'elemento è una Misura) o Misure (se l'elemento è un Attributo dimensionale) associati con relative Funzioni di aggregazione, nel pannello "Dettaglio elemento".

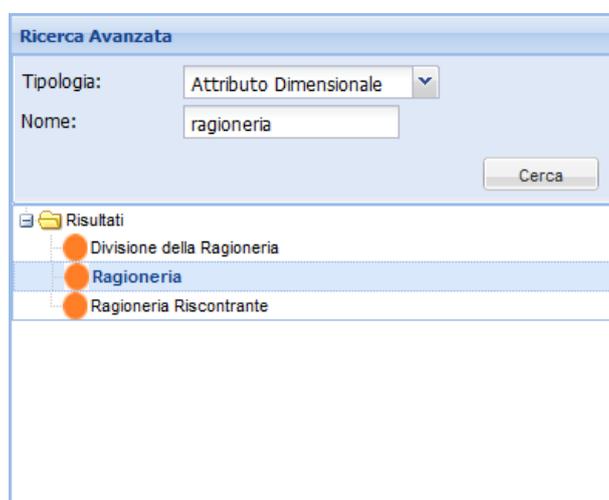


Figura 32

La Figura 32 mostra la ricerca di Attributi dimensionale il cui nome contenga la stringa "ragioneria". La ricerca ha prodotto una lista di tre elementi.

2.1.3 Schermata Grafo principale

Il pannello "Grafo principale" visualizza il contenuto della finestra di selezione del pannello "Vista generale". Utilizzando le barre di scorrimento verticale e orizzontale è possibile visualizzare la restante parte del grafo.

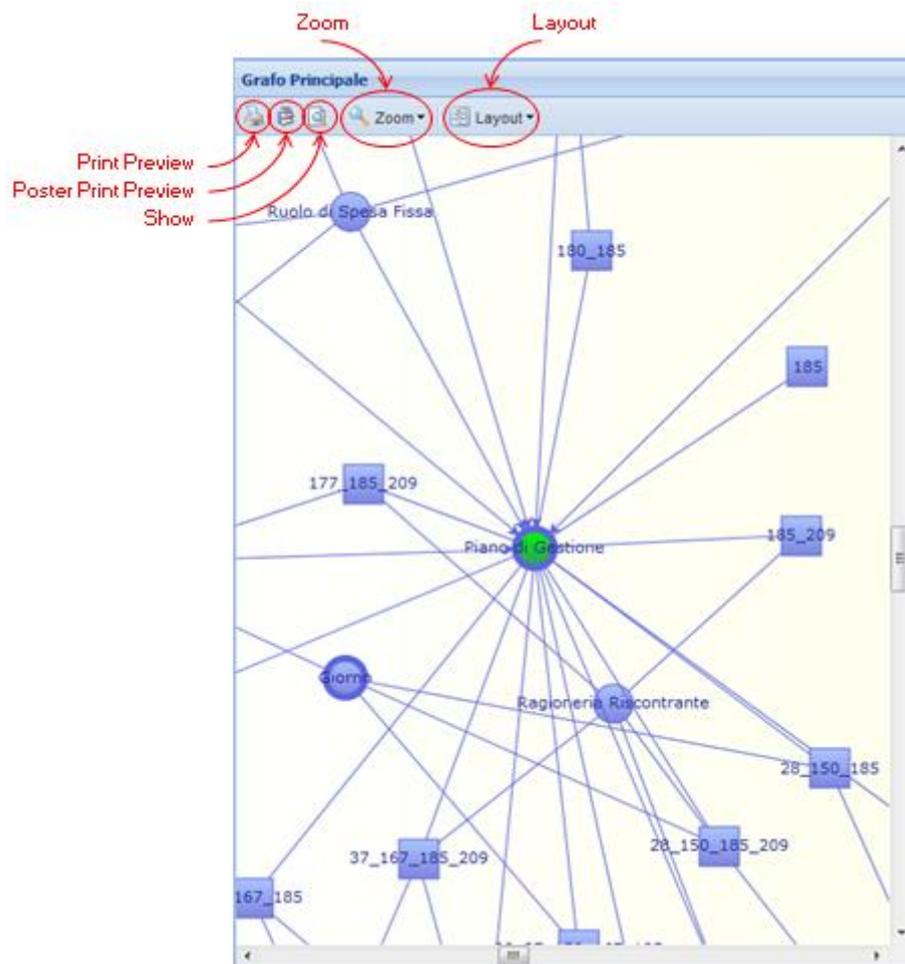


Figura 33

In Figura 33 viene mostrata la barra dei menu la quale mette a disposizione le seguenti funzioni:

- Print Preview: mostra il contenuto di "Grafo principale" in una nuova scheda del browser;
- Poster Print Preview, mostra il contenuto di "Grafo principale" in una nuova scheda mediante un numero di pagine specificate dall'utente;
- Show, mostra il contenuto del pannello Vista generale in una nuova scheda del browser;
- Zoom, applica il livello di zoom desiderato dall'utente;
- Layout, ridispone gli elementi del grafo secondo un'impaginazione differente.

La simbologia utilizzata per le entità presenti nel grafo è indicata in Figura 34.



Figura 34

2.1.4 Schermata Dettaglio elemento

In Figura 35 è riportata un'istanza del pannello "Dettaglio elemento", relativamente alla selezione dell'Attributo Dimensionale "Ragioneria".

Nel riquadro in alto appaiono i campi sui quali può essere effettuata la query associata all'attributo. In tal caso cliccando sul bottone "Esegui" verrà visualizzata la schermata relativa alla query, ristretta ai campi "Descrizione Ragioneria" e "Tipo Ragioneria".

Nel riquadro in basso appaiono gli Attributi Descrittivi associati all'attributo dimensionale.

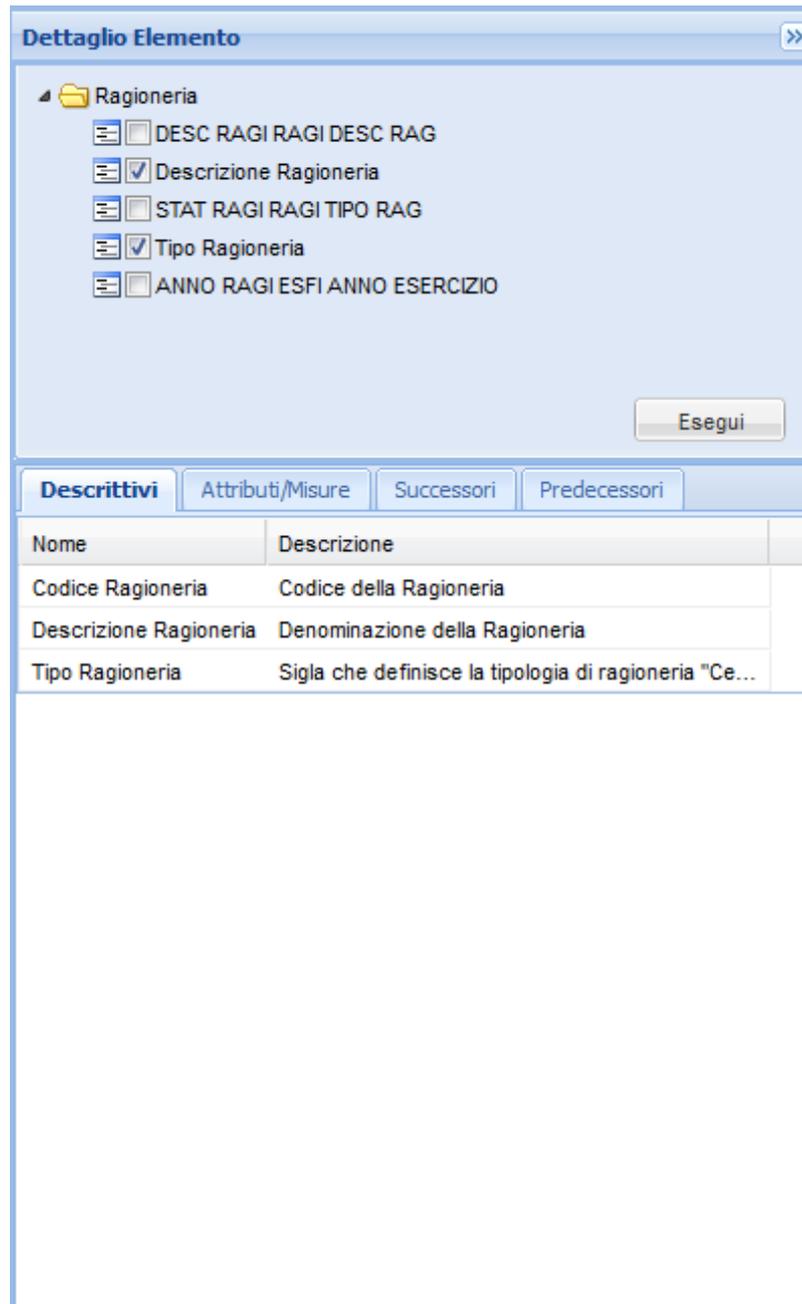


Figura 35

Cliccando sul tab "Attributi/Misure" viene visualizzata la lista delle Misure con le relative funzioni di aggregazione dell'attributo "Piano di Gestione" (vedi Figura 36).

Cliccando sul tab "Successori" viene visualizzato il grafo associato agli Attributi Dimensionali Successori dell'attributo "Piano di Gestione" (vedi Figura 37).

Cliccando sul tab "Predecessori" viene visualizzato il grafo associato agli Attributi Dimensionali Predecessori dell'attributo "Piano di Gestione" (vedi Figura 38).

Descrittivi	Attributi/Misure	Successori	Predecessori
	Variazione in ...	SUM	
	Variazione in ...	SUM	
	Variazione in ...	SUM	
	Variazione in ...	SUM	
	Emesso Lordo ...	SUM	
	Emesso Netto O...	SUM	
	Emesso Note di I...	SUM	
	Pagato Note di...	SUM	
	Emesso Ordini d...	SUM	
	Importo Preleva...	SUM	
	Pagato Ordini ...	SUM	
	Prenotato Ordini...	SUM	
	Residui Accer...	SUM	

Figura 36

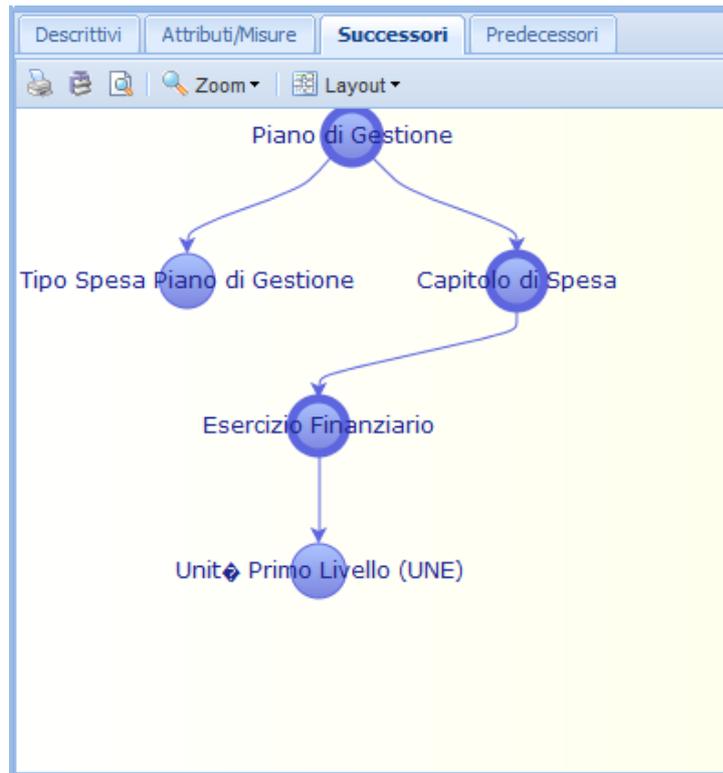


Figura 37

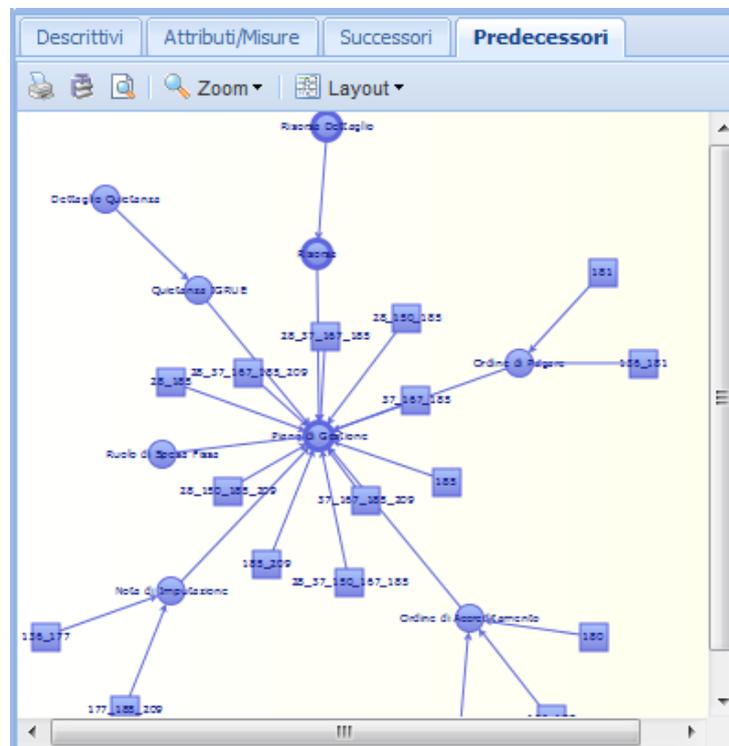


Figura 38

2.2 Schermata Visualizzazione query

In Figura 39 è riportato un esempio di visualizzazione del risultato della query ottenuta dalla selezione dei campi, relativi alle colonne della tabella mostrata, in riferimento all'Attributo Dimensionale "Ragioneria".

Ragioneria

DESC_RAGI_RAGI_DESC_RAG	DESCRIZIONE_RAGIONERIA	STAT_RAGI_RAGI_TIPO_RAG	TIPO_RAGIONERIA
RAGIONERIA CENTRALE PRESSO IL MINISTERO DELL'ECONOMIA E DELLE FINANZE	RAGIONERIA CENTRALE PRESSO IL MINISTERO DELL'ECONOMIA E DELLE FINANZE	RC	RC
RAGIONERIA CENTRALE PRESSO IL MINISTERO DELL'ECONOMIA E DELLE FINANZE	RAGIONERIA CENTRALE PRESSO IL MINISTERO DELL'ECONOMIA E DELLE FINANZE	RC	RC
RAGIONERIA CENTRALE PRESSO IL MINISTERO DELL'ECONOMIA E DELLE FINANZE	RAGIONERIA CENTRALE PRESSO IL MINISTERO DELL'ECONOMIA E DELLE FINANZE	RC	RC
RAGIONERIA CENTRALE PRESSO IL MINISTERO DELLE INFRASTRUTTURE	RAGIONERIA CENTRALE PRESSO IL MINISTERO DELLE INFRASTRUTTURE	RC	RC
RAGIONERIA CENTRALE PRESSO IL MINISTERO DELLE INFRASTRUTTURE	RAGIONERIA CENTRALE PRESSO IL MINISTERO DELLE INFRASTRUTTURE	RC	RC
RAGIONERIA CENTRALE PRESSO IL MINISTERO DEL LAVORO E DELLA PREVIDENZA SOCIALE	RAGIONERIA CENTRALE PRESSO IL MINISTERO DEL LAVORO E DELLA PREVIDENZA SOCIALE	RC	RC
RAGIONERIA CENTRALE PRESSO IL MINISTERO DEL LAVORO E DELLA PREVIDENZA SOCIALE	RAGIONERIA CENTRALE PRESSO IL MINISTERO DEL LAVORO E DELLA PREVIDENZA SOCIALE	RC	RC
RAGIONERIA CENTRALE PRESSO IL MINISTERO DEL LAVORO E DELLA PREVIDENZA SOCIALE	RAGIONERIA CENTRALE PRESSO IL MINISTERO DEL LAVORO E DELLA PREVIDENZA SOCIALE	RC	RC
RAGIONERIA CENTRALE PRESSO IL MINISTERO DELL'UNIVERSITA' E DELLA RICERCA SCIENTIFICA E TECNOLOGICA	RAGIONERIA CENTRALE PRESSO IL MINISTERO DELL'UNIVERSITA' E DELLA RICERCA SCIENTIFICA E TECNOLOGICA	RC	RC
RAGIONERIA CENTRALE PRESSO IL MINISTERO DELL'UNIVERSITA' E DELLA RICERCA SCIENTIFICA E TECNOLOGICA	RAGIONERIA CENTRALE PRESSO IL MINISTERO DELL'UNIVERSITA' E DELLA RICERCA SCIENTIFICA E TECNOLOGICA	RC	RC

16,872 records estratti, in visualizzazione 1 di 10.

[Prima](#)
[Precedente](#)
[1](#)
[2](#)
[3](#)
[4](#)
[5](#)
[6](#)
[7](#)
[8](#)
[Successiva](#)
[Ultima](#)

Figura 39

3 Release future

È auspicabile la realizzazione di nuove release del prodotto per l'implementazione delle due macrofunzionalità esposte nei sottoparagrafi successivi.

3.1 Modifica della struttura del DW RGS

Gli utenti tecnici hanno intuito la possibilità di poter utilizzare il sistema per la modifica della struttura del DW RGS. Si pensa dunque che sarà possibile realizzare una nuova funzionalità tramite la quale l'utente tecnico, in base alle analisi da lui effettuate sfruttando la visione integrata e semplificata del grafo, potrà realizzare nuovi fatti e dimensioni. Questo semplificherà notevolmente il processo operativo attuale, diminuendo tempi, quindi costi, e diminuendo drasticamente la possibilità del verificarsi di errori.

3.2 Query ontologiche

Gli utenti di business potranno effettuare query ontologiche.

Tipicamente, l'oggetto dell'integrazione di differenti sistemi informativi riguarda la cooperazione di dati eterogenei contenuti in diverse basi di dati distribuite sia all'interno che all'esterno dell'organizzazione. Tali informazioni vengono integrate e presentate agli utilizzatori del sistema per mezzo della definizione di uno schema dati globale, vedi Figura 40.

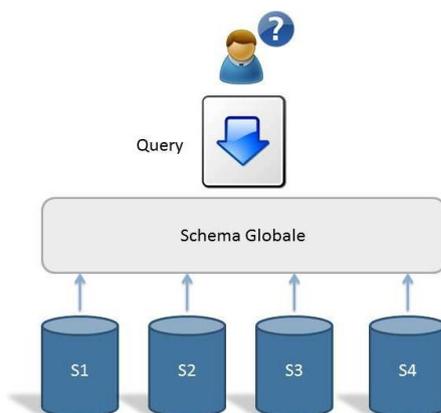


Figura 40

Il punto debole della soluzione classica sopra descritta deriva dalla mancata concettualizzazione della visione integrata dell'informazione in quanto lo schema globale (sia esso virtuale o materializzato) è espresso in termini di modello logico dei dati (tipicamente secondo il formalismo relazionale). La potenza espressiva di tali tipologie di modelli consente un livello di astrazione non sufficiente a descrivere un dominio complesso avendo a disposizione costrutti fortemente influenzati da problematiche di implementazione e di performance.

Ulteriore aspetto penalizzante deriva dal fatto che le regole caratterizzanti il dominio di interesse (regole di business e regole di integrazione) vengono colte ed implementate esclusivamente attraverso la realizzazione di opportune procedure software e mai espresse in maniera univoca e formale. Questo introduce notevoli problematiche di semantica del dato ponendo importanti barriere nell'accesso all'informazione in maniera univoca e coerente.

Al fine di mitigare le problematiche sopra espresse la comunità scientifica internazionale ha recentemente proposto l'utilizzo delle ontologie per l'accesso alle basi dati, vedi Figura 41. Tale approccio va sotto il nome di OBDA (Ontology-Based Data Access) e si poggia sull'idea fondamentale di esprimere lo schema globale attraverso una ontologia. La visione integrata che il sistema fornisce ai suoi utilizzatori non è più espressa da semplici strutture dati ma viene formalizzata attraverso una descrizione semantica dei concetti e delle relazioni proprie del dominio di interesse. L'accesso ai dati viene garantito da opportuni mapping che mascherano completamente la struttura fisica dei sistemi sorgente. Come ulteriore elemento caratterizzante ogni query espressa innesca un meccanismo di "ragionamento" sull'ontologia producendo risultati ottimali anche in presenza di dati incompleti o di bassa qualità.

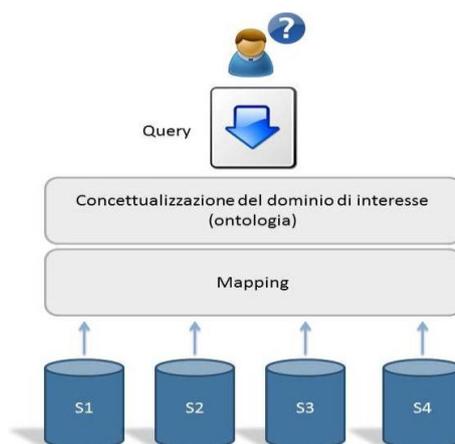


Figura 41

Per meglio comprendere quanto finora espresso e soprattutto quanto si esprimerà nel seguito si ritiene utile introdurre un esempio che mostri la differenza di una interrogazione espressa su una ontologia dall'analogha query espressa attraverso l'utilizzo diretto delle basi dati.

La risposta ad una query posta su un classico sistema di basi dati viene elaborata secondo l'assunzione di "mondo chiuso" (CWA, Closed World Assumption) che afferma che tutto quello che non è espresso è necessariamente falso, ovvero, nelle interrogazioni, non vi è indeterminazione: si considera vero quello che è presente sulla base di dati e falso quello che non c'è. Come immediato corollario, in un sistema di basi dati le informazioni soddisfano pienamente lo schema che li contiene; la struttura dello schema e le sue regole non sono utili né utilizzate per determinare la risposta ad una interrogazione.

In maniera totalmente opposta un'ontologia elabora il risultato di una interrogazione secondo l'assunzione di "mondo aperto" (OWA, Open World Assumption) che afferma che tutto ciò che è contenuto nell'ontologia è vero, mentre tutto ciò che non vi è contenuto non è né vero né falso, ma semplicemente non si conosce. Questo implica che il sistema assume i dati come potenzialmente inconsistenti applicando regole di inferenza per derivare assunzioni esprimibili a partire da quelle contenute nell'ontologia.

Si supponga di operare nel dominio espresso in Figura 42 che asserisce che: ogni persona ha un padre che a sua volta è una persona.

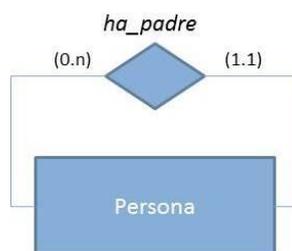


Figura 42

Si supponga di avere a disposizione una base dati che contiene i seguenti valori:

Persona: {*Fabrizio, Antonio, David*}

ha_padre: { (*Fabrizio, Antonio*), (*Antonio, David*) }

ovvero *Fabrizio, Antonio* e *David* sono persone, il padre di *Fabrizio* è *Antonio* ed il padre di *Antonio* è *David*.

Si eseguano le seguenti query sui due sistemi (per semplicità espresse in linguaggio naturale):

Query 1: Quali sono tutte le persone che hanno un padre?

Query 2: Quali sono le persone che hanno un padre e un nonno?

Database

Query 1: *Fabrizio, Antonio*

Query 2: *Fabrizio*

Ontologia

Query 1: *Fabrizio, Antonio, David*

Query 2: *Fabrizio, Antonio, David*

Come si vede le risposte ottenibili dal sistema di base dati, per l'assunzione di CWA, sono esclusivamente quelle contenute nel database stesso. Dal momento che non è esplicitamente contenuta l'informazione di quale sia il padre di *David*, si assume che *David* non abbia padre e di conseguenza *Antonio* non abbia un nonno. Completamente diverso è l'approccio OWA ontologico che pur partendo dai stessi dati incompleti, riesce a derivare informazioni aggiuntive in virtù della possibilità di inferire nuova informazione. Nello specifico, dal momento che il modello asserisce che tutte le persone hanno un padre, automaticamente tutte le persone

presenti nella base dati hanno un padre (query q1). Allo stesso modo dal momento che tutte le persone hanno un padre e che un padre è una persona, si inferisce che ogni padre ha un padre ovvero che tutte le persone hanno un nonno.

L'idea specifica, relativamente all'applicazione GAD, è permettere all'utente di business di formulare query ontologiche attraverso l'interfaccia Web realizzata. L'utente, interagendo con il grafo, il quale rappresenta il modello dei dati globale su esposto, confeziona una query che il sistema traduce in una interrogazione sull'ontologia mediante il linguaggio SPARQL. L'architettura OBDA, mediante la definizione di mapping sui sistemi sorgente, estrae i dati di interesse, li arricchisce con le regole di inferenza proprie del dominio e li restituisce all'utente.

Conclusioni

La Specifica Funzionale riportata nel Capitolo 3 è stata approvata dalla committenza in data 30 maggio. Ciò conferma che i requisiti raccolti sono definitivi e stabili. Le tempistiche del progetto sono al momento rispettate. Appare dunque lecito pensare che il progetto non subirà rallentamenti, dunque la consegna dovrebbe avvenire, come stabilito, il 1° ottobre 2012.

GAD è solo uno degli esempi che si potrebbero fare per motivare l'impegno da parte della Pubblica Amministrazione ad un'informatizzazione sempre più spinta, al fine di supportare al meglio sia i processi operativi che quelli decisionali.

L'infrastruttura tecnica del Sistema Informativo della RGS è all'avanguardia. Tuttavia, gli elementi necessari a condurre un'azienda, sia essa pubblica o privata, verso il successo, non sono solo tecnologici. Le parole chiave devono essere miglioramento continuo delle performance dei processi, profusa cultura aziendale a tutti i livelli, continuo riallineamento di Organizzazione, Sistema Informativo ed ICT, fare della strategia un processo continuo.

Bibliografia

L. Astegiano, S. Coletta, S. Ferrari, F. Frezza, S. Lomuscio, L. Perugini, A. Stati (2006). *Il Data Warehouse ed i Sistemi Informativi Direzionali della Ragioneria Generale dello Stato*. Quaderni Consip

M. Golfarelli, S. Rizzi (2006). *Data Warehouse – teoria e pratica della progettazione*. McGraw - Hill

Indice delle tabelle

Tabella 1 - operatori di aggregazione	79
Tabella 2 - definizioni ed acronimi	89
Tabella 3 - CDU001: verificare identità	100
Tabella 4 - CDU002: scegliere contesto	101
Tabella 5 - CDU003: visualizzare Vista generale	102
Tabella 6 - CDU004: visualizzare Grafo	103
Tabella 7 - CDU005: ricercare Attributi Dimensionali.....	104
Tabella 8 - CDU006: ricercare Misure	105
Tabella 9 - CDU007: visualizzare Attributi Dimensionali.....	106
Tabella 10 - CDU008: visualizzare Gruppi di Misure.....	107
Tabella 11 - CDU009: visualizzare Attributi Descrittivi	108
Tabella 12 - CDU010: visualizzare Attributi Dimensionali e Funzioni di Aggregazione	109
Tabella 13 - CDU011: visualizzare Misure e Funzioni di Aggregazione.....	110
Tabella 14 - CDU012: visualizzare Successori	111
Tabella 15 - CDU013: visualizzare Predecessori	112
Tabella 16 - CDU014: visualizzare Query	113
Tabella 17 - CDU015: visualizzare Misure.....	114