



UNIVERSITÀ DI PISA

**DIPARTIMENTO DI INGEGNERIA DELL'ENERGIA DEI SISTEMI
DEL TERRITORIO E DELLE COSTRUZIONI**

**RELAZIONE PER IL CONSEGUIMENTO DELLA
LAUREA MAGISTRALE IN INGEGNERIA GESTIONALE**

**Realizzazione e governance
di un moderno sistema di analytics per il controllo di
gestione in ambito banking**

SINTESI

RELATORI

Prof. Riccardo Dulmin
*Dipartimento di Ingegneria dell'Energia,
dei Sistemi, del Territorio
e delle Costruzioni*

Ing. Daniele Barone
Deloitte Consulting S.r.l S.B.

IL CANDIDATO

Simona Conversano
s.conversano@studenti.unipi.it

Sessione di Laurea Magistrale del 23/11/2022

Realizzazione e governance di un moderno sistema di analytics per il controllo di gestione in ambito banking

Simona Conversano

Sommario

Questo lavoro di tesi è stato sviluppato a partire dal percorso di tirocinio da me svolto presso Deloitte Consulting S.r.l S.B., azienda leader nella fornitura di servizi di consulenza di ampio genere. Il contenuto di questa tesi verte su un obiettivo cardine ossia sull'implementazione di attività di Data Extraction, Data Cleaning, Data Preparation e Data Visualization per un Gruppo Bancario. Il mio contributo al progetto si può sintetizzare in alcune attività principali come: a) mantenimento di una comunicazione attiva e continua con il cliente grazie alla predisposizione di prefissati momenti di incontro volti alla definizione dei requisiti e alla condivisione dello stato di avanzamento dei lavori; b) trascrizione formale dei requisiti e mantenimento di una documentazione funzionale e tecnica costantemente aggiornata; c) utilizzo del linguaggio PL/SQL per la manipolazione dei dati e interrogazione di questi ultimi; d) creazione di un modello dati Tabular e di un sistema di Reportistica associato; e) supporto operativo al cliente nella fase di UAT (User Acceptance Testing); f) definizione delle attività preparatorie per il rilascio della soluzione; g) progettazione di possibili sviluppi futuri.

Abstract

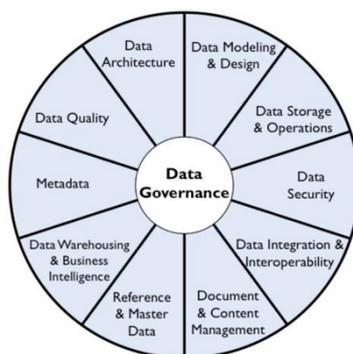
This dissertation is developed by the internship I carried out at Deloitte Consulting S.r.l S.B., leader company in the provision of wide-ranging consulting services. The key objective of the thesis is the implementation of Data Extraction, Data Cleaning, Data Preparation and Data Visualization activities for a Banking Group. My contribution to the project can be summarized in some key activities: a) maintenance of active and continuous communication with the customer thanks to the preparation of predetermined meetings aimed at defining requirements and sharing the progress of the work; b) formal transcription of requirements and maintenance of a functional and technical documentation that is constantly updated; c) use of the PL / SQL language for the manipulation of data; d) creation of a Tabular data model and an associated reporting system; e) operational support to the customer in the UAT (User Acceptance Testing) phase; f) definition of preparatory activities for release and go-live; g) planning of possible future developments.

1 La necessità di gestire i dati nel Contesto Bancario

La trasformazione digitale ha interessato ampiamente il settore Financial Services, comprendente banche e assicurazioni. I grandi istituti bancari, infatti, rappresentano ad oggi gli enti maggiormente interessati al tema del Data Management, in quanto assistono, ogni giorno, al proliferare di milioni di dati provenienti da diverse fonti. Le banche riscontrano, quindi, la forte necessità di governare questa mole di informazioni al fine di soddisfare uno tra gli obiettivi di business fondamentali: ottenere una base dati affidabile, che possa supportare l'adozione di precise soluzioni strategiche e competitive in tempo reale. Nasce quindi l'esigenza di gestire i propri dati al fine di generare conoscenza. Attraverso quest'ultima, è possibile migliorare il processo di decision making, uno tra i vantaggi più ingenti derivanti dall'utilizzo dei dati. Il processo decisionale deve, infatti, essere ottimizzato attraverso il governo dei dati e la trasformazione di questi ultimi in informazioni attendibili e di valore.

2 Fase 0: Attività Preparatorie

L'obiettivo fondamentale del lavoro di tesi è stato l'implementazione operativa dei requisiti di un progetto di Data Management. Per questo motivo, è stato effettuato un approfondimento iniziale del tema. DAMA international, organizzazione leader nello sviluppo di temi di Data Management, lo definisce come «lo sviluppo, l'esecuzione e la supervisione di piani, politiche, programmi e procedure che garantiscono, controllano, proteggono e migliorano il valore dei data asset e delle informazioni per tutta la durata del loro ciclo di vita».



Copyright© 2017 DAMA International

Figura 1. DAMA Wheel

Nel framework esposto nel DAMA-DMBOK, chiamato DAMA *Wheel* (Figura 1), vengono definite le *Knowledge Area* del Data Management, ossia le *best practice* per la completa gestione dei dati.

L'approccio di ogni organizzazione al Framework proposto da DAMA international segue percorsi differenti; ogni azienda definisce, infatti, delle priorità di azione e, rispetto ad esse, focalizza la propria attenzione. A tal proposito, il progetto in esame verte principalmente su tematiche di Data Modeling; l'obiettivo di questo lavoro di tesi è infatti la predisposizione di un modello dati che contenga informazioni che siano contemporaneamente accurate, complete e consistenti. Tale processo si rende necessario soprattutto a causa dell'enorme mole di dati depositati nel database del Gruppo Bancario, che richiedono di essere sistematicamente gestiti. Lo scopo finale è rendere possibile una visualizzazione efficace della natura dei dati, in modo tale da fornire una panoramica globale del contesto al quale si riferiscono ed eventuali direzioni di miglioramento. La reportistica si propone quindi come mezzo con cui adottare decisioni supportate dall'informazione ed individuare eventuali trend o anomalie.

Le principali tecnologie scelte a supporto del progetto in esame sono:

- **Toad for Oracle** (Data Extraction)

Il tool viene utilizzato per la prima fase di Data Extraction. L'estrazione dei dati avviene, infatti, dal database del Gruppo Bancario, al fine di importare, sul successivo modello tabular, esclusivamente il perimetro di dati interessato.

- **Microsoft Visual Studio** (Data Modeling)

Visual Studio rappresenta un ambiente di sviluppo che, nel caso in esame, viene utilizzato per la manipolazione e modellazione dei dati; il dataset, infatti, viene sottoposto ad opportune attività di pulizia e ad un allargamento volto alla creazione di nuovi campi funzionali a supporto dell'operatività del cliente.

- **Power Bi** (Data Visualization)

L'obiettivo del progetto è rendere possibile una visualizzazione dei dati attraverso apposite Dashboard, dopo aver reso tali dati affidabili e di qualità, tramite i processi precedentemente citati. Il fine ultimo risulta, quindi, fornire una chiara rappresentazione del contesto al quale i dati si riferiscono. Tale attività viene effettuata su Power Bi che, con le sue molteplici funzionalità, permette un'ampia esplorazione dei dati al fine di proporre un'analisi che possa servire come base per l'ottimizzazione dei processi aziendali, per la definizione di nuove scelte

strategiche e per la conoscenza globale dei propri dati.

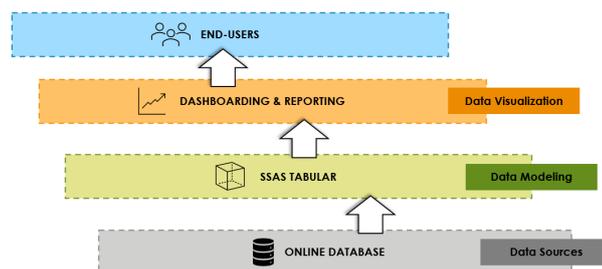


Figura 2. Architettura di Progetto

Per progetti della tipologia in esame, si propone uno standard architettonico come quello riportato in Figura 2, che espone globalmente tutti i layer interessati. Dopo aver effettuato la connessione ad un'apposita base dati, viene creato il modello SSAS Tabular dove viene effettuata la connessione al perimetro dati interessato e la successiva modellazione completa dei dati. Il penultimo layer, collegato al modello tabulare creato, viene utilizzato esclusivamente per scopi di visualizzazione con l'obiettivo di predisporre un'efficace interfaccia utente in accordo con i requisiti concordati.

3 Fase 1: Attività Applicative

Dopo aver ben compreso il contesto di riferimento, hanno avuto inizio le attività operative. Il progetto è stato affrontato tramite una metodologia comune ai servizi erogati da Deloitte e che rappresenta ad oggi un framework cardine per i progetti futuri ed in essere.

3.1 Analisi della situazione AS-IS e comprensione dello scenario di riferimento

Per comprendere al meglio la situazione AS IS, è stato necessario uno scambio informativo con la società cliente, al fine di comprendere principalmente i bisogni di progetto. Il Gruppo Bancario ha messo a disposizione dei consulenti Deloitte la connessione da remoto ai server del sistema bancario, fornendo una piena accessibilità ai database e ai programmi da utilizzare. L'esigenza di progetto riscontrata è stata correlata alla necessità di effettuare puntuali attività di gestione dei dati per una duplice finalità.

Il primo obiettivo risulta quello di semplificare le attività del cliente, in modo tale che possa rapidamente valutare, tramite un adeguato sistema di reportistica, le criticità dei dati finanziari, senza correre il rischio di intraprendere decisioni errate. Il sistema di reportistica

appare, quindi, come uno strumento efficace che presenta una situazione globale tramite grafici e tabelle, nonché filtri e misure economiche.

Il secondo obiettivo è legato alle operazioni tecniche che precedono la generazione della reportistica e che sono legate alle attività di estrazione del dato e pulizia di esso.

3.2 Pianificazione Roadmap di Progetto

Ogni progetto necessita della strutturazione di una roadmap operativa, redatta al fine di consolidare scadenze concordate e di generare un'organizzazione del lavoro più strutturata. Pertanto, il piano viene condiviso nella prima riunione di progetto che coinvolge le principali parti interessate. Si sviluppa quindi un vero e proprio masterplan che mostra il cronoprogramma sintetizzato per le quattro milestone cruciali: raccolta dei requisiti, analisi funzionale, sviluppo e, infine, fase di test utente e di go live.

La prima fase è rappresentata dalla raccolta dei requisiti, intesi come tutte le specifiche richieste dal cliente che devono essere operativamente implementate. Questa fase si concretizza in un continuo scambio informativo con il cliente mediante incontri volti alla formalizzazione dei requisiti in esame. L'analisi funzionale, attività immediatamente successiva, è volta, a calare le specifiche teoriche in una serie di attività pratiche da implementare in seguito. Il deliverable di questa fase è dato da un documento formale, condiviso con il cliente e contenente una descrizione mirata e puntuale di tutte le specifiche emerse in fase di raccolta dei requisiti. Le attività tecniche ed operative, successive alla validazione del documento precedentemente citato, occupano l'intervallo temporale più esteso del progetto e si concretizzano in tutti i singoli task che concorrono alla costruzione del risultato finale e che verranno miratamente esposti nei paragrafi successivi. Tutti gli sviluppi sono opportunamente riportati in un documento tecnico condiviso con il cliente.

L'ultima fase corrisponde al rilascio del progetto, con una precedente fase di test e controllo, attestante la corretta implementazione di tutti gli interventi richiesti e, quindi, il pieno soddisfacimento dei requisiti.

3.2.1 Raccolta dei Requisiti e Analisi Funzionale

Tramite incontri sistematici con il cliente si formalizzano tutte le specifiche di progetto. Una chiara definizione e condivisione dei requisiti (sintetizzati in Figura 3) risulta, infatti, fondamentale per una corretta implementazione degli sviluppi operativi.

ID	ATTIVITA'
1	Query di import dal database con opportuna perimetrazione richiesta
2	Attività di pulizia dei dati e gestione dei campi non adeguatamente valorizzati
3	Creazione di due tabelle statiche a supporto degli sheet successivamente creati in Power Bi
4	Creazione di colonne aggiuntive calcolate
5	Creazione di misure calcolate
6	Generazione di un foglio Excel collegato alla base dati tabular
7	Connessione dell'ambiente di Power Bi con la base dati creata in Visual Studio in modalità "live"
8	Costruzione Report Power Bi
9	Creazione del meccanismo di aggiornamento dei dati
10	Creazione e condivisione della documentazione ufficiale di Analisi Funzionale e Analisi Tecnica

Figura 3. Sintesi dei Requisiti Raccolti

3.2.2 Sviluppi operativi e Analisi Tecnica

Dopo aver completato la fase di raccolta dei requisiti, hanno inizio le attività operative. La connessione ad una base dati è il primo step che viene effettuato dopo la creazione di un nuovo file di progetto; da Visual studio viene settata la connessione al Database Oracle di riferimento dal quale vengono importati i dati.

Dopo aver scelto la modalità di connessione alla fonte dati, viene definito il perimetro dell'operatività tramite una query di import realizzata attraverso una porzione di codice di tipo PL/SQL. La fase di import avviene tramite l'utilizzo di Power Query, interfaccia grafica che rende possibile l'importazione e la trasformazione dei dati. Viene quindi importata la tabella alimentante, contenente il perimetro di interesse, che rappresenta la base per la creazione delle due tabelle calcolate richieste. A partire dalla tabella originaria, infatti, vengono definite due tabelle statiche di dettaglio, ottenute filtrando la base dati per un determinato campo di interesse. Inoltre, dalla raccolta dei requisiti emerge la richiesta relativa all'aggiunta di nuove colonne calcolate ad integrazione della base dati già esistente. Tali colonne vengono valorizzate tramite opportune logiche condivise dal cliente (esempio riportato in Figura 4)

regola 2° livello: SOTTOVOCE DI BILANCIO_CALC			
if:	&		then:
BILECO_CODE_1	utenti particolari	RIGHT (MACCOA;3)	
15	*BROKEN	*	15-rateo#BROKEN
80	*BROKEN	*	80-RN#BROKEN
90	*BROKEN	*	90-DFV#BROKEN
15	ACC002BDTA	*	15-rateo#integrazione continuing
80	ACC002BDTA	*	80-RN#integrazione continuing
90	ACC002BDTA	*	90-DFV#integrazione continuing
15	<> utenti particolari	INT	15-rateo
15	<> utenti particolari	CAM	15-rateo#CA
80	<> utenti particolari	CCS	80-RN#comp cambi ccs
80	<> utenti particolari	NPV	80-RN
90	<> utenti particolari	CAM	90-DFV#CA
90	<> utenti particolari	CCS	90-DFV#comp cambi ccs
90	<> utenti particolari	INT	90-DFV
90	<> utenti particolari	NPV	90-DFV
15	<> utenti particolari	FEW	15-rateo#unwinding fees
15	<> utenti particolari	FEK	15-rateo#unwinding fees
15	<> utenti particolari	FEU	15-rateo#unwinding fees
15	<> utenti particolari	FEP	15-rateo#unwinding fees
80	<> utenti particolari	FEW	80-RN#unwinding fees
80	<> utenti particolari	FEK	80-RN#unwinding fees
80	<> utenti particolari	FEU	80-RN#unwinding fees
80	<> utenti particolari	FEP	80-RN#unwinding fees
90	<> utenti particolari	FEW	90-DFV#unwinding fees
90	<> utenti particolari	FEK	90-DFV#unwinding fees

Figura 4. Requisito relativo alla generazione di una nuova colonna calcolata

La logica di valorizzazione, esposta nell'immagine, rappresenta la base per la generazione del codice in linguaggio DAX predisposto nel modello dati tabular (Figura 5).

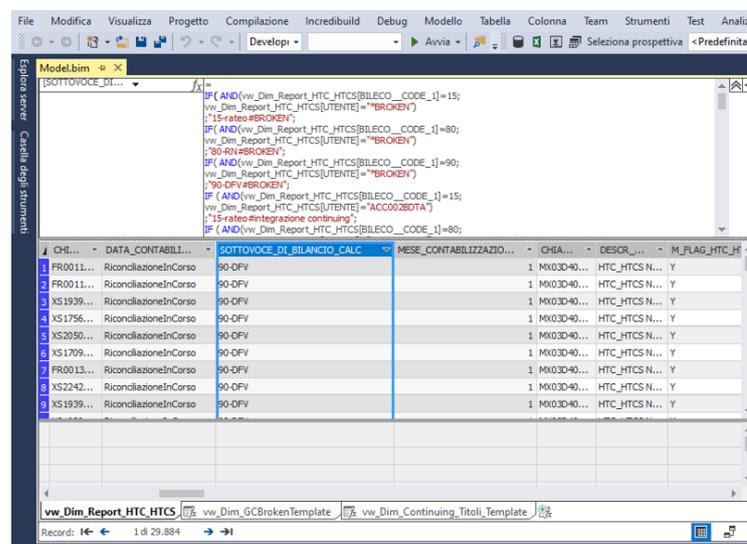


Figura 5. Codice DAX per la generazione della nuova colonna calcolata

La caratteristica principale dei dati importati risulta essere la dinamicità: per tale ragione, un sistema di questo tipo deve necessariamente prevedere un meccanismo che possa aggiornare i dati in ogni momento. A supporto di tale operatività, viene costruita una colonna calcolata, con l'obiettivo di fornire un'indicazione dell'ultimo aggiornamento dati effettuato e, quindi, dell'orizzonte temporale con il quale il modello risulta essere allineato. Tale informazione risulta essere fondamentale per un controllo dell'effettivo funzionamento

del meccanismo, ma anche per fornire una notifica al cliente che, in ogni foglio della reportistica, può facilmente visualizzare la data e l'orario dell'ultimo refresh del modello dati. La colonna viene creata attraverso una porzione di codice aggiunto in Power Query, mentre la costruzione pratica del meccanismo di aggiornamento, che è alla base, verrà successivamente spiegata.

Dopo aver terminato la costruzione di tutti i nuovi campi richiesti, vengono definite le misure del modello intese come tutti i valori aggregati di colonna. Tale azione risulta di supporto al cliente nell'individuazione di eventuali segnali di allarme e per riscontrare facilmente anomalie nei dati finanziari.

Dopo aver terminato la fase di manipolazione dei dati, attraverso tutte le elaborazioni precedentemente discusse, viene predisposto per il cliente un file Excel contenente l'intero modello dati costruito su Visual Studio. Il foglio Excel viene, quindi, collegato al database del modello tabulare ed agganciato in "Live Connection" in modo tale da fornire in real-time una base dati aggiornata. La consegna di questo documento avviene con l'obiettivo di ottimizzare la User-Experience, fornendo al cliente la possibilità di manipolare i dati in maniera autonoma.

Dopo aver terminato la parte di sviluppi, relativa alla costruzione di un modello dati in accordo con i criteri di Data Quality, si procede con la creazione della Reportistica, attraverso Power Bi. Il primo step da implementare, anche in questo ambiente, è la connessione ad una base dati; in questo caso, però, l'origine dati corrisponde al progetto precedentemente creato tramite Visual Studio. Si creano, quindi, visualizzazioni tabellari e grafici per ogni foglio richiesto. Ognuno di essi riflette le specifiche fornite e contiene l'esposizione di tabelle Pivot con opportuni filtri di pagina e/o di tabella. Viene, inoltre, resa possibile la visualizzazione, per ogni foglio, dell'ultimo aggiornamento dati effettuato (informazione presente in Figura 6 come "DATA ULTIMO REFRESH").

MACROVOCE: DERIVATI OTC FLAG HTC HTCS: Y DATA ULTIMO REFRESH: 20/09/2022 17:00:56

COD_LEGAME_CALC	SIST_PROVENIENZA_ECOSYS	FASE	FAMILY	UNIT	FORMA_TECNICA	DTA_RIC	PL_EUR
28650	MR3	NOFX	DCBAB	OTC	31/01/2022	16.228.06	
28661	MR3	NOFX	DCBAB	OTC	31/01/2022	65.277.51	
28662	MR3	NOFX	DCBAB	OTC	31/01/2022	46.733.58	
28663	MR3	NOFX	DCBAB	OTC	31/01/2022	24.019.86	
28664	MR3	NOFX	DCBAB	OTC	31/01/2022	50.631.20	
28665	MR3	NOFX	DCBAB	OTC	31/01/2022	50.596.20	
28666	MR3	NOFX	DCBAB	OTC	31/01/2022	17.896.80	
28667	MR3	NOFX	DCBAB	OTC	31/01/2022	12.692.91	
28669	MR3	NOFX	DCBAB	OTC	31/01/2022	63.859.68	
28670	MR3	NOFX	DCBAB	OTC	31/01/2022	12.035.29	
28671	MR3	NOFX	DCBAB	OTC	31/01/2022	17.553.88	
28672	MR3	NOFX	DCBAB	OTC	31/01/2022	32.111.04	
28673	MR3	NOFX	DCBAB	OTC	31/01/2022	33.165.06	
28674	MR3	NOFX	DCBAB	OTC	31/01/2022	22.373.54	
28675	MR3	NOFX	DCBAB	OTC	31/01/2022	6.056.85	
TOTAL						538.158.608,66	

ANAGRAFICA ANAGRAFICA-CHECK GC BROKEN-CHECK GC BROKEN-TEMPLATE INTEGRAZIONE CONTINUING CONTINUING TITOLI-TEMPLATE APERTURA MI

Figura 6. Reportistica in Power Bi

Dopo aver modellato la base dati estratta e collegato la reportistica associata, si procede con la creazione e lo sviluppo del meccanismo di aggiornamento dati. La logica da implementare parte dalla presenza di una tabella detta “Tabella di Pooling” e dall’inserimento, in quest’ultima, di righe dette “token”. La tabella appena citata viene aggiornata a partire da uno script che, una volta lanciato, la interroga e controlla l’ultimo record inserito, focalizzando l’attenzione sulla valorizzazione della colonna “Stato”. Se lo stato corrisponde ad “Inserito”, viene avviato il processo di refresh, che permette un aggiornamento dei dati presenti nel modello; a questo punto lo stato passa a “In lavorazione” e poi a “Completato”. L’evento trigger che innesca il processo è, dunque, l’inserimento del token, innescato manualmente; è il cliente, infatti, che, in maniera autonoma, premendo un apposito bottone sulla propria interfaccia web, dà inizio all’aggiornamento dati. Quest’azione scatena l’inserimento di un token che assume inizialmente lo stato “Inserito” e, al termine del processo di aggiornamento, assume lo stato “Completato”. Al raggiungimento di quest’ultimo avrà termine l’operazione di aggiornamento dati e verrà inviata una mail al cliente con due possibili esiti: “Refresh eseguito con successo” oppure “Refresh eseguito con errore”. Il processo è schematizzato in Figura 7.

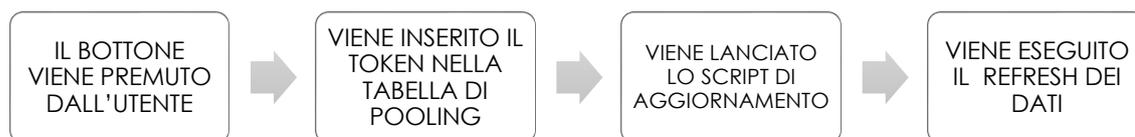


Figura 7. Logica di Refresh

3.2.3 Rilascio e UAT (User Acceptance Testing)

Il rilascio della soluzione proposta ha inizio con l'inserimento di tutti gli oggetti creati, che concorrono alla generazione della soluzione richiesta, all'interno di un'apposita interfaccia web, messa a disposizione dal Gruppo Bancario. Per la tracciabilità di tutte le attività promosse viene, quindi, creata la cosiddetta Unità di Cambiamento, ossia un pacchetto di richieste relative agli oggetti rilasciati.

La fase di UAT ha inizio al termine di tutte le attività di sviluppo e dell'inserimento degli oggetti creati nel portale precedentemente citato; in questa fase, il cliente effettua dei test al fine di validare la correttezza della soluzione proposta, per procedere, poi, con l'ultima fase di rilascio definitivo della soluzione. L'interazione con il cliente, in questa fase, risulta, infatti, frequente ed è volta all'affiancamento, con l'obiettivo di verificare tutte le funzionalità richieste e la conseguente aderenza ai requisiti concordati in fase di raccolta requisiti. La fase di UAT termina con l'esito positivo di tutti i casi d'uso implementati e con il successivo rilascio della soluzione.

4 Conclusioni

Il lavoro di tesi presentato ha portato all'implementazione di molteplici attività trasversali; il mantenimento e la cura di uno scambio informativo attivo e puntuale con l'azienda cliente, lo sviluppo tecnico dei requisiti richiesti attraverso l'utilizzo di vari tool e la stesura di documenti ufficiali sono solo alcune delle attività che hanno portato alla definizione della soluzione finale. L'obiettivo di aiutare il Gruppo Bancario nella definizione di un framework di governo del dato è stato perseguito e correttamente raggiunto per il perimetro del progetto in esame. La soluzione rilasciata inoltre si è dimostrata uno strumento fondamentale per la semplificazione delle innumerevoli attività manuali svolte dal cliente nell'operatività quotidiana. Oltre al beneficio apportato in termini di tempo, la cura e la gestione oculata dei dati ha reso possibile la definizione di una base dati affidabile e di qualità sulla quale improntare corrette scelte strategiche e di business. Attraverso il Report creato in Power Bi, inoltre, il cliente è stato in grado di cogliere prontamente tutte le anomalie presenti nei dati attraverso la sola visualizzazione grafica e tabellare.

Relativamente al caso in esame, viene proposta una seconda fase ricca di "Evolutive" ossia interventi di evoluzione ed ampliamento rispetto alla soluzione già implementata. Gli sviluppi richiesti mirano quindi ad allargare il perimetro d'azione e ad arricchire il modello

esistente fornendo una visualizzazione, tramite l'aggiunta di nuovi fogli nella reportistica esistente, di più ampio spettro. Le necessità che muovono le nuove richieste condivise dal cliente coincidono con il bisogno di accedere in maniera rapida e continua ad uno strumento che possa fornire una panoramica degli aspetti critici e di miglioramento diminuendo tutte le possibili manualità effettuate dal cliente.