

# **IT RISK and Data Quality Risk**

**Claudio Ruffini**

**22 Giugno 2016**



LA FINANZA. INTERPRETATA A REGOLA D'ARTE.

**Augeos**



# Il concetto di Data Quality nel contesto IT

- La recente attenzione rivolta dal legislatore sulla «Data Quality» testimonia la centralità della **qualità dei dati nella moderna Governance dei rischi**.
- Con l'aggiornamento alla Circolare 263, la Banca d'Italia ha ufficialmente introdotto tale concetto, imponendo corrette modalità di **misurazione della qualità**, individuazione di soglie limite, anche e soprattutto con riferimento al mondo alla gestione del rischio informatico.
- Tuttavia, individuare **dimensioni, metodologie e metriche** efficaci presuppone una notevole visione strategica e una certa competenza tecnica che non sempre sono a portata di mano.



# Dimensioni della Data Quality

## COBIT (confronto 4.1 – 5)

### Cataloghi

- **Efficacia:** *l'informazione è efficace se soddisfa le esigenze del consumatore delle informazioni, che utilizza le informazioni per uno specifico compito. Se il consumatore delle informazioni riesce a eseguire il compito grazie alle informazioni recepite, allora l'informazione è efficace. Ciò corrisponde ai seguenti obiettivi di qualità dell'informazione: quantità appropriata, pertinenza, comprensibilità, interpretabilità, obiettività;*

**Efficienza:** *mentre l'efficacia considera l'informazione come un prodotto, l'efficienza si riferisce più al processo di ricezione e utilizzo delle informazioni, adottando una visione dell'informazione come "servizio". Se le informazioni che soddisfano le esigenze di informazione del consumatore vengono ottenute e utilizzate in modo semplice (comportando quindi poche risorse, sforzo fisico, sforzo cognitivo, tempo e denaro), allora l'uso delle informazioni è efficiente. Ciò corrisponde ai seguenti obiettivi di qualità dell'informazione: credibilità, accessibilità, facilità di utilizzo, reputazione;*

**Integrità:** *l'informazione è integra quando è priva di errori e completa. Ciò corrisponde ai seguenti obiettivi di qualità dell'informazione: completezza, precisione;*

**Affidabilità:** *l'affidabilità è spesso vista come sinonimo di precisione, tuttavia, si può anche dire che l'informazione è affidabile se è considerata vera e credibile. Rispetto all'integrità, l'affidabilità è più soggettiva, più legata alla percezione che non al dato di fatto. Ciò corrisponde ai seguenti obiettivi di qualità dell'informazione: credibilità, reputazione, obiettività;*

**Disponibilità:** *La disponibilità è uno degli obiettivi di qualità dell'informazione sotto la voce di accessibilità e sicurezza;*

**Riservatezza:** *La riservatezza corrisponde all'obiettivo di qualità dell'informazione relativo alla limitazione dell'accesso;*

**Conformità:** *La conformità intesa nel senso che l'informazione deve essere conforme alle specifiche riguarda ognuno degli obiettivi di qualità dell'informazione, a seconda dei requisiti. La conformità alle normative è spesso un obiettivo o requisito dell'utilizzo delle informazioni, non tanto una qualità intrinseca delle informazioni.*



# Dimensioni della Data Quality

ISO/IEC 25012 2008

- Caratteristiche *inerenti* il dato:
  1. **Accuratezza**: perfetta rispondenza con il mondo reale che rappresenta;
  2. **Attualità**: il tempo in cui il dato è utilizzato;
  3. **Coerenza**: dato non contraddittorio con altri dati;
  4. **Completezza**: presenza di tutti gli attributi necessari;
  5. **Credibilità**: provenienza da fonte certa.
- Caratteristiche *inerenti e dipendenti* dal sistema:
  1. **Accessibilità**: il dato è accessibile a tutti;
  2. **Comprensibilità**: il significato del dato immediato e chiaro;
  3. **Conformità**: il dato risponde a regolamentazioni;
  4. **Efficienza**: il dato è utilizzabile con risorse;
  5. **Precisione**: il dato è del livello di misura richiesto;
  6. **Riservatezza**: il dato può essere utilizzato;
  7. **Tracciabilità**: gli accessi al dato sono registrati.
- Caratteristiche *dipendenti* dal sistema:
  1. **Disponibilità**: dato disponibile e re-interrogabile;
  2. **Portabilità**: capacità del dato di migrare da un sistema a un altro;
  3. **Ricoverabilità**: dato in ambiente sicuro e recuperabile.

Sappiamo  
dunque  
COSA  
misurare



Ma il COME?

Esigenze:

- Sensibilità
- Costo
- Usabilità
- Stabilità
- Scientificità
- ...



# Indicatori DQ

- Facciamo qualche esempio semplice
  - Regole di integrità
  - Indicatori che verificano che una informazione/documento abbia la struttura o uno schema secondo standard
  - Scostamento rispetto a DB benchmark



# Metodologie e Metriche della Data Quality

*Batini e Scannapieco* classificano le metodologie per la qualità dei dati secondo differenti criteri:

- **Misurazione vs Miglioramento:** le prime sono metodologie utili a valutare la qualità dei dati, le seconde a migliorarla;
- **Generali vs Specializzate:** le prime comprendono un'ampia gamma di fasi e attività, le seconde riguardano attività specifiche su uno o più domini applicativi
- **Data-driven vs Process-driven:** le prime sono metodologie costituite dall'acquisizione diretta dei dati e basate su record matching e vincoli di integrità, le seconde sono incentrate sui processi e sulla qualità di questi;



# Dilemma

## Qualità del dato:

- PRO: è una tecnica che va sempre bene anche quando i dati vengono acquisiti dall'esterno.
- CONTRO: è generica e spesso ha bisogno di un confronto con il dato corretto che è difficile da ottenere



## Qualità del processo che produce il dato

- PRO: miglioramenti strutturali e permanenti se diventano obiettivi per il processo stesso
- CONTRO: funziona solo per i dati prodotti dalla propria organizzazione



- Misurare la qualità dei dati vuol dire anche dover intervenire e correggere?
- A chi e per cosa possono servire i risultati dell'analisi della qualità dei dati?
- I valori delle misure della qualità possono servirmi per migliorare la mia operatività?
- Qual è il costo della qualità? Quali i benefici quantificabili?



# Tre ambiti di applicazioni

DQ come  
KRI per i  
Rischi  
Operativi

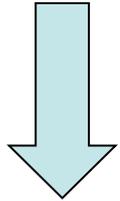
DQ come  
parte di una  
analisi dei  
rischi IT

Monitor degli  
eventi  
operativi e  
Alert

Migliorare la governance dei Rischi e  
Controlli della banca.



# Tre ambiti di applicazioni



DQ come  
KRI per i  
Rischi  
Operativi

DQ come  
parte di una  
analisi dei  
rischi IT

Monitor degli  
eventi  
operativi e  
Alert

Migliorare la governance dei Rischi e  
Controlli della banca.



# Correlazione eventi di perdita / KRI

- Nella moderna logica della gestione del rischio, i Key Risk Indicators (KRI) fungono da indicatori sintetici in grado di rilevare un rischio in un dato momento.
- Con riferimento al rischio (operativo, IT, ...), i KRI possono adottare i risultati delle metodologie per la qualità del dato in un cruscotto sintetico



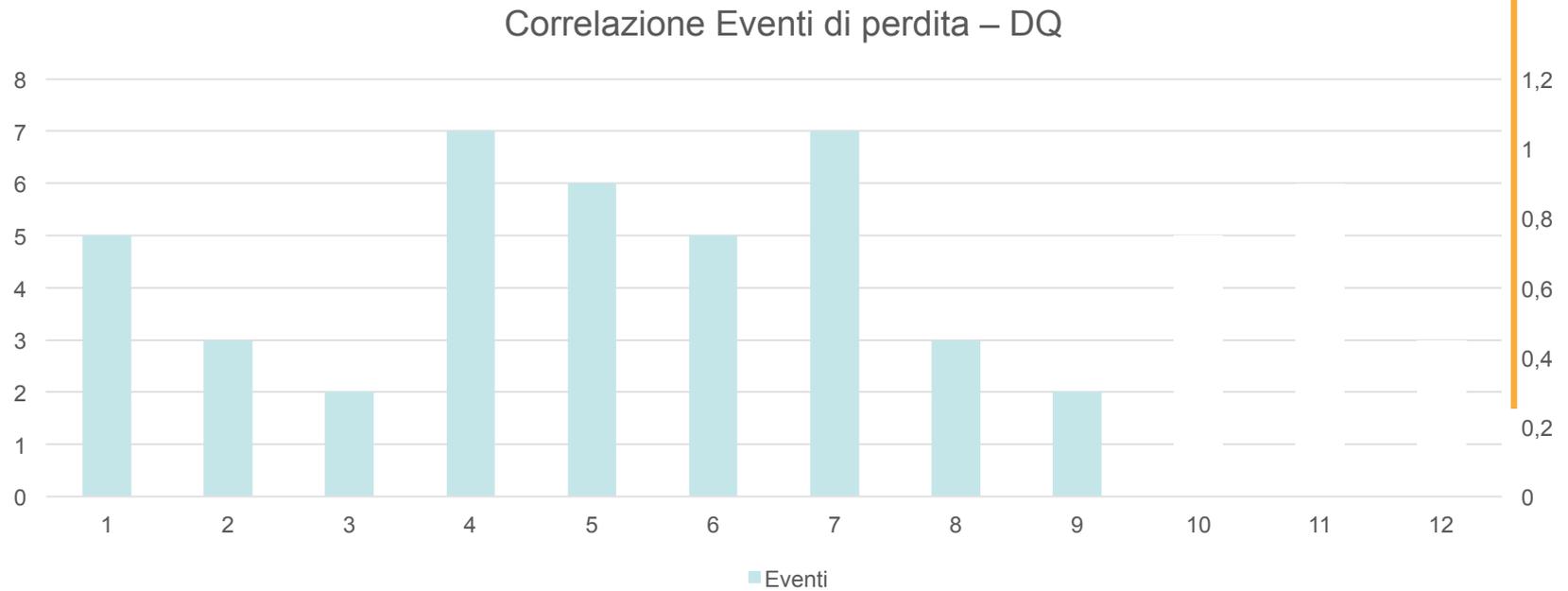
# Correlare dati



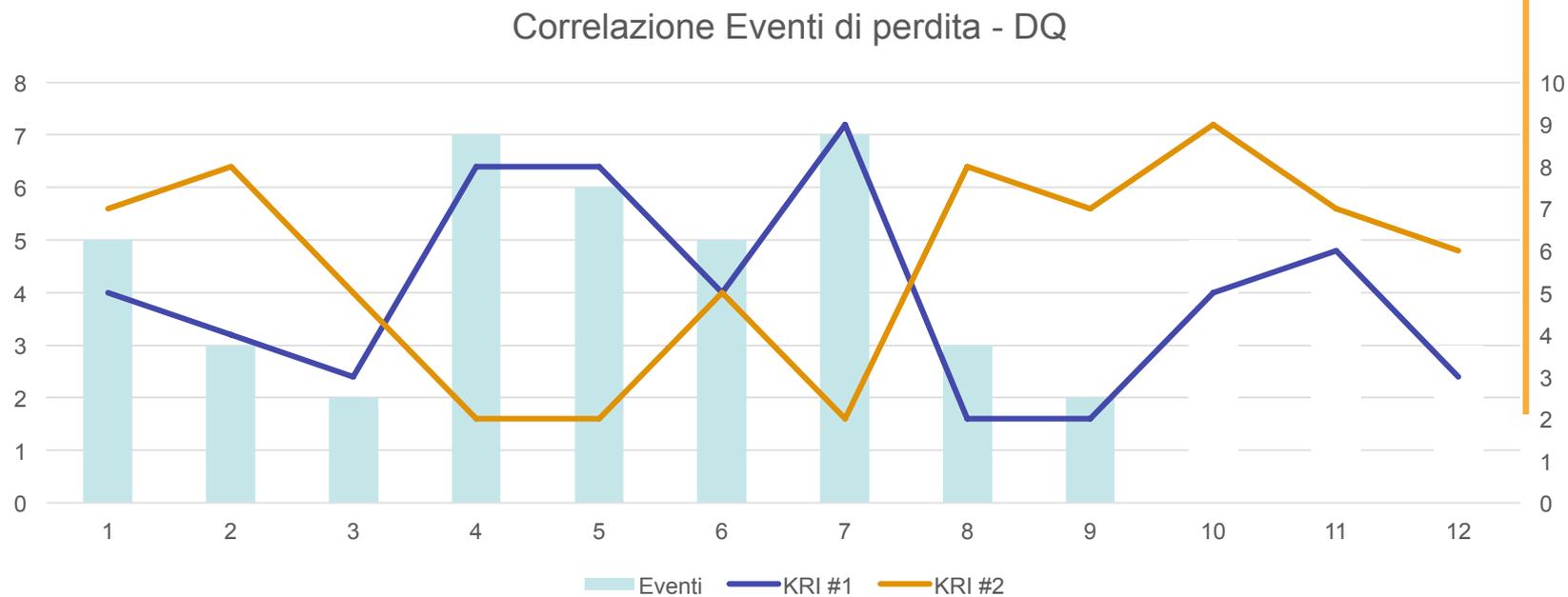
- L'idea chiave è quella di studiare le misure storiche rilevate di DQ in relazione agli eventi di perdita nello stesso periodo di tempo
- Mettere a disposizione un ambiente di analisi che permetta di selezionare, analizzare **correlazioni** ed effettuare simulazioni di indicatori predittivi



# Correlazione eventi di perdita / DQ

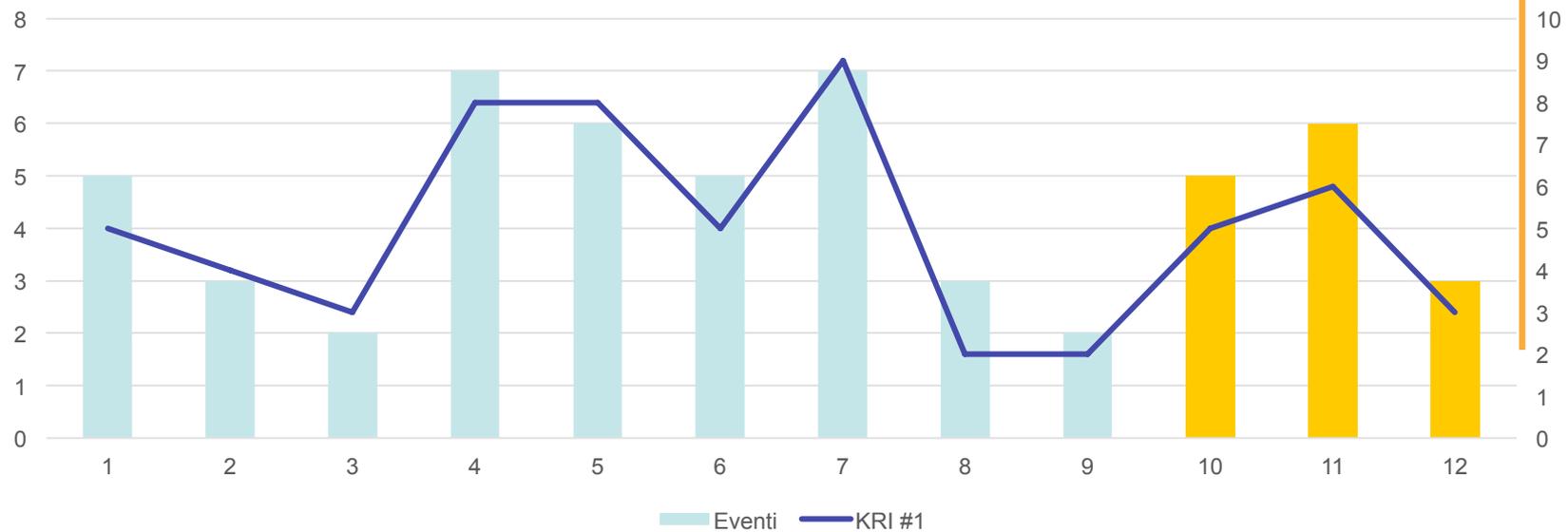


# Correlazione eventi di perdita / DQ



# Correlazione eventi di perdita / DQ

Correlazione Eventi di perdita – DQ che diventa KRI



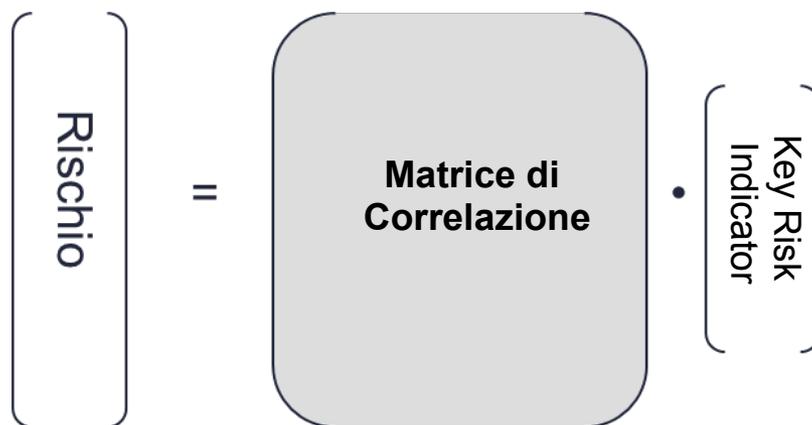
# Problema delle soglie

Per determinare delle soglie in grado di fornire un alert da possibili eventi di perdita è possibile utilizzare anche la **metodologia Augeos BT4Risk®** già descritta in un precedente intervento DIPO di qualche anno fa e frutto di un progetto di ricerca effettuato con l'università Bicocca di Milano



## Risk measurement: misure di KRI

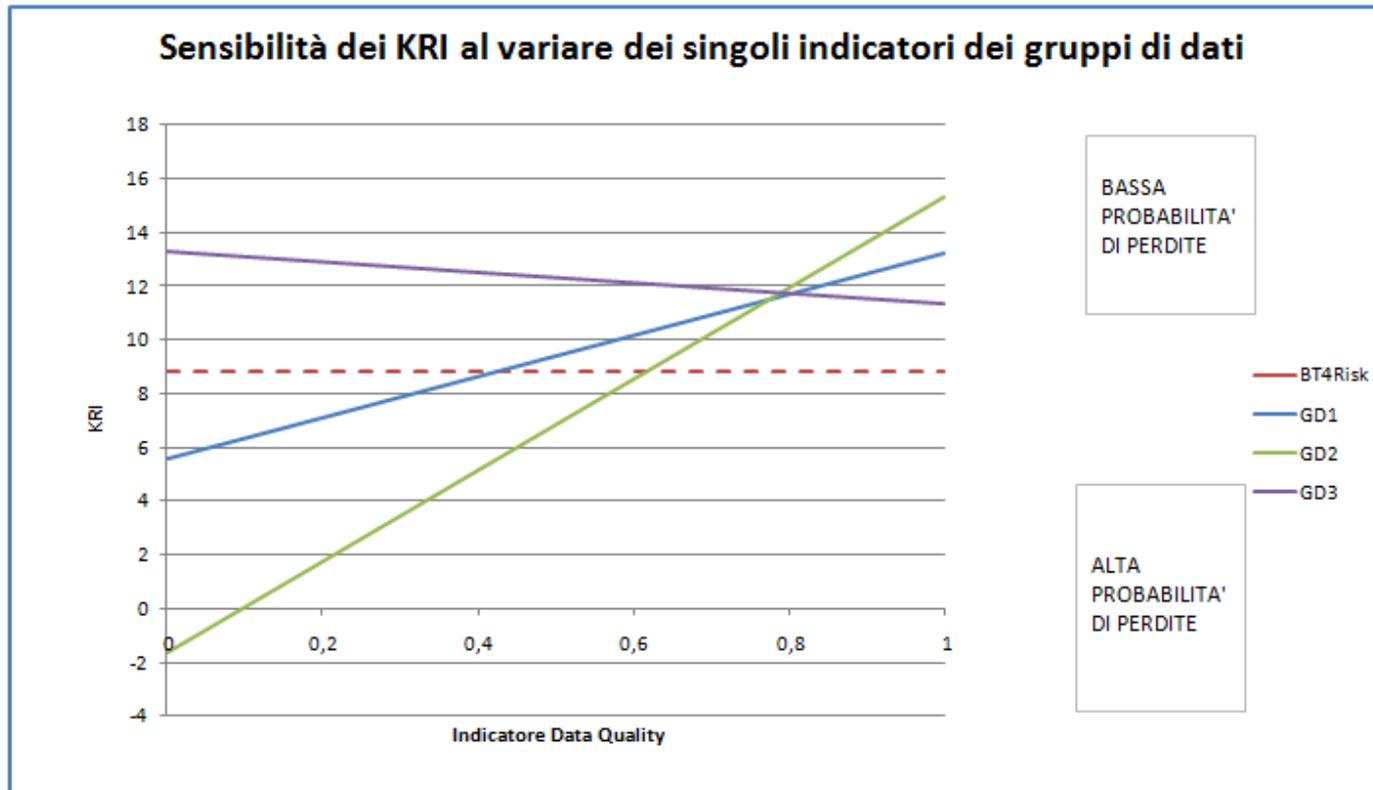
- al variare del valore del rischio si esaminano come variano gli indicatori che rappresentano i KRI rilevati tramite apposite sonde permettendo di ottenere coefficienti di correlazione



- una volta ottenuta la Matrice di Correlazione tra KRI e probabilità di rischio sarà possibile predire la variazione del Rischio a partire dalle misure di KRI



# Risultati

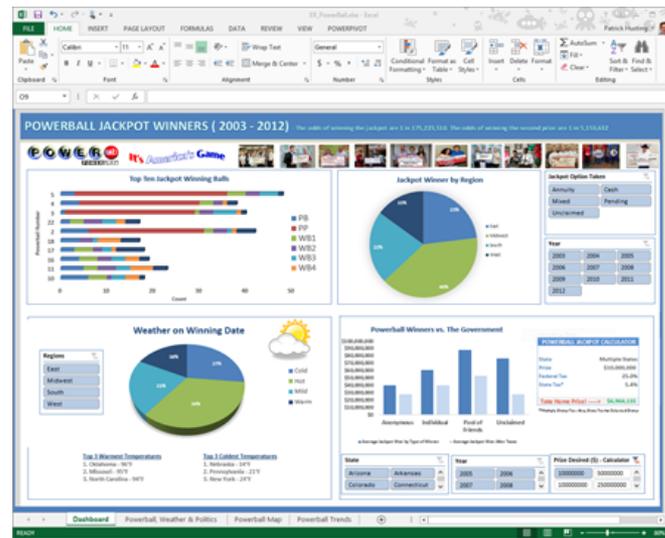


il gruppo di dati 3 non concorre alla generazione di alert, mentre si verifica una relazione negativa tra l'indicatore di qualità dei database 1 e 2 e il valore dell'indicatore KRI assegnato, il quale conduce a segnali di alert al di sotto di certi valori.

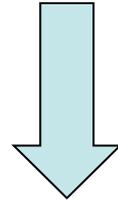


# Risultati

Le misure di qualità dei dati e delle informazioni sono degli ottimi KRI predittivi per i rischi operativi



# Tre ambiti di applicazioni



DQ come  
KRI per i  
Rischi  
Operativi

DQ come  
parte di una  
analisi dei  
rischi IT

Monitor degli  
eventi  
operativi e  
Alert

Migliorare la governance dei Rischi e  
Controlli della banca.



# IT Risk Mapping – Analisi Top down

L'analisi **TOP-DOWN** del rischio IT è focalizzata sulle vulnerabilità degli asset informatici e le possibili minacce correlate per formare uno scenario di rischio potenziale;

	Threat cat 1	Threat Cat 2	Threat Cat 3	...	Threat Cat N
Asset n. 1					X
Asset n. 2		X			
Asset n. 3	X		X		
...					
Asset n. N			X		



# Tipi di Risorse di un processo

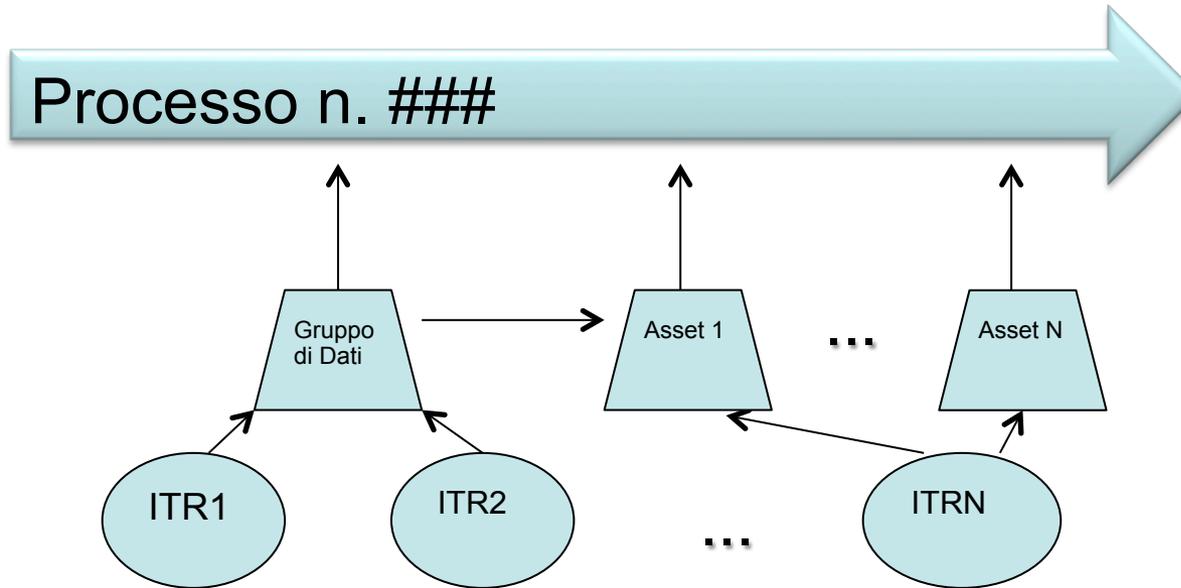
## Esempi di Risorse

- Asset informatici It o Hardware
- Asset Software, prodotti o servizi
- Dati e informazioni
- Strumenti o macchinari
- Persone
- ...

**IT Risk  
Analysis**



# IT Risk Mapping



I Dati e le informazioni rientrano a pieno titolo tra gli elementi di un **IT Risk Management**. Il raccordo con i processi avviene identificando preventivamente le risorse associate a ciascun processo.

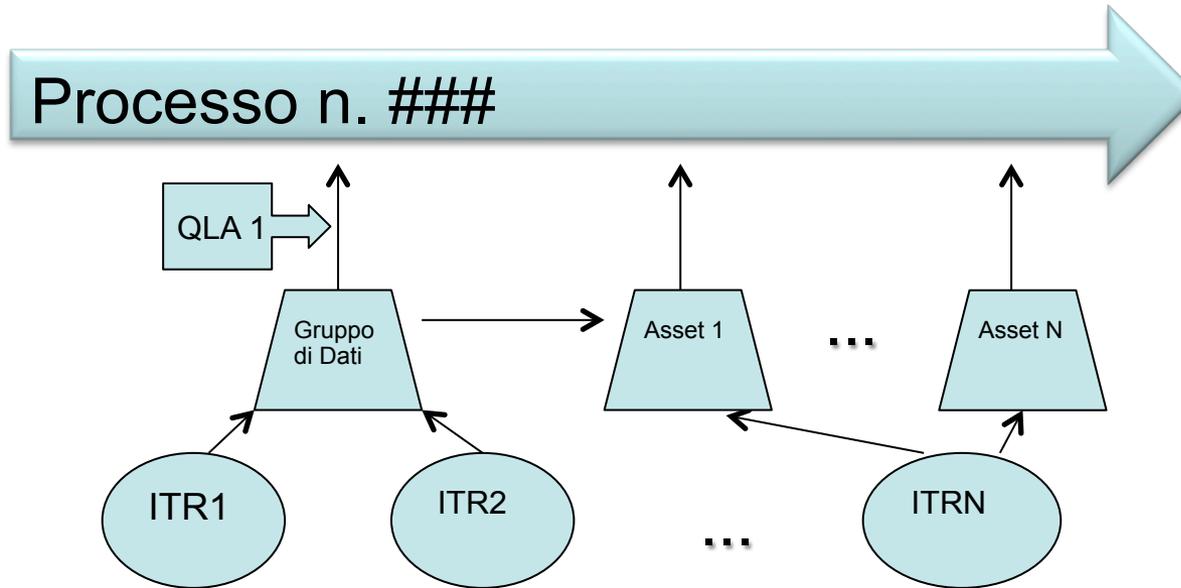


# Raccordo al processo

**Il processo** è l'elemento chiave a cui ricondurre tutte le analisi e gli elementi dei vari modelli (framework) delle singole funzioni della banca



# IT Risk Mapping



Il legame con il processo avviene con un **QLA** (Quality Level Agreement) che rappresenta la base per l'elemento di scenario analysis del processo



# Esempio

- Il processo per il calcolo del RAF per esempio ha bisogno di un gruppo di dati amministrativi dei titoli con informazioni relative ad almeno 2 dei 3 rating disponibili con al massimo un ritardo di 2 gg.
- Il QLA del gruppo di dati di rating è per esempio del tipo
  - Disponibilità 99%
  - Completezza 66%
  - Attualità max 2 gg
  - Accuratezza: 95%

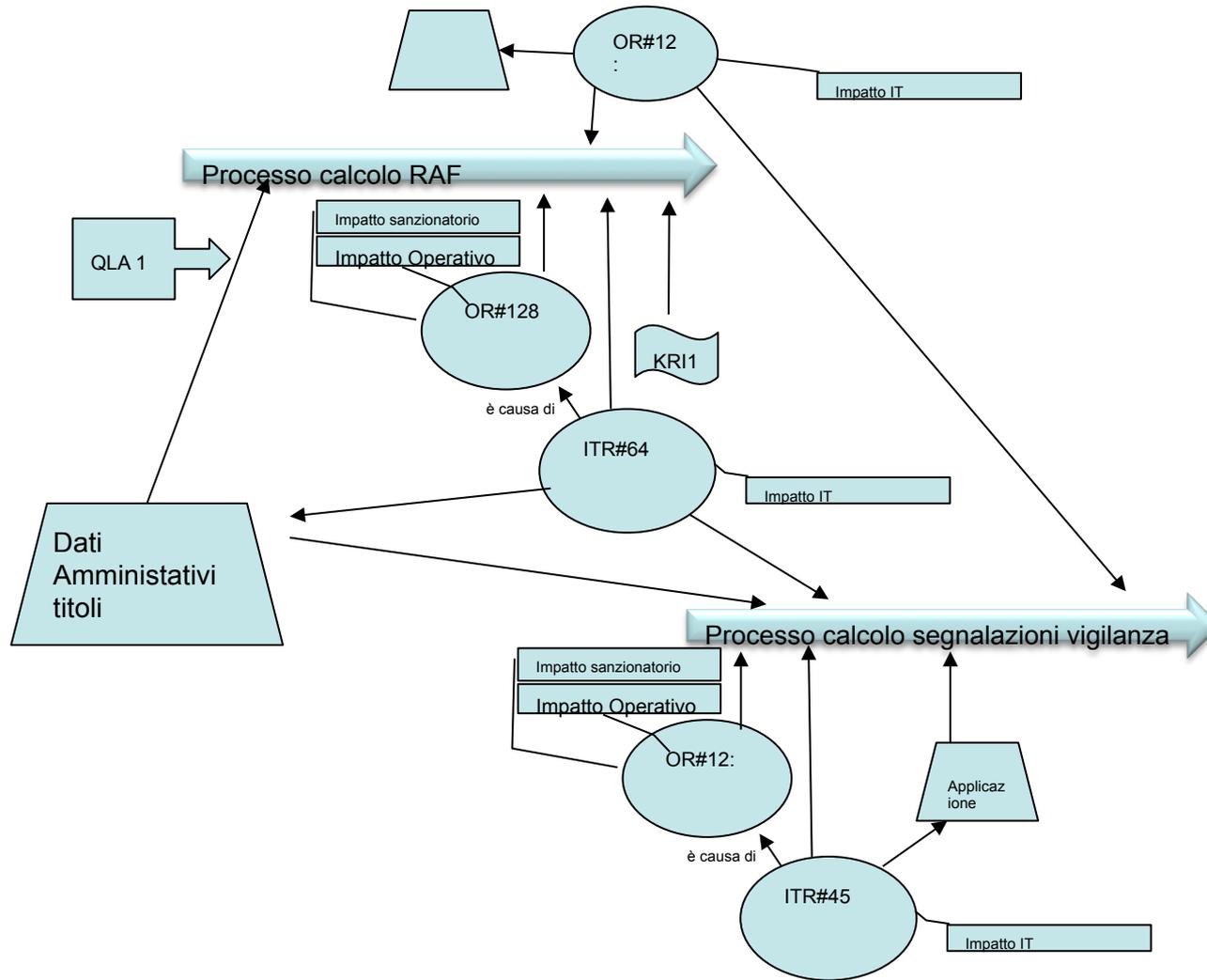


# Analisi di rischio DQ

- A questo punto dell'analisi possiamo individuare dei possibili **scenari di rischio operativo da DQ** per ciascuno dei possibili eventi di non rispetto dei QLA nell'ipotesi che ciascuno delle relative violazioni possa incidere sugli obiettivi del processo
- E' possibile inserire dei KRI per ciascuno dei valori di qualità dei dati
- E' possibile individuare dei **controlli** con delle possibili contromisure per mitigare il rischio corrispondente



# Network Analysis Risk event

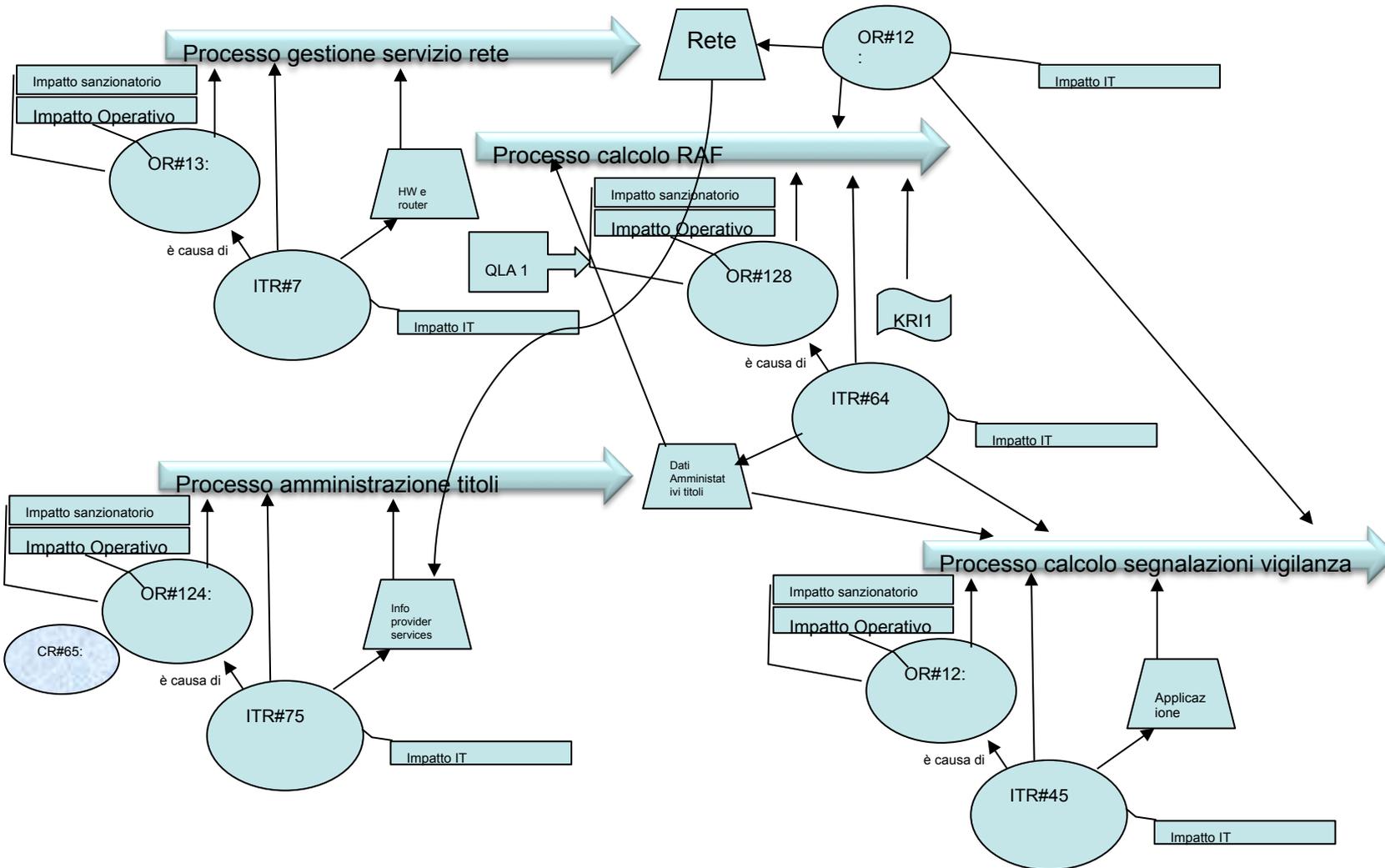


# Qualità del processo che produce il dato

- Per rendere più persistenti le azioni di mitigazioni occorre proseguire con l'analisi e occuparsi del **processo che genera il gruppo di dati** che ci interessa.
- Tarare gli obiettivi del processo sugli obiettivi di QLA per il gruppo di dati (in genere il massimo tra tutti i QLA a cui il gruppo di dati è associato)



# DQ Risk Network Analysis



Tramite la DQ Risk Network Analysis  
possiamo creare delle azioni di  
mitigazioni non solo sui dati ma anche sui  
processi che generano i dati  
**rendendo più efficace e persistente  
l'azione di mitigazione**



# IT Risk Mapping – Bottom-up

Con l'analisi **BOTTOM-UP** procediamo a ritroso partendo dalla descrizione di un evento accaduto e procediamo con raffinamenti successivi alla modellazione o completamento della DQ Risk Network Analysis



# Esempio di Evento

*Il 17/5/2016 alle ore 17 il Server 120.50.37.2 preposto alla raccolta dei flussi informativi da Info provider esterno ha smesso di funzionare causando un blocco nel processo di valorizzazione Fair Value per 1w.*



# Analisi semantica

Il 17/5/2016 alle ore 17 il **Server 120.50.37.2** preposto alla raccolta dei **flussi informativi** da Info provider esterno ha smesso di funzionare causando un blocco nel **processo di valorizzazione Fair Value** per 1w.

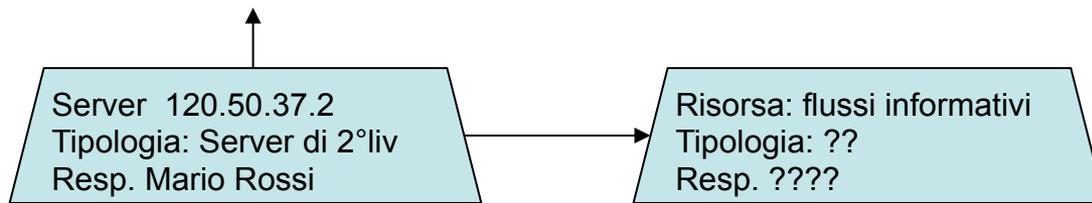
Il Server 120.50.37.2 è presente nel catalogo degli asset quindi viene collegato.

La risorsa «flussi informativi» deve essere completata con informazioni complementari per raffinare l'indagine

Supporto di tool semantici semi-automatici



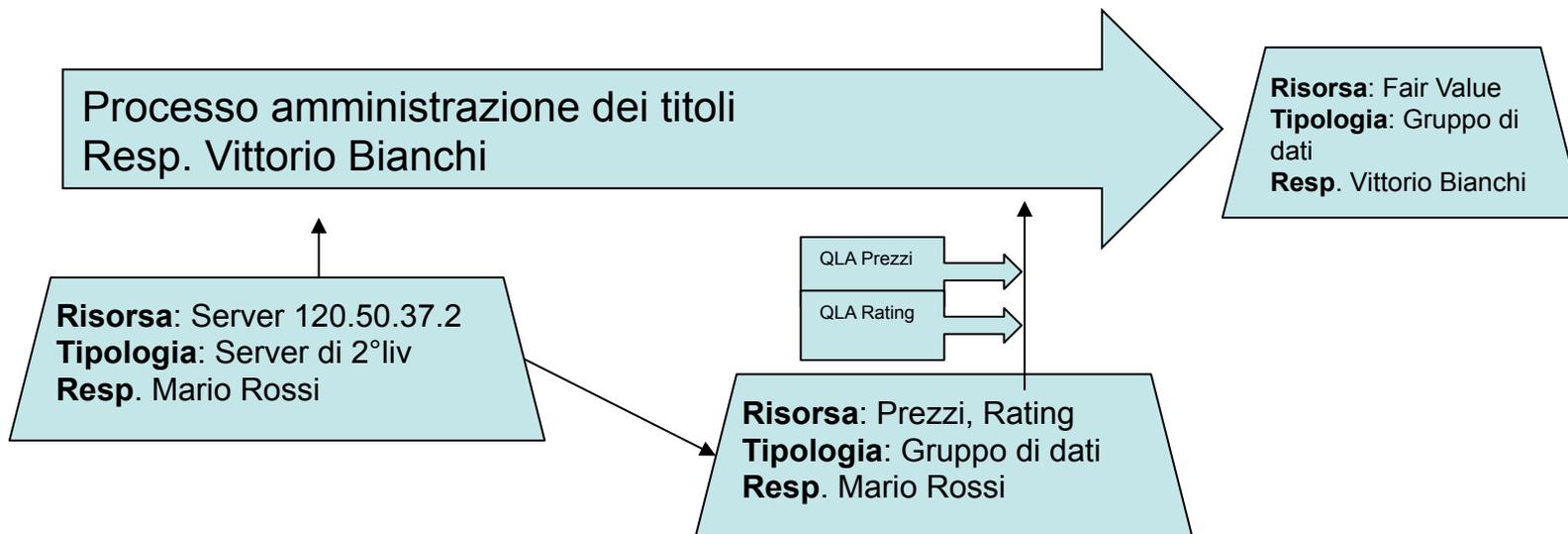
Processo amministrazione dei titoli  
Resp. Vittorio Bianchi



Il 17/5/2016 alle ore 17 il **Server 120.50.37.2** preposto alla raccolta dei **flussi informativi** da Info provider esterno ha smesso di funzionare causando un blocco nel processo di valorizzazione Fair Value per 1w.

Questo evento mi permette di fare delle analisi sullo schema che individuano un oggetto Risorsa probabilmente di tipo «Gruppo di Dati». Non solo ma mi dice anche a chi posso chiedere informazioni sia a Mario Rossi sia a Vittorio Bianchi





Il 17/5/2016 alle ore 17 il **Server 120.50.37.2** preposto alla raccolta dei **flussi informativi** da Info provider esterno ha smesso di funzionare causando un blocco nel processo di valorizzazione Fair Value per 1w.

Con interviste o suggerimenti del sistema si riesce a completare la risorsa



# Benefici

Posso introdurre dei **controlli permanenti** più efficaci avendo maggiore consapevolezza dei Dati che sto usando.

Riesco ad avere **una mappa dei processi e dei rischi** più rispondente alla realtà e focalizzata

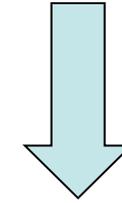
Posso determinare dei **KRI più mirati** per prevenire problemi operativi



L'analisi Bottom-up degli eventi su una mappa di **Risk Network Analysis** mi permette di migliorare la governance e prevenire possibili perdite operative.



# Tre ambiti di applicazioni



DQ come  
KRI per i  
Rischi  
Operativi

DQ come  
parte di una  
analisi dei  
rischi IT

Monitor degli  
eventi  
operativi e  
Alert

Migliorare la governance dei Rischi e  
Controlli della banca.



- Come posso valorizzare lo sforzo della DQ Risk Network Analysis ?
- Come posso utilizzare queste informazioni per migliorare la mia operatività



- Utilizziamo queste informazioni attraverso un monitor di eventi (simile ad un incident management System)
- Immaginiamo di poter avere una fase di raccolta di eventi di varie origini
  - Log
  - Mail
  - Verbali
  - Segnalazioni da LDC



- Tramite la semplice descrizione e i pochi dati che si hanno a disposizione (autore, fonte, data, ...) viene individuato un **referente** a cui associare l'evento.
- Il referente si ritroverà un evento in un monitor interattivo e dovrà procedere alla classificazione e individuazione degli elementi della Risk Network



# Esempio

Token 134

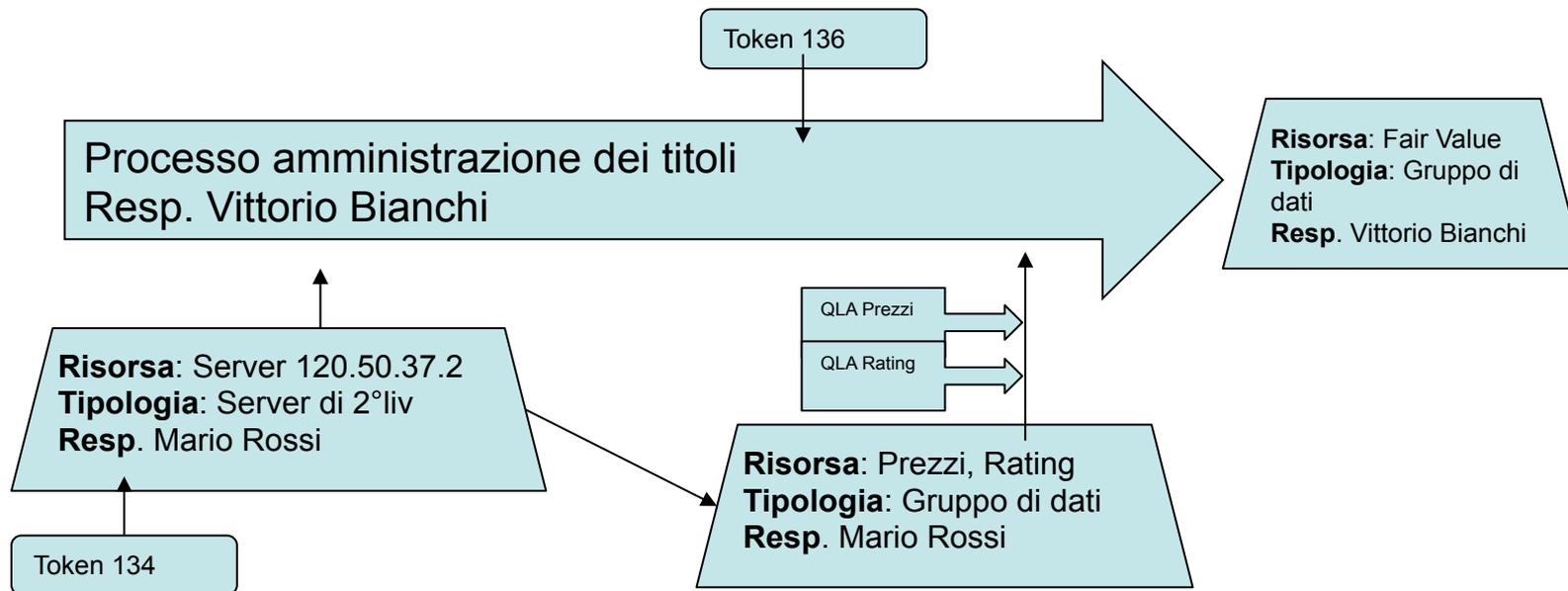
**Autore:** log

Il 17/5/2016 alle ore 17 il Server 120.50.37.2 preposto alla raccolta dei flussi informativi da Info provider esterno ha smesso di funzionare.

**Referente** Mario Rossi

Mario Rossi conferma che l'evento è afferente il Server di propria competenza (che rientra tra quelli censiti nella propria rete di competenza, altrimenti poteve censirlo).





Il 17/5/2016 alle ore 17 il **Server 120.50.37.2** preposto alla raccolta dei **flussi informativi** da Info provider esterno ha smesso di funzionare causando un blocco nel processo di valorizzazione Fair Value per 1w.



# Esempio

Mario Rossi giudica che il fermo macchina fa violare il QLA che vede associato al gruppo di dati relativo a Prezzi e al Rating

Token 136

**Autore:** Mario Rossi

*Violazione del QLA relativo a rating. Indisponibilità dei dati per un periodo superiore ai 2gg*

**Referente:** Vittorio Bianchi



# Esempio

Token 136

**Autore:** Mario Rossi

Violazione del QLA relativo a rating. Indisponibilità dei dati per un periodo superiore ai 2gg

**Referente:** Vittorio Bianchi

Vittorio Bianchi conferma che l'evento è afferente il Processo di propria competenza (che rientra tra quelli censiti nella propria area di competenza, altrimenti poteve censirlo).

Vittorio Bianchi eventualmente chiede info a Mario Rossi.

Vittorio Bianchi giudica se l'indisponibilità del rating causa una violazione dei suoi QLA

Eventualmente prende delle **contromisure** o censisce un **task di mitigazione** o di rinegoziazione dei QLA con Mario Rossi



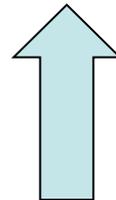
# Tre ambiti di applicazioni

DQ come  
KRI per i  
Rischi  
Operativi

DQ come  
parte di una  
analisi dei  
rischi IT

Monitor degli  
eventi  
operativi e  
Alert

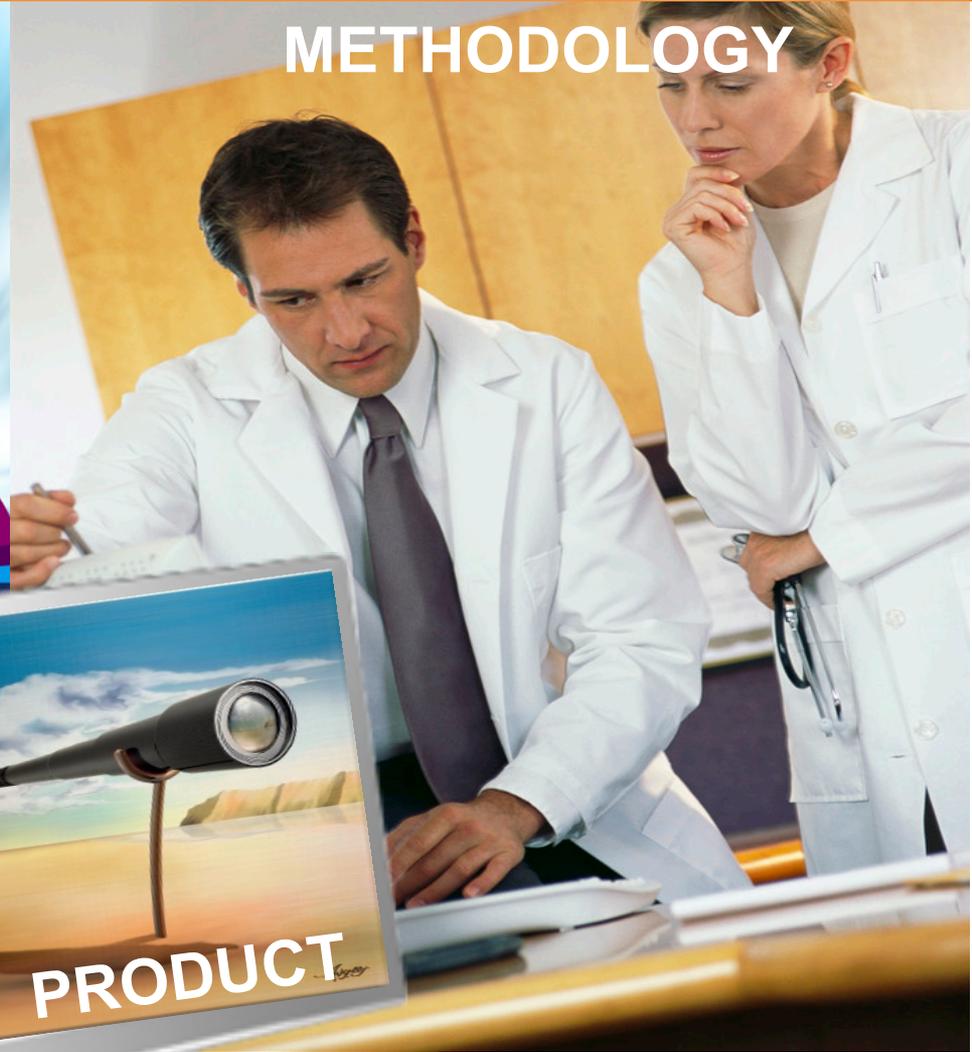
Migliorare la governance dei Rischi e  
Controlli della banca.



# AUGEOS SOLUTION

METHODOLOGY

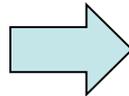
DATA



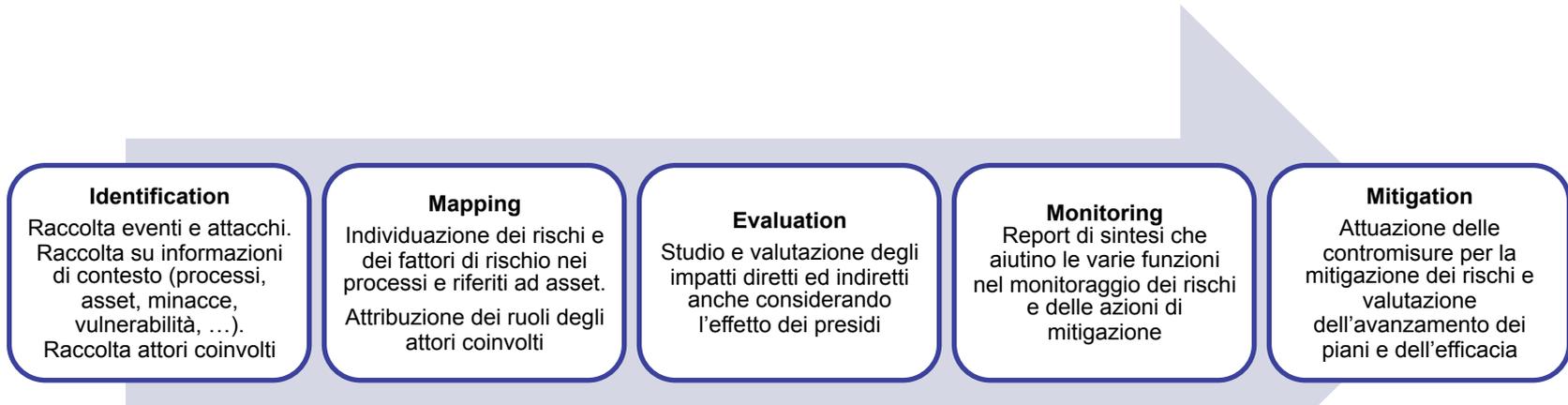
# Governo del Rischio – Metodologie di Analisi

## Analisi minacce di rischio informatico e correlazione con Rischio Operativo:

- Definizione tassonomia minacce
- Correlazione con Rischio Operativo



Approccio metodologico



### Maturity Model Approach

Profondità di analisi in funzione del livello richiesto

Risk Based Approach

Functional Risk Approach

Control in Action Approach



## Risk Executive Dashboard



RiskShelter



Normageos



AIT Risk

## GRC Plus



# Augeos

**Vi aspettiamo allo stand  
Augeos**



**Alcune innovazioni sono più incisive di altre**

