

Bando Unico R&S - Regione Toscana



La Piattaforma per l'Analisi dei Dati Basata sui Grafi

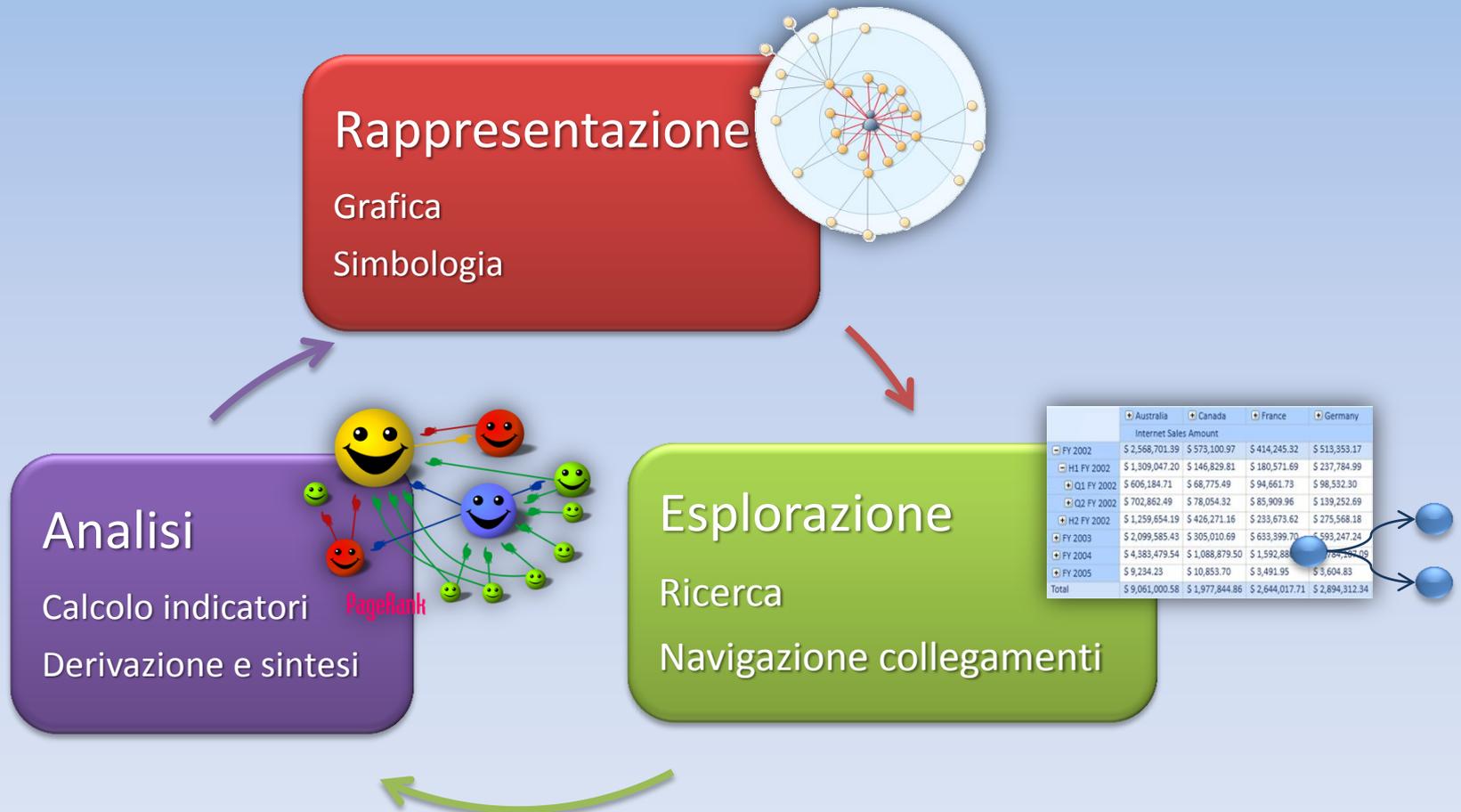
Ernesto Lastres
Sistemi Territoriali S.r.l.

Workshop 26/01/2012, Pisa

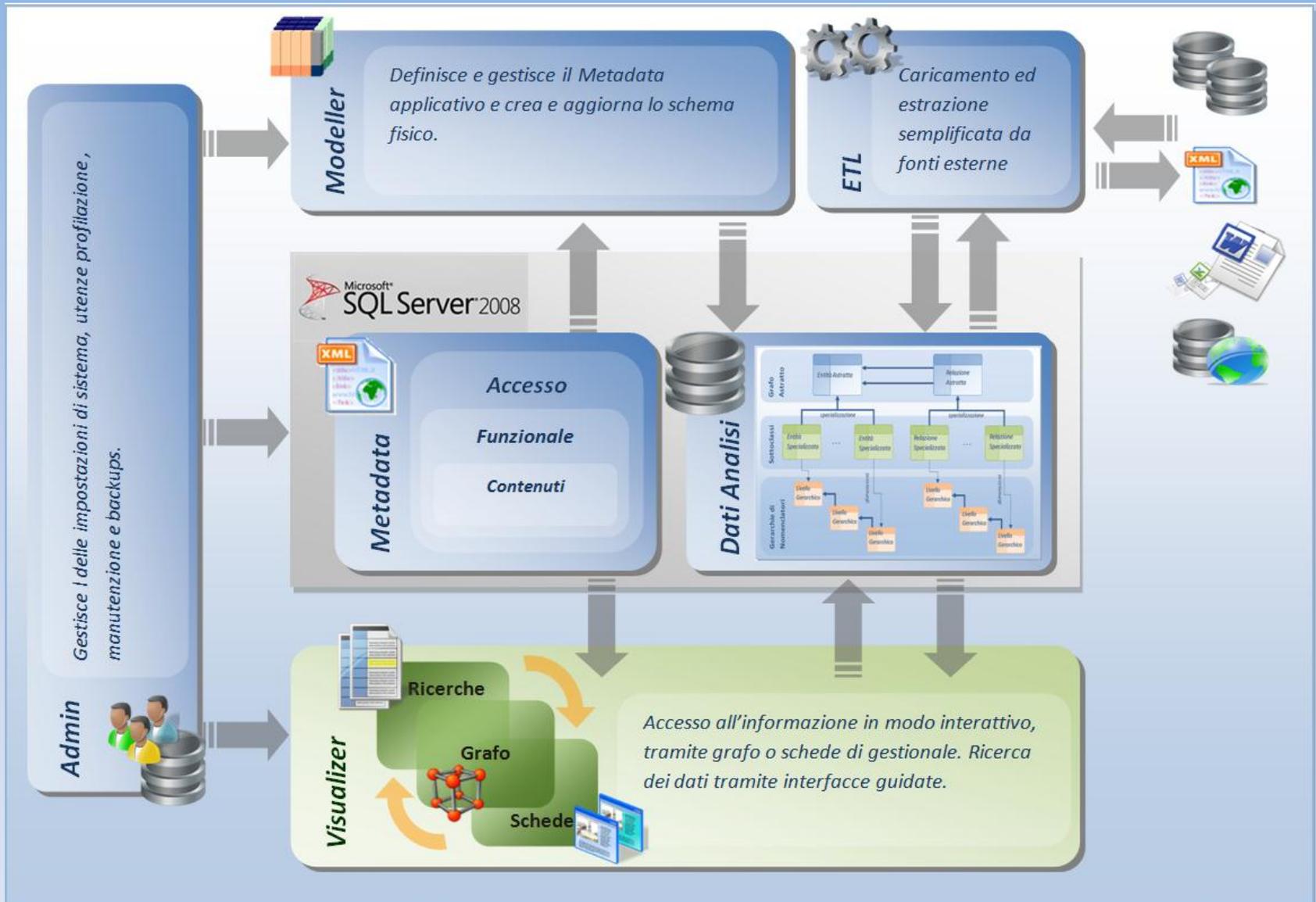
- Diverse contesti richiedevano sforzi di implementazione e riutilizzo.
- Ogni volta bisogna rifare e riprogettare.
- Campi applicativi:
 - Telefonia
 - Analisi patrimoniale
 - Frode assicurativa
 - Gioco online
 - Analisi clinica
- *Serve qualcosa che semplifichi il riuso e faciliti la personalizzazione in contesti applicativi diversi.*

- Disegno e progettazione di un framework di Business Intelligence (BI) per il supporto alle decisioni basato sulle Reti Sociali.
- Porre le basi per un prodotto informatico innovativo che sia ben integrato ed applicabile in diversi contesti.
- Fare ricerca su metodologie e procedure innovative di visualizzazione, navigazione ed analisi di valenza generale.
- Sperimentare l'uso del framework su dati clinici e sanitari.

- Nel fare ciò abbiamo voluto centrare tre aspetti fondamentali molto collegati.



Framework BI

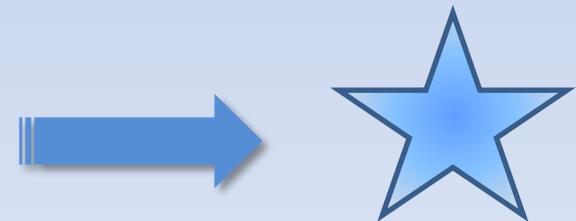


- Permette di definire il metadata che descrive l'applicazione.
- Principalmente tre tipi di metadata:
 - *Contenuti*: lo schema del database (il modello dati).
 - *Funzionale*: come i dati vengono fruiti.
 - *Accesso*: accesso a dati e funzione in base al profilo.
- Il metadata viene gestito in uno schema XML.
- Il modeler si configura come un sistema di Authoring indirizzato agli esperti di dominio.

- **Contenuto**
 - Tipi di entità e relazioni, attributi, descrizioni, alias, domini, gerarchie dimensionali, variabili di analisi, ecc.
- **Funzionale**
 - **Gestione alfanumerica**
 - Visibilità, sola lettura, ordine, categoria (raggruppamento), sezione (scheda), tipo di controllo UI, formato, dipendenze, *nullable*, domini, valore di *default*, espandibile, link, ecc.
 - **Rappresentazione grafica**
 - Colori (bordo e ripieno), figura geometrica, bitmap, trasparenze, caratteri, ecc.
 - **Ricerche e Analisi**
 - Tipi di algoritmi, mapping delle variabili, impostazione filtri, ecc.

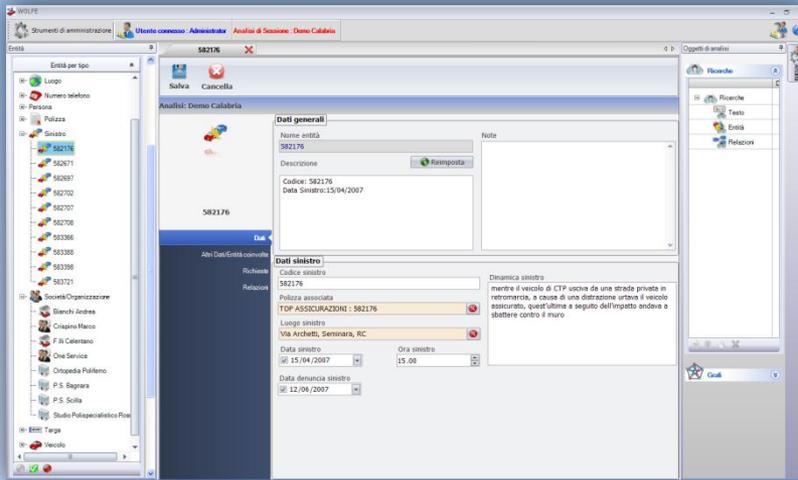
- Abbiamo cercato di dare molta importanza alla rappresentazione come chiave di lettura dei dati.
- La simbologia viene definita in modo dichiarativo.
- Il metadata definisce il modo in cui devono essere rappresentati i dati graficamente.
- I simboli possono caratterizzare (o essere mappati) su attributi degli elementi che rappresentano.

```
<Path x:Name="Path" Stretch="Fill" StrokeThickness="2" StrokeLineJoin="Round" Stroke="#FF118FBF"
  Data="F1 M 54.3333,93L 1,52.3333L 51.6667,1.00003L 113,31L 153.667,9.00003L 163,121.667L 54.3333,93 Z "
  <Path.Fill>
    <RadialGradientBrush RadiusX="0.506173" RadiusY="0.679558" Center="0.5,0.5" GradientOrigin="0.5,0.5">
      <RadialGradientBrush.GradientStops>
        <GradientStop Color="#FF1954A2" Offset="0"/>
        <GradientStop Color="#FF8CA9D0" Offset="0.503597"/>
        <GradientStop Color="#FFFFFF" Offset="1"/>
      </RadialGradientBrush.GradientStops>
    </RadialGradientBrush>
  </Path.Fill>
</Path>
```

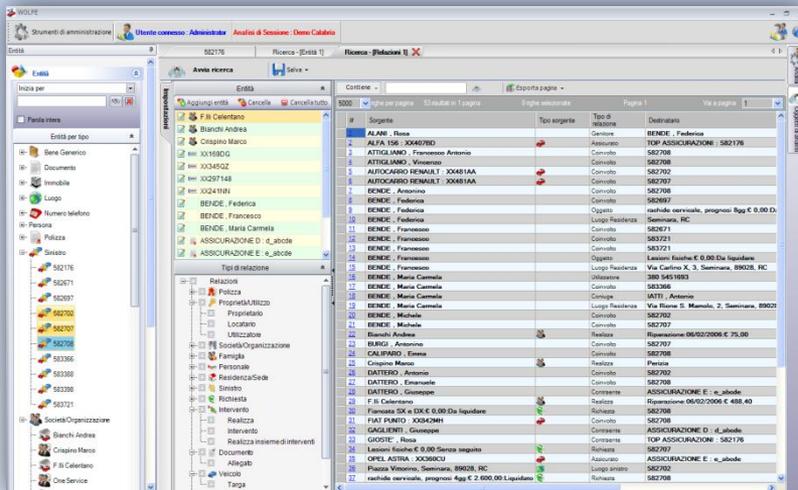


- Principale applicazione che va all'utente finale.
- Permette l'accesso all'informazione e funzionalità.
- Si configura dinamicamente in base al metadata applicativo.
- Costituito da tre moduli principali:
 - Anagrafica
 - Grafo
 - Strumenti di Analisi
- Premette di creare dei progetti che possono essere salvati e configurati

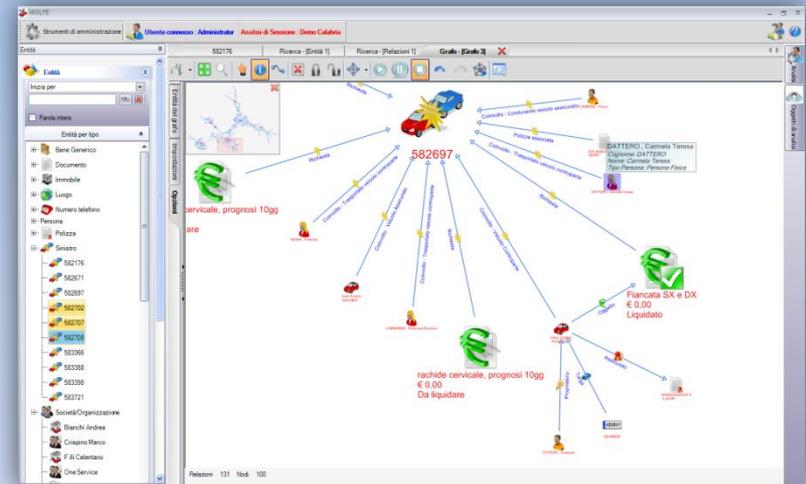
Anagrafica



Strumenti di Analisi



Grafo



- Schede di dettaglio per ogni entità e relazione.
- Viene generata dinamicamente dal visualizzatore.
- Sono auto-configurabili in base al metadata applicativo.
- Permettono di vedere e modificare gli attributi di entità e relazioni.
- Permette di navigare i collegamenti in modo «ipertestuale»
- Permette di comporre le relazioni in modo semplice collegando gli elementi col mouse.

- Permette di esplorare i collegamenti in modo grafico.
- Il grafo si costruisce dinamicamente.
- Si possono aggiungere elementi al grafo in tanti modi
 - Da ricerche
 - Espandendo un nodo
 - Da un altro grafo
- È possibile anche esplorare le proprietà degli elementi
- Si possono applicare dei filtri
- Calcolare cammini di collegamento
- Rappresentare i dati in diversi layout.
- Creare tematizzazioni (classificazioni) di colori e dimensioni sul valore delle variabili degli elementi.

MainWindow

Selected Analysis: Assititi

NodeSearch_1 | EdgeSearch_2 | Graph_1

Save | Close | Nodes: 59 | Edges: 75

Tematizzazione archi e nodi in su dimensione e colere in funzione delle variabili

Tools

Navigation - Info

Filters

Classification

Options

Node Color

On / Off Apply

Configuration

Use only loaded D:

Var: Node_Weight

From: [Blue]

To: [Red]

Class: 3 Infinite

Node Size

On / Off Apply

Configuration

Use only loaded D:

Var: Node_Weight

From: 50.00

To: 90.00

Class: 3 Infinite

Edge Color

On / Off Apply

Configuration

Use only loaded D:

Var: Edge_Weight

From: [Green]

To: [Yellow]

Class: 3 Infinite

Edge Size

On / Off Apply

Configuration

Use only loaded D:

Var: []

MainWindow

Selected Analysis: Assistiti

NodeSearch_1 EdgeSearch_2 Graph_1

Save Close Nodes: 59 Edges: 75

Navigation Info

Layout Type: **HIERARCH**

Separation

Commands

Tools

Analysis Objects

Animazioni e impostazioni di differenti layout.

Ricerche

- Permettere di realizzare delle ricerche che daranno come risultato diversi insiemi di dati in griglie e tabelle.
- Applicare delle selezioni, filtri, ordinamenti o attivare delle “schede” di tipo gestionale:
 - **Ricerca Entità**. Individua le entità in base a combinazioni di parametri associati agli attributi.
 - **Ricerca Relazioni**. Individua le relazioni in base a combinazioni di parametri associati agli attributi e a insiemi di entità.
- I risultati delle ricerche potranno a sua volta essere utilizzati per costruire altre ricerche.

Ricerche

MainWindow

Selected Analysis: Assistiti

NodeSearch_1 EdgeSearch_2

Execute Search Save Close Initial Entity Set Constrains: 0

50 items for page Page 1 of 19568 978353 elements in 19568 page(s)

Drag a column header here to group by that column.

Source Type	Source Label	Relation Type	Relation Label	Target Type	Target Label	Start Time	End Time	Rel_Dis
Ass	827804 (10)	assume (1)	assume (1)	Farm A02BA (2430)	12/24/2007	12/24/2007		
Ass	523265 (11)	assume (1)	assume (1)	Farm J01FA (27304)	4/27/2007	4/27/2007		
Ass	1857770 (5)	assume (1)	assume (1)	Farm H02AB (23200)	12/18/2007	12/18/2007		
Ass	919224 (14)	assume (1)	assume (1)	Farm A02BC (23338)	10/29/2007	10/29/2007		
Ass	2929161 (6)	assume (1)	assume (1)	Farm H02AB (23200)	7/2/2007	7/2/2007		
Ass	2877432 (5)	assume (1)	assume (1)	Farm G03CA (1836)	6/12/2007	6/12/2007		
Ass	1587950 (6)	assume (1)	assume (1)	Farm J01CR (31904)	11/24/2007	11/24/2007		
Ass	1065598 (11)	assume (1)	assume (1)	Farm C07AA (1248)	5/7/2007	5/7/2007		
Ass	3709251 (7)	assume (1)	assume (1)	Farm R03AK (6757)	10/12/2007	10/12/2007		
Ass	2909812 (3)	assume (1)	assume (1)	Farm J01FA (27304)	9/11/2007	9/11/2007		
Ass	460862 (7)	assume (1)	assume (1)	Farm J01MA (19450)	2/5/2007	2/5/2007		
Ass	1434478 (7)	assume (1)	assume (1)	Farm M01AX (13670)	8/6/2007	8/6/2007		
Ass	476992 (6)	assume (1)	assume (1)	Farm J01DA (22543)	11/8/2007	11/8/2007		
Ass	1970279 (4)	assume (1)	assume (1)	Farm J01CR (31904)	6/16/2007	6/16/2007		
Ass	581324 (13)	assume (1)	assume (1)	Farm D07AC (1049)	7/7/2007	7/7/2007		
Ass	1944673 (3)	assume (1)	assume (1)	Farm J01MA (19450)	2/14/2007	2/14/2007		
Ass	332850 (10)	assume (1)	assume (1)	Farm H02AB (23200)	5/22/2007	5/22/2007		
Ass	1050134 (30)	assume (1)	assume (1)	Farm B01AA (2881)	1/10/2007	1/10/2007		
Ass	1535483 (5)	assume (1)	assume (1)	Farm # (2588)	10/12/2007	10/12/2007		
Ass	3795325 (8)	assume (1)	assume (1)	Farm N02BE (7128)	4/18/2007	4/18/2007		
Ass	2618018 (16)	assume (1)	assume (1)	Farm J01MB (498)	7/17/2007	7/17/2007		
Ass	1298663 (6)	assume (1)	assume (1)	Farm M01AB (18906)	4/23/2007	4/23/2007		
Ass	327761 (3)	assume (1)	assume (1)	Farm G03DA (750)	7/18/2007	7/18/2007		

Search

- Search
- Entities
- Relations

Graph

Graph

Algorithm

- Algorithm
- Page Rank Centrality
- EigenVector Centrality
- Weighted Page Rank Ce
- Betweenness Centrality

- *Altre ricerche utili*

- ***Frequenza***: statistiche di partecipazione dei soggetti/entità nelle relazioni (quante volte ha partecipato e con quale ruolo).
- ***Incrocio***: Tutti i soggetti/entità in comune a più situazioni date (ricavandone i ruoli).
- ***Patterns***: Ricerca di schemi comuni secondo certi criteri.
- ***Composizione***: di ricerche con operazioni insiemistiche/algebriche.

Calcolo di indicatori

- *Meccanismo generale per calcolare e aggiungere nuove variabili di analisi su sottoinsiemi di dati.*
- *Le nuove variabili sono disponibili per l'analisi a livello di grafo (tematizzazione e filtri) e ricerche.*
- *Grafi e ricerche possono essere «inviare» agli algoritmi di calcolo.*
 - **Centralità:** *Calcolo di indici di centralità.*
 - *Degree*
 - *Betweenness*
 - *PageRank*
 - *Eigenvector*
 - **Gruppi densi:** *Calcolo di indici di addensamento e identificazione dei cluster.*

- Introdurre nuovi strumenti di ricerca che lo rendano più versatile.
- Migliorare ancora l'usabilità tramite la grafica e l'introduzione di nuovi controlli utente.
- Esperimentare l'applicazione in diversi settori applicativi.
- Industrializzare la soluzione.