

FantaTitolari



Gestione Avanzata Dei Dati

Giuseppe Cantisani

Che cos'è il fantacalcio?

- È un gioco che ogni anno coinvolge milioni di appassionati di calcio
- Ogni giocatore ha a disposizione 500 fantamilioni da spendere per costruire una squadra
- La rosa è composta tipicamente da 3 portieri, 8 difensori, 8 centrocampisti, 6 attaccanti
- L'esito di una partita si basa sulle reali prestazioni di 11 calciatori schierati su 25 disponibili
- Sono ammessi i seguenti moduli: 4-4-2, 4-3-3, 4-5-1, 5-3-2, 5-4-1, 3-4-3, 3-5-2
- Sono ammesse 3 sostituzioni dalla panchina
- I voti sono assegnati da una fonte sportiva scelta dalla Lega
- Obiettivo: arrivare 1° in classifica

Specifica del problema - 1

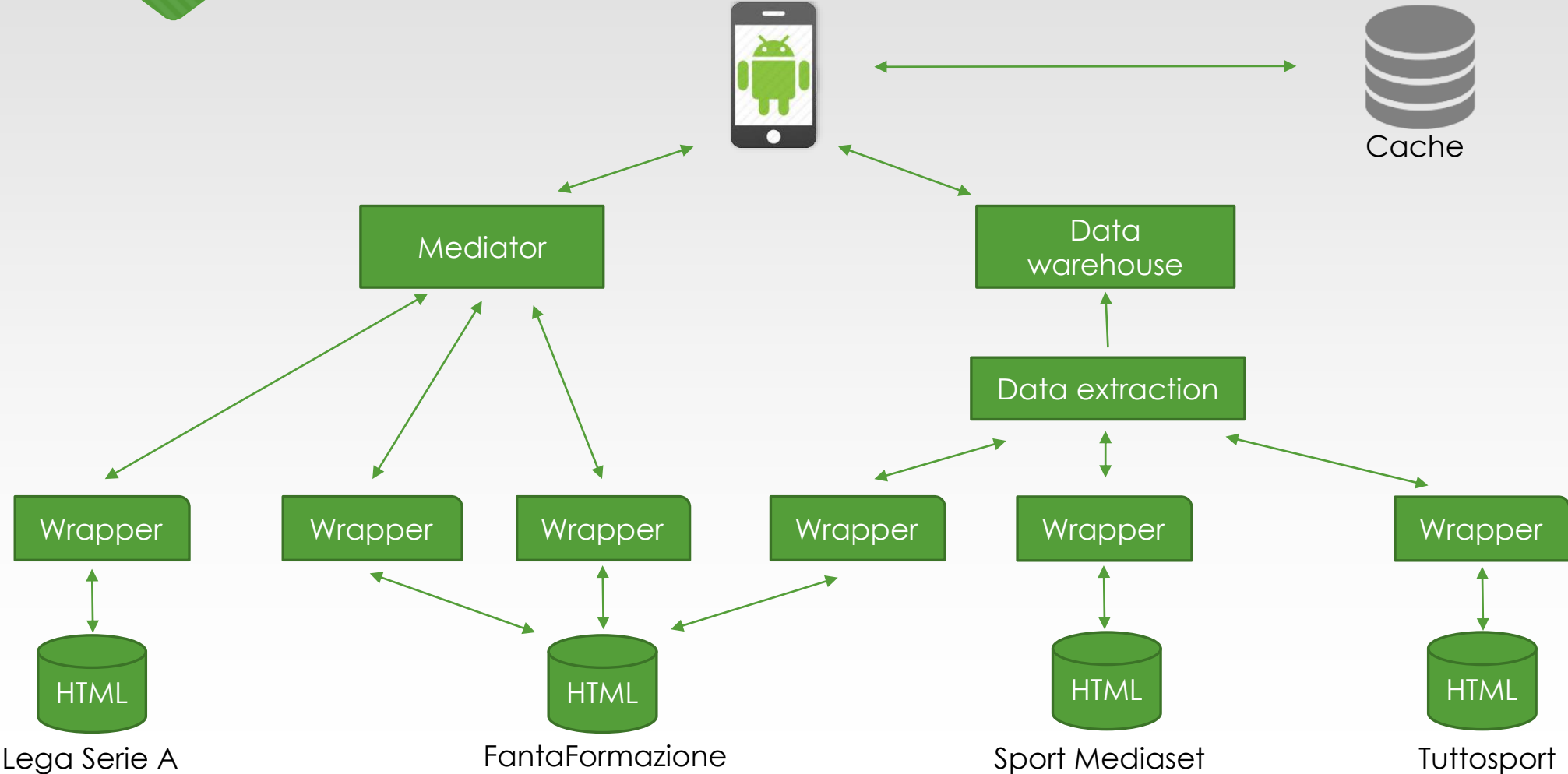
- Quale modulo devo adottare? Chi devo schierare? Quale fonte devo consultare? Quali sono le fonti affidabili?
- Fantagazzetta, Gazzetta dello Sport, Corriere dello Sport, Tuttosport, Sky, Sport Mediaset sono le fonti più importanti per la consultazione delle probabili formazioni.
- È buona norma consultare più di una fonte prima di decidere chi schierare
- Più informazioni = maggiore probabilità di ottenere informazioni corrette sul giocatore
- Problema: Consultare 6 siti per ogni giocatore $6 \cdot 25 = 150$ visualizzazioni delle stesse pagine per decidere la formazione -> Operazione molto laboriosa!

Specifica del problema - 2

FantaTitolari risolve il problema e fa molto di più!

- Permette l'inserimento di una o più rose senza alcuna registrazione al sistema
- Integra in un'unica interfaccia l'informazione (titolare, riserva, indisponibile) di tutti i giocatori della rosa per tutte le fonti citate
- Mostra la classifica della Serie A
- Mostra il calendario della giornata da disputare
- Mostra la classifica dei marcatori e assist man
- Mostra le statistiche dei migliori e peggiori della Serie A
- Mostra la lista dei giocatori indisponibili con tipo di infortunio e rientro previsto

Architettura



Cache

- Il risultato di ogni query viene salvato in un DB SQLite usato come cache
- Ad ogni query è associato un tempo di scadenza:
 - Disponibilità di ogni giocatore in rosa : periodici (dopo gli allenamenti)
 - Classifica Serie A , calendario, statistiche giocatori: variabile (Dopo l' ultima partita di giornata)
 - Indisponibili: 1 volta al giorno (a fine giornata dopo gli allenamenti)

Fonti - 1

- FantaFormazione (<http://www.fantaformazione.com>)
 - È un sito web per la gestione del fantacalcio
 - Utilizzato per ottenere la lista completa dei calciatori di Serie A con relative statistiche, l'informazione (titolare, riserva, indisponibile) di tutti i giocatori per le fonti sportive: Gazzetta dello Sport , Corriere dello sport, FantaGazzetta e Sky, la lista degli indisponibili, il calendario di giornata
 - Web scraping – cambi periodici



Fonti - 2

- Tuttosport (<http://www.tuttosport.com>)
 - È un quotidiano sportivo con sede a Torino, è il terzo giornale sportivo più venduto in Italia
 - Utilizzato per ottenere l'informazione (titolare, riserva, indisponibile) di tutti i giocatori
 - Web scraping – cambi periodici
- Sport Mediaset (<http://www.sportmediaset.mediaset.it>)
 - È una rubrica sportiva realizzato e curata da Mediaset.
 - Utilizzato per ottenere l'informazione (titolare, riserva, indisponibile) di tutti i giocatori
 - Web scraping – cambi periodici

The logo for Tuttosport, featuring the word "TUTTOSPORT" in a bold, brown, sans-serif font with a slight shadow effect.The logo for Sport Mediaset, featuring the words "SPORT" and "MEDIASET" in white, bold, sans-serif font, each enclosed in a blue rectangular box with a white border.

Fonti - 3

- Lega Serie A (<http://www.legaseriea.it>)
 - È l'organo che gestisce i più importanti tornei calcistici per club in Italia
 - Utilizzato per ottenere la classifica della Serie A
 - Web scraping – cambi periodici



Collegamento fonti (join)

- La lista completa dei giocatori viene scaricata al primo avvio dell'app da Fantaformazione
- Uso il nome del giocatore per estrarre l'informazione (titolare, riserva, indisponibile) per le fonti Gazzetta dell Sport, Corriere dello Sport, Sky, FantaGazzetta da FantaFormazione
- Uso lo stesso nome per estrarre l'informazione (titolare, riserva, indisponibile) da Tuttosport e Sport Mediaset

Descrizione schemi locali - 1

○ **FantaFormazione (FF)**

- **lista** (Giocatore, Ruolo, Squadra, ImgSq, MV, Giocate, NAss, NGol, Fg, Gds, Sky, Cds, Data)
- **calendario** (Squadra1, Img1, Squadra2, Img2, DataInc)
- **indisponibili** (Giocatore, TipoInf, DataR)

○ **SportMediaset (SM)**

- **lista** (Giocatore, Disp)

Descrizione schemi locali - 2

- **Tuttosport (TS)**

- **lista** (Giocatore, Disp)

- **Lega Serie A (LSA)**

- **classifica** (Pos, Squadra, Img, Punti, G, V, N, P, F, S)

Wrapper - 1

- I wrapper sono stati realizzati con Import.io
- FantaFormazione:
`http://www.fantaformazione.com/fantacalcio.calendari/probabili/Home.htm? (calendario) - >Tw(squadra1, img1, squadra2, img2, dataInc), Ew(visivo)`

Genera la seguenti API: https://api.import.io/store/connector/ea28a68d-0fae-467b-96b0-76b3dd618c90/_query?input=webpage/url:http%3A%2F%2Fwww.fantaformazione.com%2Ffantacalcio.calendari%2Fprobabili%2FHome.htm%3F&&_apikey=f92e2a3c256340f7a6015c435441c5f3426c424c5d814eac393c8882bc3a3c3b757082b31da8d7d5102f60bf0e987a8055b447cfd44a01fb5225df00a4eb761c76f6926b2d7b456ca342c775ac91d80

risultato: `{"offset":0,"results":[{"squadra1":"Bologna","incontro":"sabato 6 febbraio alle 18:00","squadra2":"Fiorentina","img2":"http://www.fantaformazione.com/modules/app/fantacalcio/views/imgs/fiorentina.png","img1":"http://www.fantaformazione.com/modules/app/fantacalcio/views/imgs/bologna.png"},...]}`

- Le tabelle **lista** e **indisponibili** sono costruite allo stesso modo

Wrapper - 2

- SportMediaset: http://www.sportmediaset.mediaset.it/calcio/calcio/serie-a-le-formazioni_1083183-201502a.shtml ->Tw(lista), Ew(visivo)

Genera la seguente API: https://api.import.io/store/connector/acf4c7b9-d278-4a98-b7ee-5c449651b7d7/_query?input=webpage/url:http%3A%2F%2Fwww.sportmediaset.mediaset.it%2Fcalcio%2Fcalcio%2Fserie-a-le-formazioni_1083183-201502a.shtml&&apikey=f92e2a3c256340f7a6015c435441c5f3426c424c5d814eac393c8882bc3a3c3b757082b31da8d7d5102f60bf0e987a8055b447cffd44a01fb5225df00a4eb761c76f6926b2d7b456ca342c775ac91d80

risultato: {"offset":0,"results":[{"lista":"Carpi (4-4-1-1): Belec; Zaccardo, S. Romagnoli, Gagliolo, Letizia; Pasciuti, Cofie, Lollo, Martinho; Lasagna; Borriello. A disp.: Brkic, Gabriel Silva, Spolli, Marrone, Lazzari, Di Gaudio, Mbakogu, Matos. All.: Castori Squalificati: - Indisponibili: Fedele, Bubnjic, Bianco Milan (4-4-2): Donnarumma; Abate, Alex, A. Romagnoli, De Sciglio; Cerci, Kucka, Montolivo, Bonaventura; Bacca, Niang. A disp.: Abbiati, Livieri, Calabria, Zapata, Mexes, Poli, José Mauri, De Jong, Honda, Suso, Luiz Adriano. All.: Mihajlovic Squalificati: - Indisponibili: Menez, Balotelli, Diego Lopez, Bertolacci, Antonelli"},...]}

- Per poter ottenere **lista** (giocatore, disp) ho usato uno script php che invoca l'API e per ogni squadra estrae i titolari (A disp.), riserve (All) e gli indisponibili (Squalificati, Indisponibili)

Wrapper - 3

- Tuttosport: : [http://www.tuttosport.com/calcio/serie-a/probabili-formazioni/2015/12/01-6371070/inter_-_genoa/->Tw\(lista\), Ew\(visivo\)](http://www.tuttosport.com/calcio/serie-a/probabili-formazioni/2015/12/01-6371070/inter_-_genoa/->Tw(lista), Ew(visivo))

Genera la seguente API: https://api.import.io/store/connector/a33b7282-14a7-4d1b-a29a-335d3ed97a59/?query=input=webpage/url:http%3A%2F%2Fwww.tuttosport.com%2Fformazione%2Fcalcio%2Fserie-a%2Fprobabili-formazioni%2F2015%2F12%2F01-6371070%2Finter_-_genoa%2F&&apikey=f92e2a3c256340f7a6015c435441c5f3426c424c5d814eac393c8882bc3a3c3b757082b31da8d7d5102f60bf0e987a8055b447cffd44a01fb5225df00a4eb761c76f6926b2d7b456ca342c775ac91d80

```
risultato:{"offset":0,"results":[{"lista":"Inter"}, {"lista":"Genoa"}, {"lista":"Titolari"}, {"lista":["Handanovic","1","1","Perin"]}, {"lista":["D'Ambrosio","33","11","Figueiras"]}, {"lista":["Alex Telles","12","93","Laxalt"]}, {"lista":["Medel","17","3","Ansaldi"]}, {"lista":["Miranda","25","8","Burdisso"]}, {"lista":["Murillo","24","5","Izzo"]}, {"lista":["Brozovic","77","88","Rincón"]}, {"lista":["Melo","83","20","Tino Costa"]}, {"lista":["Icardi","9","13","Gakpe"]}, {"lista":["Ljajic","22","22","Lazovic"]}, {"lista":["Palacio","8","10","Perotti"]}, {"lista":"Riserve"}, {"lista":["Carrizo","30","23","Lamanna"]}, {"lista":["Ranocchia","23","27","Ujkani"]}, {"lista":["Montoya","14","4","De Maio"]}, {"lista":["Juan Jesus","5","16","Capel"]}, {"lista":["Jovetic","10","21","Pandev"]}, {"lista":["Guarín","13","77","Tachtsidis"]}, {"lista":["Manaj","97","18","OlivierN tcham"]}]}
```

- La tabella lista restituisce una lista di informazioni per singola partita (nell'esempio Inter – Genoa), per ottenere **lista** (giocatore, disp) ho usato uno script php che ripulisce i dati estraendo i titolari (dopo l'oggetto che contiene il valore Titolari), riserve (dopo l'oggetto che contiene il valore Riserve), i restanti giocatori non presenti nella lista sono indisponibili.

Wrapper - 4

- Lega Serie A: <http://www.legaseriea.it/it/serie-a-tim/classifica> (classifica) -> Tw(pos, squadra, img, punti, g, v, n, p, f, s)
-), Ew(visivo)

Genera la seguente API:https://api.import.io/store/connector/71bc2228-dcdf-48b1-b992-b6397f56f8c4/_query?input=webpage/url:http%3A%2F%2Fwww.legaseriea.it%2Fit%2Fserie-a-tim%2Fclassifica&&apikey=f92e2a3c256340f7a6015c435441c5f3426c424c5d814eac393c8882bc3a3c3b757082b31da8d7d5102f60bf0e987a8055b447cffd44a01fb5225df00a4eb761c76f6926b2d7b456ca342c775ac91d80

```
risultato:{"offset":0,"results":[{"img":"http://www.legaseriea.it/uploads/default/attachments/squadre/squadre_m/12/images/logo_piccolo/157/original/napoli.png","subiti":"19","giocate":"23","fatti":"52","img/_title":"NAPOLI","pos":"1","perse":"2","squadra":"1 Napoli","pareg":"5","img/_source":"/uploads/default/attachments/squadre/squadre_m/12/images/logo_piccolo/157/original/napoli.png","vinte":"16","punti":"53"},... ]}
```


Schema globale

- **statistiche** (Giocatore, Ruolo, Squadra, ImgSq, MV, Giocate, NAss, NGol)
- **info_rose** (Giocatore, Fg, Gds, Sky, Cds, Sm, Ts, Data)
- **indisponibili** (Giocatore, Tipolnf, DataR)
- **classifica** (Pos, Squadra, Img, Punti, G, V, N, P, F, S)
- **calendario** (Sq1, Img1, Sq2, Img2, DataInc)
- **incontro** (Giocatore, ImgSq1, ImgSq2)

Mapping GAV - 1

- **statistiche** (Giocatore, Ruolo, Squadra, ImgSq, MV, Giocate, NAss, NGol) :- **FF.lista** (Giocatore, Ruolo, Squadra, ImgSq, MV, Giocate, NAss, NGol, L1, L2, L3, L4, L5)
- **info_rose** (Giocatore, Fg, Gds, Sky, Cds, Sm, Ts, Data) :- **FF.lista** (Giocatore, L1, L2, L3, L4, L5, L6, L7, Fg, Gds, Sky, Cds, Data), **SM.lista** (Giocatore, Sm), **TS.lista**(Giocatore, Ts)
- **indisponibili** (Giocatore, Tipolnf, DataR) :- **FF.indisponibili** (Giocatore, Tipolnf, DataR)

Mapping GAV - 2

- **classifica** (Pos, Squadra, Img, Punti, G, V, N, P, F, S) :- **LSA.classifica** (Pos, Squadra, Img, Punti, G, V, N, P, F, S)
- **calendario** (Sq1, Img1, Sq2, Img2, DataInc) :- **FF.calendario** (Sq1, Img1, Sq2, Img2, DataInc)
- **incontro** (Giocatore, ImgSq1, ImgSq2) :- **FF.lista** (Giocatore, L1, Squadra, L2, L3, L4, L5, L6, L7, L8, L9, L10, L11), **FF.calendario** (Squadra, ImgSq1, L11, ImgSq2, L12)
- **incontro** (Giocatore, ImgSq1, ImgSq2) :- **FF.lista** (Giocatore, L1, Squadra, L2, L3, L4, L5, L6, L7, L8, L9, L10, L11), **FF.calendario** (Sq1, ImgSq1, Squadra, ImgSq2, L12)

Mapping LAV - 1

- **FF.lista** (Giocatore, Ruolo, Squadra, ImgSq, MV, Giocate, NAss, NGol, Fg, Gds, Sky, Cds, Data) :- **statistiche** (Giocatore, Ruolo, Squadra, ImgSq, MV, Giocate, NAss, NGol), **info_rose** (Giocatore, Fg, Gds, Sky, Cds, L1, L2, Data)
- **FF.indisponibili** (Giocatore, Tipolnf, DataR) :- **indisponibili** (Giocatore, Tipolnf, DataR)
- **SM.lista** (Giocatore, Disp) :- **info_rose** (Giocatore, L1, L2, L3, L4, Disp, L5, L6)

Mapping LAV - 2

- **TS.lista** (Giocatore, Disp) :- **info_rose** (Giocatore, L1, L2, L3, L4, L5, Disp, L6)
- **FF.calendario** (Sq1, Img1, Sq2, Img2, DataInc) :- **calendario** (Sq1, Img1, Sq2, Img2, DataInc)
- **LSA.classifica** (Pos, Squadra, Img, Punti, G, V, N, P, F, S) :- **classifica** (Pos, Squadra, Img, Punti, G, V, N, P, F, S)

Query - 1

- Ricerca dei giocatori per nome
- findPlayerByName (Giocatore, Ruolo, Squadra) :- **statistiche** (Giocatore, Ruolo, Squadra, L1, L2, L3, L4, L5)
- ```
SELECT statistiche.Giocatore, statistiche.Ruolo, statistiche.Squadra
FROM statistiche
WHERE statistiche.Giocatore MATCH 'AL'
```

# Query - 2

- Ricerca delle informazioni per un giocatore della rosa
- FindInfoByName(Giocatore, Sqlmg1, Sqlmg2, Fg, Gds, Sky, Cds, Sm, Ts, Data) :- **info\_rose** (Giocatore, Fg, Gds, Sky, Cds, Sm, Ts, Data), **statistiche** (Giocatore, L1, Squadra, L2, L3, L4, L5, L6), **incontro** (Giocatore, Sqlmg1, Sqlmg2)
- SELECT incontro.ImgSq1 as Sqlmg1, incontro.ImgSq2 as Sqlmg2, info\_rose.Fg, info\_rose.Gds, info\_rose.Sky, info\_rose.Cds, info\_rose.Sm, info\_rose.Ts  
FROM info\_rose, statistiche, incontro  
WHERE info\_rose.Giocatore = statistiche.Giocatore AND statistiche.Giocatore = incontro.Giocatore AND Giocatore='ALEX SANDRO'

# Query - 3

- Estrae la classifica
- GetClassifica(Pos, Squadra, Img, Punti, G, V, N, P, F, S) :- **classifica** (Pos, Squadra, Img, Punti, G, V, N, P, F, S)
- SELECT \*  
FROM classifica  
WHERE 1



# Query - 4

- Estrae gli indisponibili
- GetInd(Giocatore, Ruolo, Squadra, TipoInf, DataR) :- **indisponibili** (Giocatore, TipoInf, DataR), **statistiche** (Giocatore, Ruolo, Squadra, L1, L2, L3, L4, L5)
- SELECT indisponibili.Giocatore, statistiche.Ruolo, statistiche.Squadra, indisponibili.TipoInf, indisponibili.DataR  
FROM indisponibili, statistiche  
WHERE indisponibili.Giocatore = statistiche.Giocatore

# Query - 5

- Estrae giocatori con voti più alti per ogni ruolo
- GetTop(Giocatore, Ruolo, MV) :- **statistiche** (Giocatore, Ruolo, L1, L2, MV, Giocate, L3, L4)
- ```
SELECT statistiche.Giocatore, statistiche.MV
FROM statistiche
WHERE Ruolo LIKE='D%'
ORDER BY MV DESC, Giocate DESC LIMIT 6
```

Query - 6

- Estrae giocatori con voti più bassi per ogni ruolo
- GetFlop(Giocatore, Ruolo, MV) :- **statistiche** (Giocatore, Ruolo, L1, L2, MV, Giocate, L3, L4)
- ```
SELECT statistiche.Giocatore, statistiche.MV
FROM statistiche
WHERE Ruolo LIKE='D%' AND MV != 0.00
ORDER BY MV ASC, Giocate DESC LIMIT 6
```

# Query - 7

- Estrarre la classifica dei marcatori della Serie A
- `getMarcatori(Giocatore, ImgSq, NGol) :- statistiche (Giocatore, L1, L2, ImgSq, L3, L4, L5, NGol)`
- ```
SELECT statistiche.Giocatore, statistiche.ImgSq, statistiche.NGol  
FROM statistiche  
ORDER BY NGol DESC LIMIT 20
```

Query - 8

- Estrarre la classifica degli assist man della Serie A
- getAssistMen(Giocatore, ImgSq, NAssist) :- **statistiche** (Giocatore, L1, L2, ImgSq, L3, L4, NAssist, L6)
- ```
SELECT statistiche.Giocatore, statistiche.ImgSq, statistiche.NAssist
FROM statistiche
ORDER BY NAssist DESC LIMIT 20
```

# Query - 9

- Estrae il calendario della giornata corrente
- GetCalendario(Sq1, Img1, Sq2, Img2, DataInc) :- **calendario** (Sq1, Img1, Sq2, Img2, DataInc)
- SELECT \*  
FROM calendario  
WHERE 1

# GAV query unfolding - 1

- GetInd(Giocatore, Ruolo, Squadra, Tipolnf, DataR) :- **indisponibili** (Giocatore, Tipolnf, DataR), **statistiche** (Giocatore, Ruolo, Squadra, L1, L2, L3,L4,L5)
- GetInd(Giocatore, Ruolo, Squadra, Tipolnf, DataR) :- **FF.indisponibili (Giocatore, Tipolnf, DataR)**, statistiche (Giocatore, Ruolo, Squadra, L1, L2, L3, L4, L5)
- GetInd(Giocatore, Ruolo, Squadra, Tipolnf, DataR) :- FF.indisponibili (Giocatore, Tipolnf, DataR), **FF.lista (Giocatore, Ruolo, Squadra, L1, L2, L3, L4, L5, L6, L7, L8, L9, L10)**

# GAV query unfolding - 2

- GetInd(Giocatore, Ruolo, Squadra, TipoInf, DataR) :- FF.indisponibili (Giocatore, TipoInf, DataR), FF.lista (Giocatore, Ruolo, Squadra, L1, L2, L3, L4, L5, L6, L7, L8, L9, L10)
- Non sono possibili ulteriori semplificazioni



# LAV bucket algorithm - 1

- GetInd(Giocatore, Ruolo, Squadra, Tipolnf, DataR) :- **indisponibili** (Giocatore, Tipolnf, DataR), **statistiche** (Giocatore, Ruolo, Squadra, L1, L2, L3, L4, L5)
- Bucket (**indisponibili** (Giocatore, Tipolnf, DataR)):
  - **FF.indisponibili** (Giocatore, Tipolnf, DataR)
- Bucket (**statistiche** (Giocatore, Ruolo, Squadra, L1, L2, L3, L4, L5)):
  - **FF.lista** (Giocatore, Ruolo, Squadra, L1, L2, L3, L4, L5, L6, L7, L8, L9, L10)

# LAV bucket algorithm - 2

- Una riformulazione possibile:
- $q'$ (Giocatore, Ruolo, Squadra, Tipolnf, DataR) :- FF.indisponibili (Giocatore, Tipolnf, DataR), FF.lista (Giocatore, Ruolo, Squadra, L1, L2, L3, L4, L5, L6, L7, L8, L9, L10)
- $q'$ (Giocatore, Ruolo, Squadra, Tipolnf, DataR) :- **indisponibili (Giocatore, Tipolnf, DataR)**, FF.lista (Giocatore, Ruolo, Squadra, L2, L3, L4, L5, L6, L7, L8, L9, L10, L11)
- $q'$ (Giocatore, Ruolo, Squadra, Tipolnf, DataR) :- indisponibili (Giocatore, Tipolnf, DataR), **statistiche (Giocatore, Ruolo, Squadra, L1, L2, L3, L4, L5), info\_rose (Giocatore, R1, R2, R3, R4, R5, R5, R6)**

# LAV bucket algorithm - 3

- $q'$  (Giocatore, Ruolo, Squadra, Tipolnf, DataR) :- indisponibili (Giocatore, Tipolnf, DataR), statistiche (Giocatore, Ruolo, Squadra, L1, L2, L3, L4, L5), info\_rose (Giocatore, R1, R2, R3, R4, R5, R6, R7)
- $q' \subseteq \text{GetInd}$

# Tecnologie utilizzate

- Android
- Import.io
- PHP 5.6
- phpMyAdmin
- JSON

**FINE**