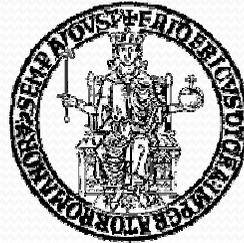


Sviluppo di strumenti grafici su una Web Application per il data mining

Università degli studi di Napoli Federico II



FACOLTÀ DI SCIENZE
MATEMATICHE, FISICHE E NATURALI
Corso di Laurea in Informatica

Relatore: Prof. Marco Faella
Tutor Aziendale: Dr. Massimo Brescia

Candidato: Giovanni Albano

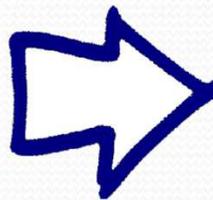
Aumento dei dati

- tabelle
- immagini
- grafici osservati o simulati
- risultato di analisi statistiche
- acquisiti da sensori di vario genere

I dati vanno immagazzinati e gestiti in qualche modo.



Conversione dei dati astrofisici in dati tabulari



	u-g	g-r	r-i	i-z	zspec
1	1,73213	0,95467	0,36712	0,30483	0,11008
2	1,47817	0,81329	0,40897	0,33138	0,05748
3	1,70356	0,90143	0,37031	0,2971	0,09349
4	2,05724	0,98329	0,38785	0,28708	0,08381
5	1,74603	0,82164	0,32789	0,27384	0,03039
6	1,87328	0,94494	0,36475	0,3271	0,0879
7	1,37929	0,88109	0,3761	0,27761	0,13275
8	1,05724	1,78063	0,5226	0,38796	0,36302
9	1,50683	0,88003	0,39445	0,26707	0,08517
10	1,78495	0,89739	0,35032	0,32353	0,07123
11	1,67323	0,85386	0,35259	0,24305	0,0851
12	1,9998	1,00836	0,40516	0,29552	0,08386
13	1,92678	1,15771	0,41964	0,3187	0,15685
14	2,16333	1,16546	0,46418	0,23726	0,17667
15	1,75237	0,83331	0,32421	0,26271	0,08203
16	1,17827	0,57508	0,18619	0,12324	0,02148
17	1,26829	0,59448	0,28959	0,13637	0,0519
18	1,79228	0,93059	0,39822	0,32444	0,11528
19	1,73924	0,88134	0,3295	0,27082	0,08612
20	1,70841	0,88142	0,3157	0,26633	0,07024
21	1,54163	1,06827	0,43559	0,24842	0,22961
22	1,92421	1,22972	0,38743	0,36923	0,15251
23	1,8015	0,98278	0,40315	0,34724	0,06885
24	1,92259	0,93589	0,38152	0,30983	0,07007
25	1,19568	0,56466	0,29216	0,16747	0,06884
26	1,91749	0,93738	0,38139	0,31645	0,08174
27	1,83843	0,94284	0,37141	0,30302	0,08258
28	1,7154	0,85239	0,37984	0,2969	0,08284
29	1,83776	0,91884	0,38131	0,27955	0,08303
30	1,81268	0,90695	0,35757	0,28843	0,08298
31	1,60364	0,8533	0,32156	0,26893	0,08353

Necessità di convertire in formato tabulare i cataloghi di immagini astronomiche -> grossa quantità di dati generati da gestire.

*Quarto paradigma della scienza

Progetto DAME (Data Mining & Exploration)



DAME Application - User: jobfmann@gmail.com LogOut

App Manuals | Model Manuals | Cloud Services | Science Cases | Documents | Info

RESOURCE MANAGER

Workspace

New Workspace

Rename | Workspace | Upload | Experiment | Delete

ppstest

File Manager

Workspace: ppstest

File	Type	Last Access	Delete
InputDataset.csv	csv	2012-05-21	✗
PPS_Train_sphere_coordinates.csv	csv	2012-09-05	✗

My Experiments

Workspace: ppstest

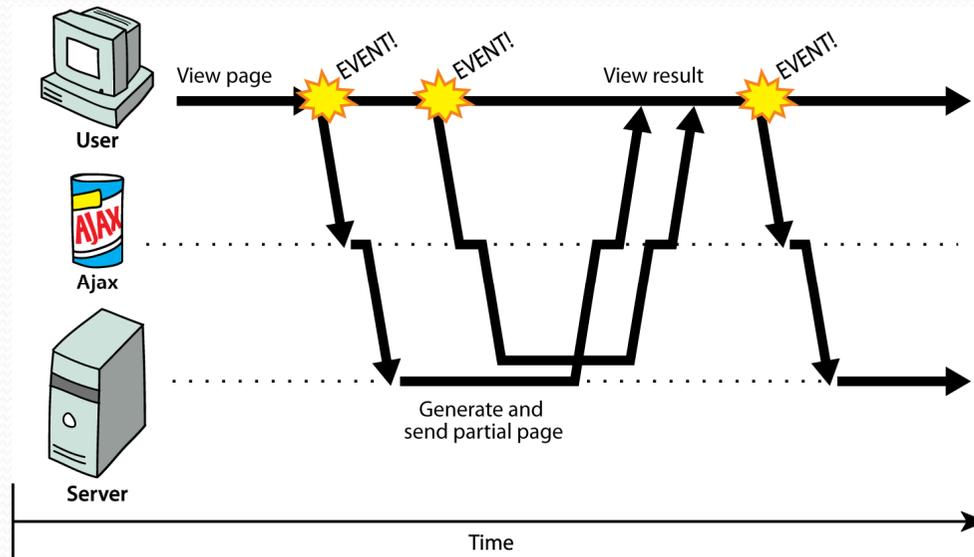
Experiment	Status	Last Access	Delete
▶ eee	ended	2012-05-21	✗
▶ prove	ended	2012-05-22	✗
▶ jfaklsfjkl	failed	2012-05-22	✗



Google Web Toolkit (GWT)

Caratteristiche:

- Compilatore Java-to-Javascript
- Ricche librerie grafiche
- Facilità di sviluppo
- Strumenti per la comunicazione asincrona



Integrazione strumenti grafici

VISUALIZZAZIONE IMMAGINI

DAME Web Application (localhost:8084/FEDev/)

DAME Application - User: jobfmann@gmail.com

App Manuals Model Manuals Cloud Services Science Cases Documents Info

RESOURCE MANAGER Histogram Scatter Plot 2D Scatter Plot 3D Line Plot Image View

Image View

Graphic Options

Workspace: alfonso

Image: sky.fits

Load Image Save Image as

CREAZIONE GRAFICI

DAME Web Application (localhost:8084/FEDev/)

DAME Application - User: jobfmann@gmail.com

App Manuals Model Manuals Cloud Services Science Cases Documents Info

RESOURCE MANAGER Histogram Scatter Plot 2D Scatter Plot 3D Line Plot

Graphic View

Histogram

Y

X

Graphic Options

Add Tab

Bin Placement

Width: 0.04

Title: Histogram

XLabel: X

YLabel: Y

Grid

Main

Workspace: WS20111221

Table: dataset_test_20.ascii

xAxis: g-f

Bar_Style: Filled

Color: Blue

Line_Width: 1

Flip

Clear tab

Plot Save Plot as Export in another window

Visualizzazione di immagini

Formati ammessi: fits-image, gif, png, jpeg

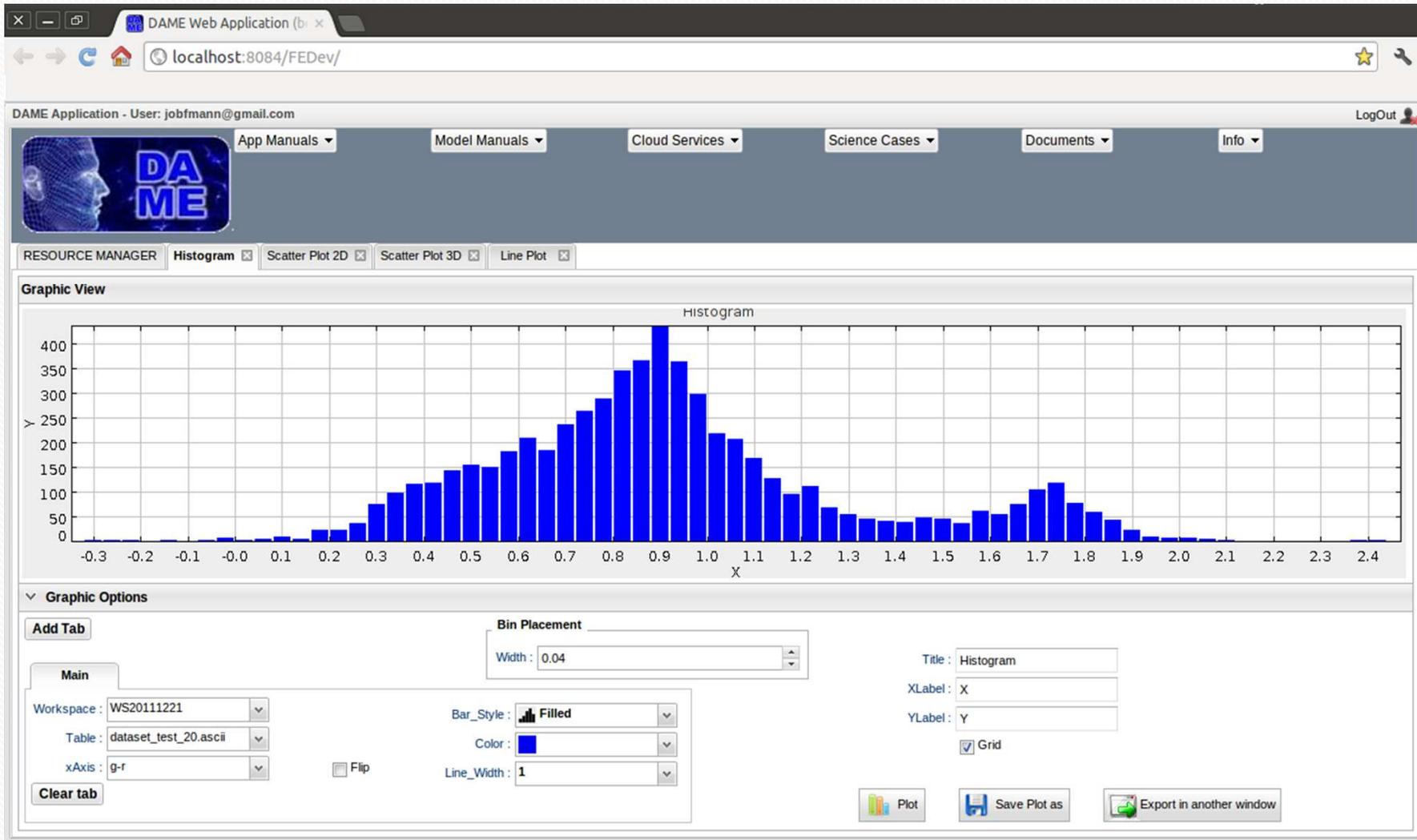
The screenshot shows a web browser window titled "DAME Web Application (b) x" with the address bar displaying "localhost:8084/FEDev/". The application header includes the user "jobfmann@gmail.com" and a "LogOut" link. A navigation menu contains "App Manuals", "Model Manuals", "Cloud Services", "Science Cases", "Documents", and "Info". Below this is a "RESOURCE MANAGER" section with tabs for "Histogram", "Scatter Plot 2D", "Scatter Plot 3D", "Line Plot", and "Image Viewer". The "Image View" section displays a large, colorful astronomical image of a nebula. Below the image, the "Graphic Options" section includes a "Workspace" dropdown set to "alfonso" and an "Image" dropdown set to "sky.fits". A "Load Image" button is on the left, and a "Save Image as" button is circled in red. To the right of the "Save Image as" button are icons for a window, a hand, and a zoom slider set to 1/100.

Creazione GRAFICI

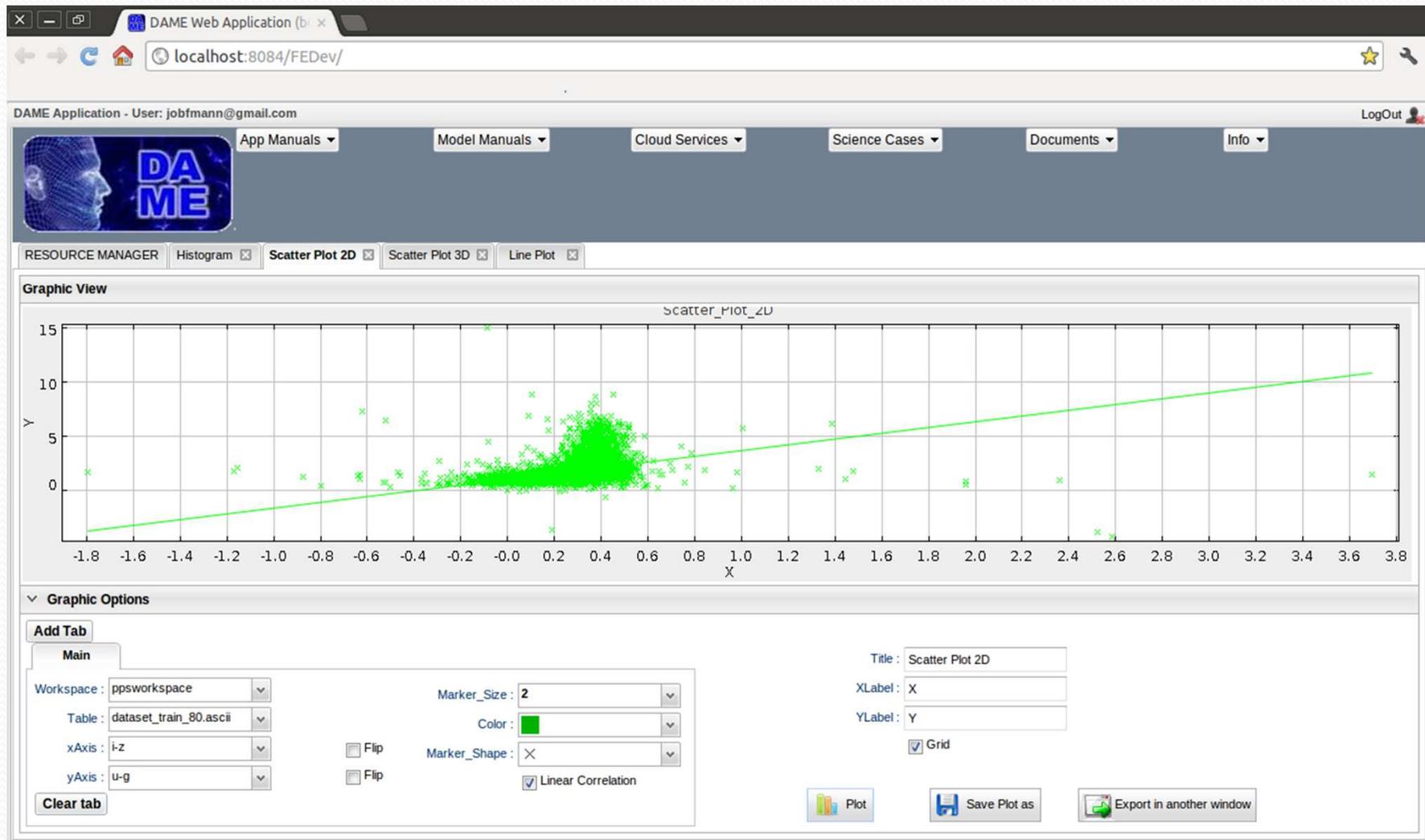
Si è dotata la Web Application di strumenti in grado di rappresentare i dati sotto forma di grafici, in particolare di:

- Histogram
- Scatter Plot 2D
- Scatter Plot 3D
- Line Plot

Histogram



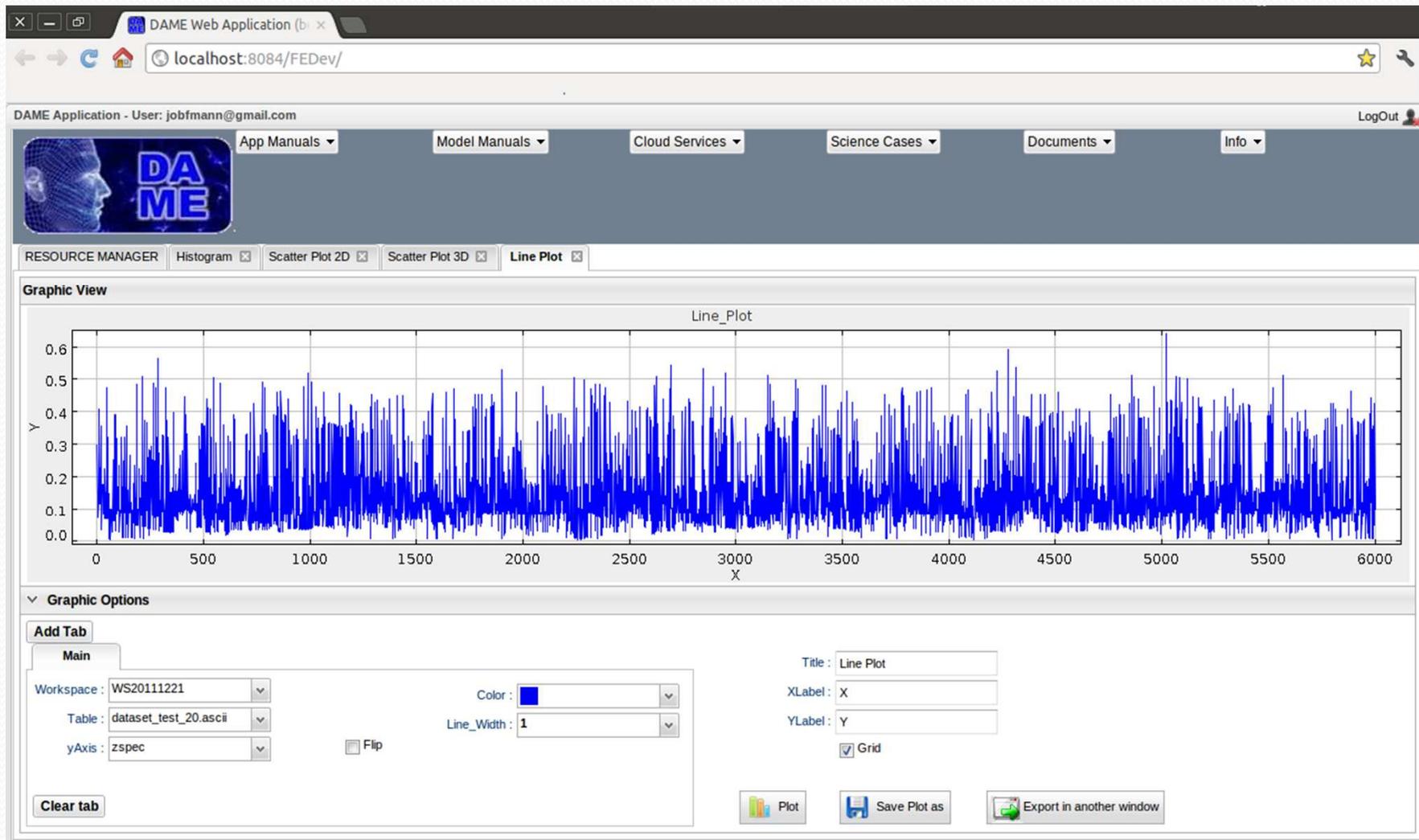
Scatter Plot 2D



Scatter Plot 3D

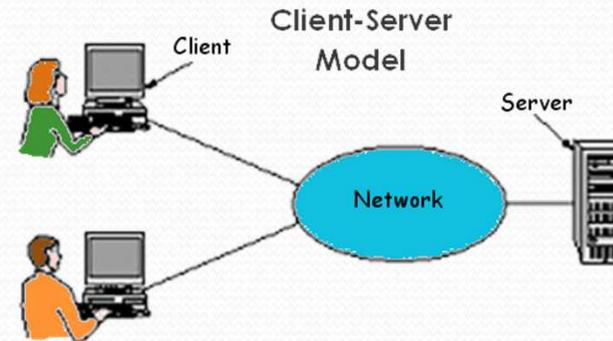
The screenshot displays the DAME Web Application interface. At the top, the browser address bar shows 'localhost:8084/FEDev/'. The application header includes the user 'jobfmann@gmail.com' and a 'LogOut' button. A navigation menu contains 'App Manuals', 'Model Manuals', 'Cloud Services', 'Science Cases', 'Documents', and 'Info'. Below this, a tabbed interface shows 'Scatter Plot 3D' as the active view. The main area, titled 'Graphic View', contains a 3D scatter plot labeled 'Scatter_Plot_3D' with red triangular markers. The plot axes are labeled X, Y, and Z, with numerical scales. Below the plot is the 'Graphic Options' panel, which is circled in red. This panel includes an 'Add Tab' section with tabs 'Main', 'A', 'B', and 'C'. The 'Main' tab is active, showing settings for 'Workspace' (manual), 'Table' (train.csv), 'xAxis' (r-i), 'yAxis' (z), and 'zAxis' (u-g). There are also 'Flip' checkboxes for each axis. Other settings include 'Marker_Size' (1), 'Color' (red), and 'Marker_Shape' (triangle). The right side of the panel has fields for 'Title' (Scatter Plot 3D), 'Phi' (210), 'Theta' (15), 'XLabel' (X), 'YLabel' (Y), and 'ZLabel' (Z). Checkboxes for 'Grid' and 'Fog' are also present. At the bottom of the panel are buttons for 'Plot', 'Save Plot as', and 'Export in another window'. Four 3D view icons are shown, with the second one highlighted in yellow.

Line Plot



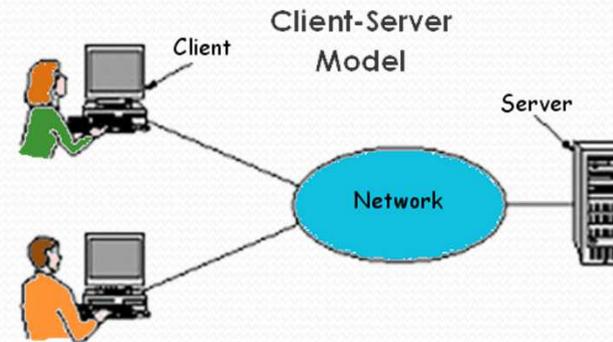
Da che lato produrre i grafici?

- Lato client
- Lato server



Da che lato produrre i grafici?

- Lato client
- Lato server



Vantaggi:

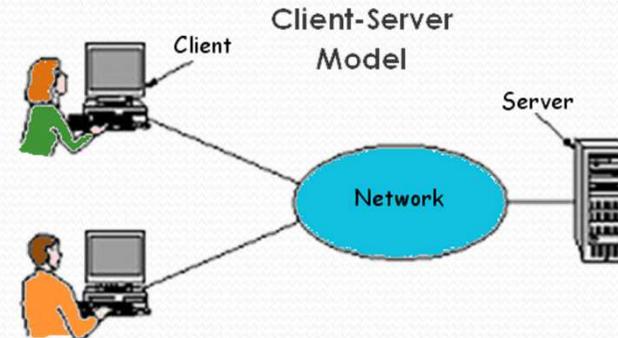
- ✓ Interattività migliorata
- ✓ Riduzione del carico di lavoro sul server

Svantaggi:

- ❖ Aumento di latenza della rete
- ❖ Riduzione delle funzionalità offerte

Da che lato produrre i grafici?

- Lato client
- **Lato server**



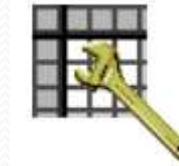
Vantaggi:

- ✓ Nessun passaggio di file dati tra client e server
- ✓ Possibilità di sfruttare le potenzialità del server
- ✓ Utilizzo di librerie grafiche complesse

Svantaggi:

- ❖ Riduzione delle risorse a disposizione degli esperimenti
- ❖ Ripetute richieste di produzione di grafici

STIL/STILTS

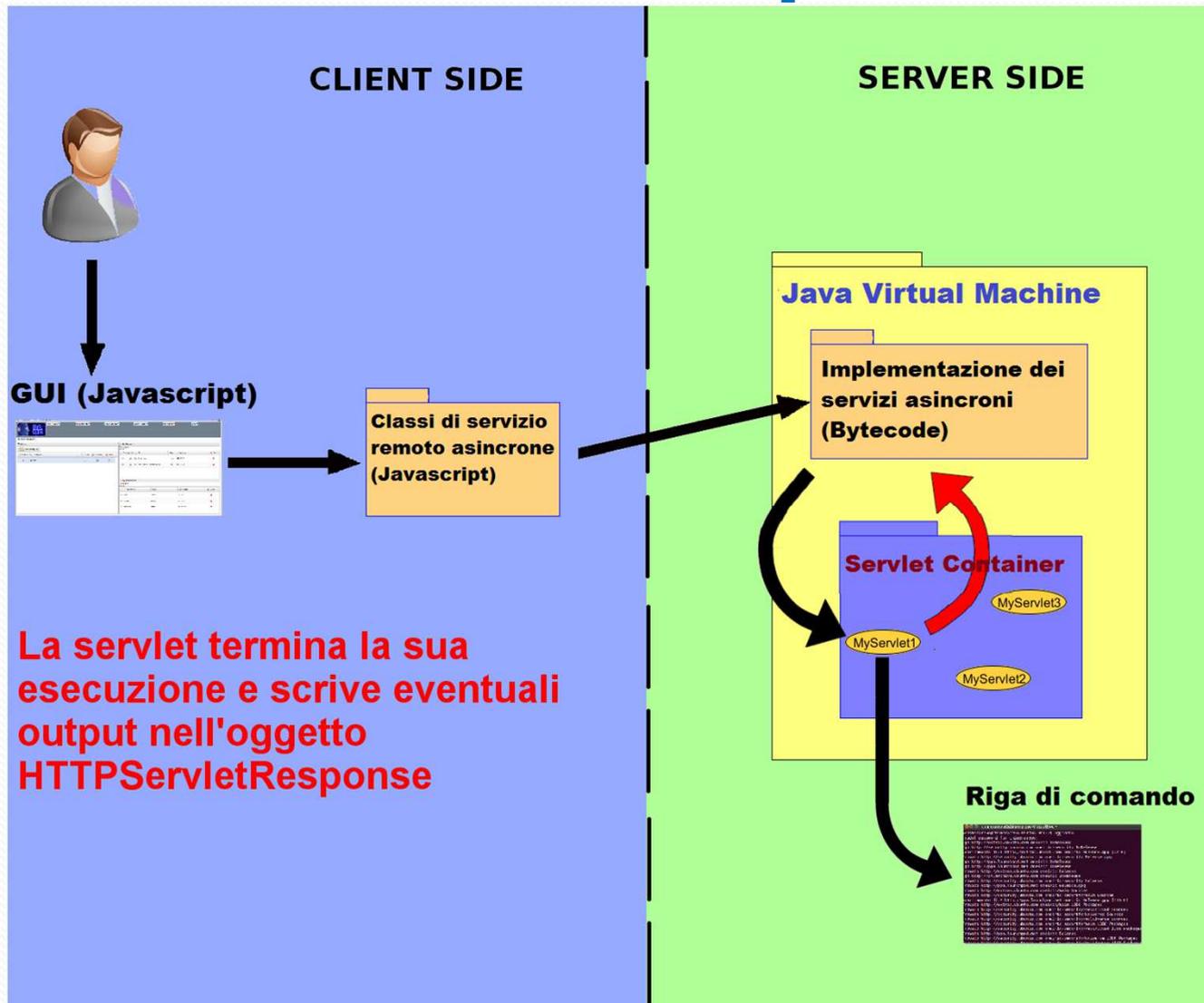


Libreria Java open source per il processing di dati tabulari e creazione di grafici

Caratteristiche:

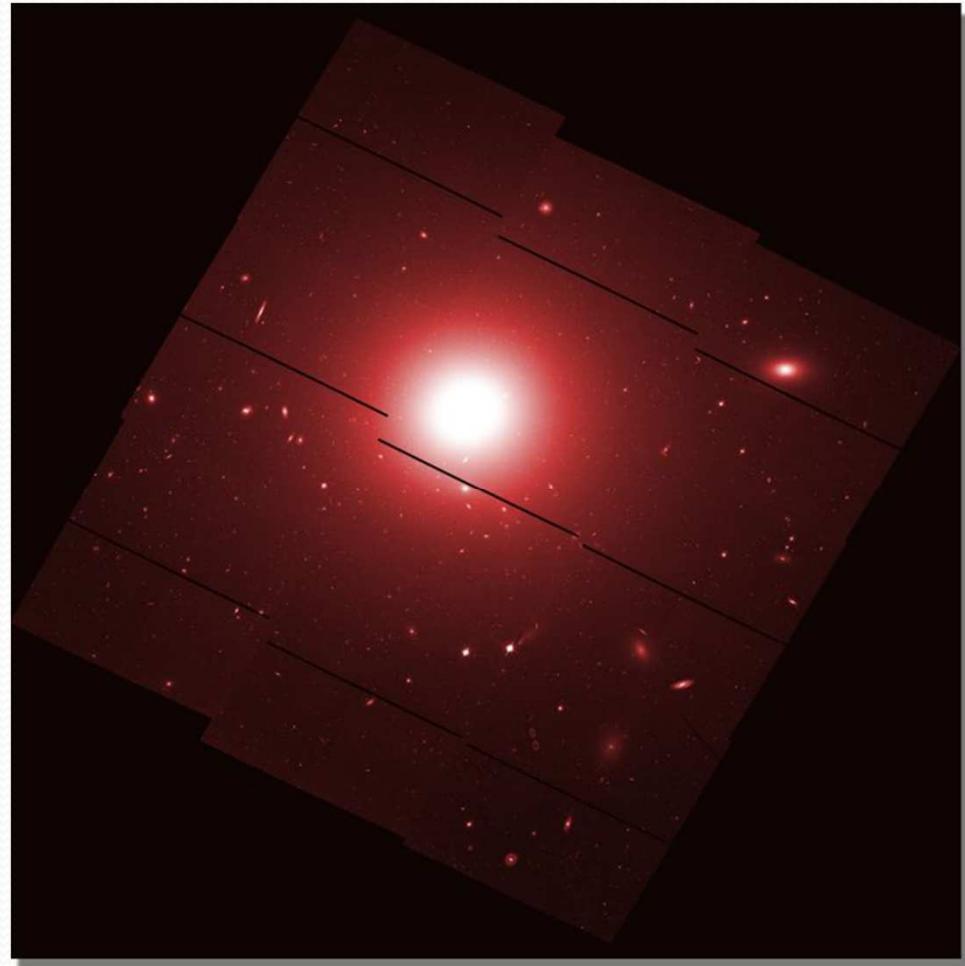
- ottenibile sotto forma di script
- facile da usare
- dimensione degli input illimitata
- buona copertura dei tipi di grafici

Interazione fra componenti



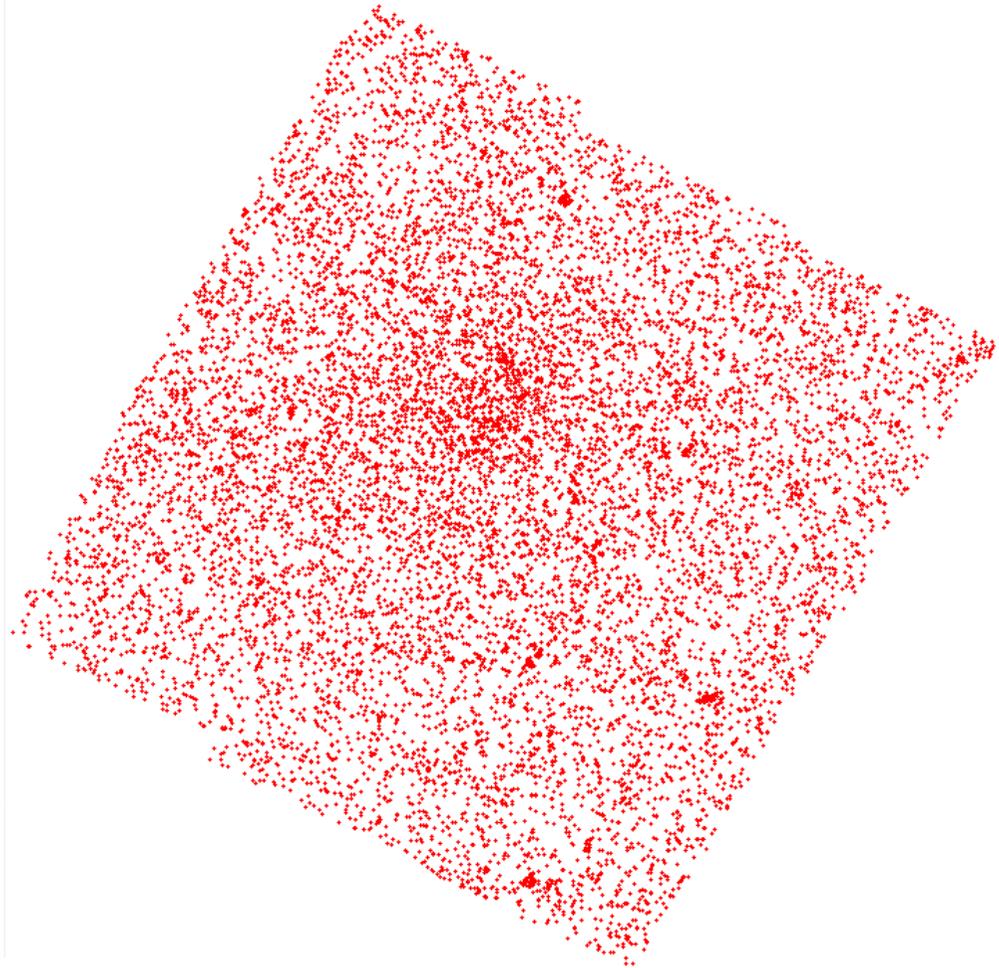
TEST DELLE FUNZIONALITÀ

Osservazione astronomica di una regione di cielo. L'obiettivo è quello di individuare eventuali ammassi globulari



TEST DELLE FUNZIONALITÀ

File di tipo FITS
ottenuto da calcoli
astrometrici
dall'immagine
precedente.
L'immagine così
prodotta evidenzia
tutti gli oggetti
presenti non
distinguibili ad occhio
nudo.



TEST DELLE FUNZIONALITÀ

Table Browser for 1: opt_and_struct_full.txt

	MAG_ISO	MAG_APER1	MAG_APER2	MAG_APER3	KRON_RADIUS	ELLIPTICITY	FWHM_IMAGE	mu0	calr_c	calr_h	calr_t	target
1	24,4753	26,7468	24,3789	0,0205	3,72	0,067	4,12	16,25	-0,1139	1,822	51,29	0
2	24,2342	26,5263	24,1632	0,0196	3,5	0,027	4,01	16,61	0,1321	1,856	35,38	0
3	23,1554	25,5964	23,1654	0,016	3,5	0,032	4,09	14,47	-0,3295	2,638	129,2	1
4	22,6316	25,3519	22,6808	0,0151	3,5	0,039	4,69	16,33	0,8065	5,002	80,45	1
5	22,4708	24,4951	22,4699	0,0216	3,5	0,066	3,45	12,81	-0,3912	-7,425	5,66	0
6	23,9033	27,5896	23,9168	0,0255	4,49	0,272	9,63	19,99	8,397	14,79	88,5	1
7	24,1972	26,4219	24,0978	0,0192	3,7	0,079	4,04	15,72	-0,1447	1,514	44,77	0
8	20,2423	22,1866	20,2963	0,017	3,5	0,03	3,23	6,68	-0,6999	-0,1492	1,899	0
9	23,5134	26,0983	23,511	0,0167	3,76	0,05	4,55	16,6	0,3777	4,75	105,8	1
10	22,5967	25,1807	22,6182	0,0147	3,5	0,021	4,45	15,98	0,6535	3,615	52,13	1
11	21,8602	23,8834	21,8737	0,0171	3,5	0,077	3,25	12,2	-0,2371	-3,39	-0,2887	0
12	24,4292	26,6224	24,3097	0,0201	4,33	0,116	3,96	16,22	-0,1139	2,635	81,01	0
13	20,3466	22,2583	20,3962	0,0161	3,5	0,047	2,89	8,9	-0,5763	0,07634	-48,54	1
14	23,8686	26,4629	23,8326	0,0179	3,92	0,144	4,53	17,16	0,5617	4,345	76,83	0
15	24,1944	26,5159	24,0998	0,0188	4,26	0,109	4,2	15,52	-0,2987	5,187	322,2	0
16	23,3385	25,6679	23,3182	0,0166	3,5	0,042	4,04	15,12	-0,1447	1,514	44,77	1
17	24,8669	27,3614	24,7216	0,0247	4,49	0,075	4,79	18,6	1,204	5,504	73,48	0
18	23,6528	26,1355	23,609	0,0177	3,68	0,091	4,36	16,69	0,4391	3,505	60,31	1
19	23,1038	26,0645	23,1542	0,016	3,95	0,056	5,66	17,97	2,176	7,619	84,62	1
20	23,435	25,2643	23,3951	0,0162	3,5	0,129	3,05	13,35	-0,3604	0,3709	-34,75	0
21	25,1723	27,781	24,9976	0,0308	4,6	0,061	5,39	19,58	2,146	6,942	71,81	0
22	24,7851	27,2661	24,6492	0,0251	4,22	0,345	5,33	17,54	1,6	6,105	71,81	0
23	22,3268	24,3309	22,3327	0,0345	3,5	0,06	3,27	14,63	0,5004	0,5712	5,914	0
24	21,8194	24,2154	21,8633	0,0139	3,5	0,063	3,97	13,76	-0,08314	2,483	70,97	1
25	19,9421	23,1152	20,3493	0,0136	3,5	0,17	4,57	14,33	1,234	20,81	755,	0

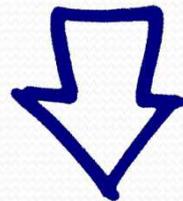
La classificazione dei punti ci permette di rappresentare ogni singolo oggetto come una riga della tabella.

TEST DELLE FUNZIONALITÀ

Pre-processing di esperimenti di data mining su tale dataset.

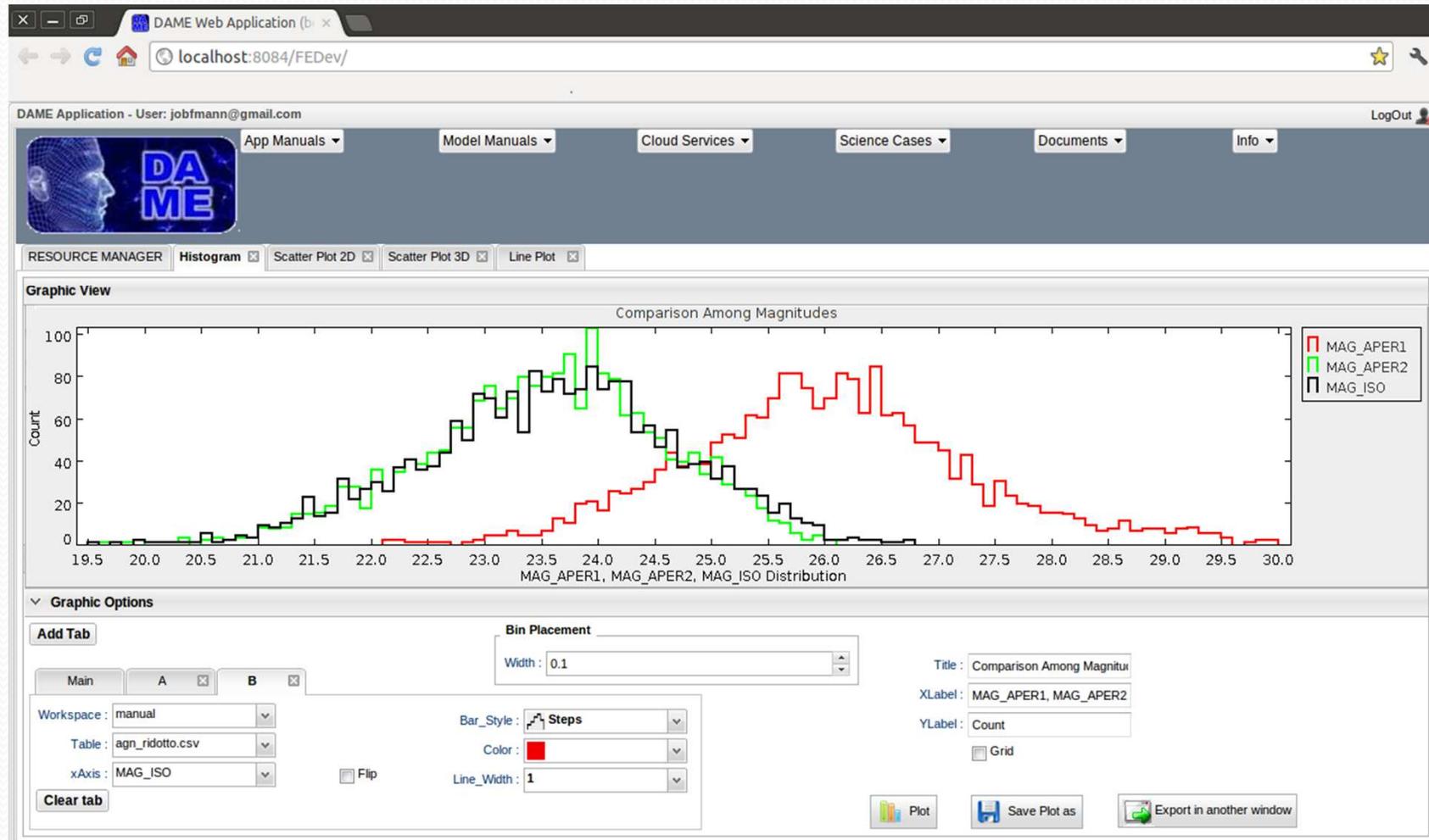


Utilizzo dei tools grafici per rappresentare i dati sotto forma di grafici di vario tipo.

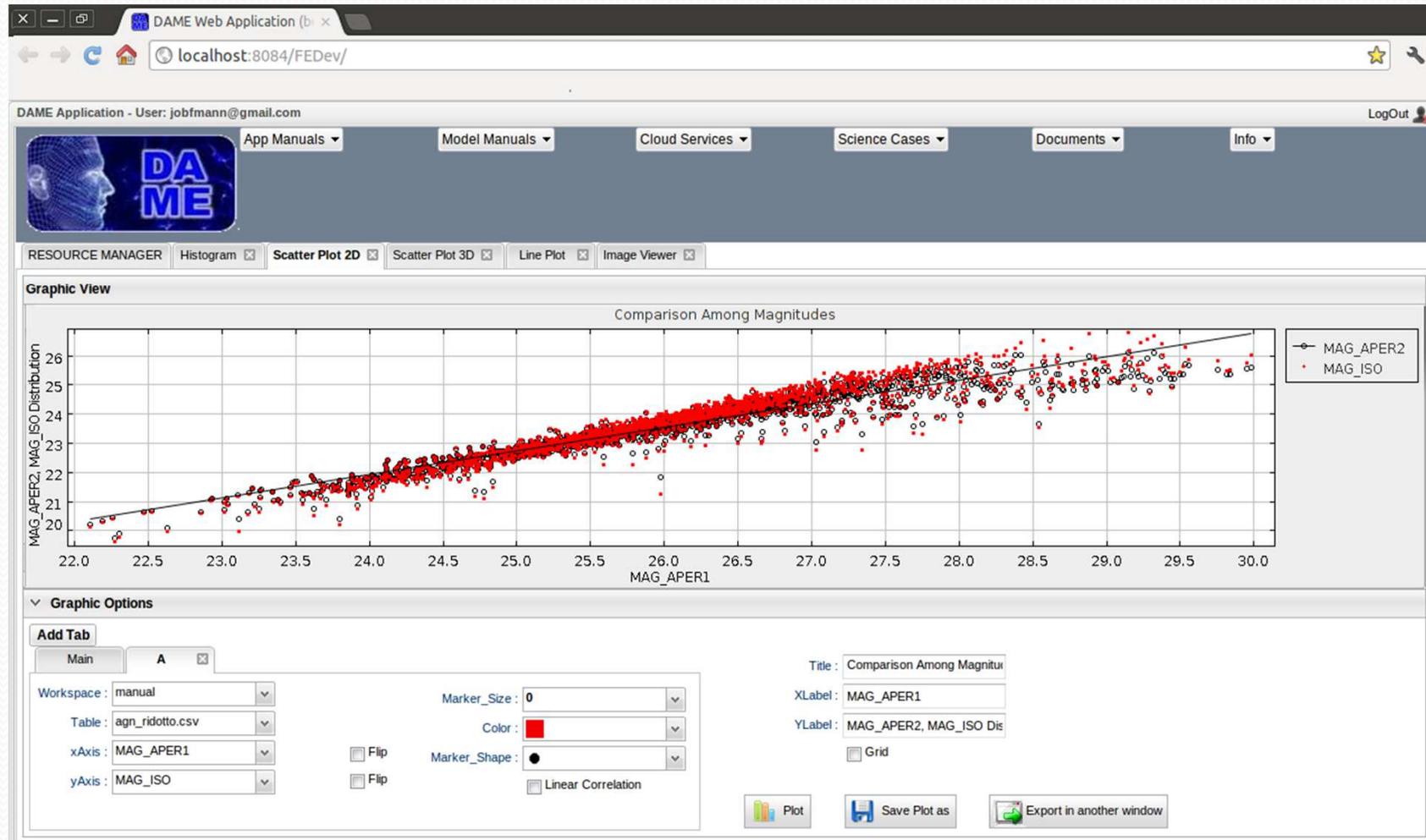


Individuazione visiva delle caratteristiche degli oggetti.

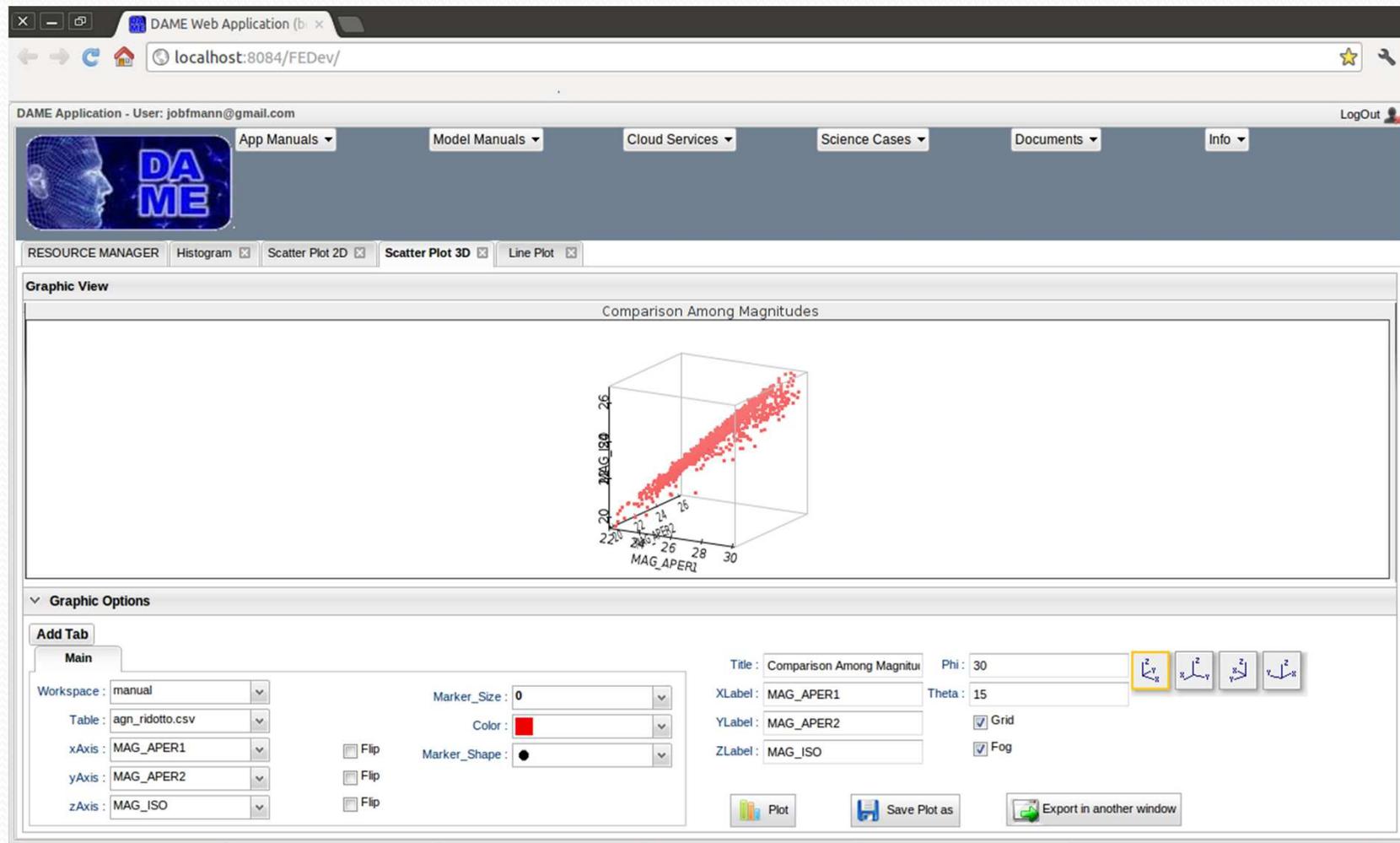
TEST DELLE FUNZIONALITÀ



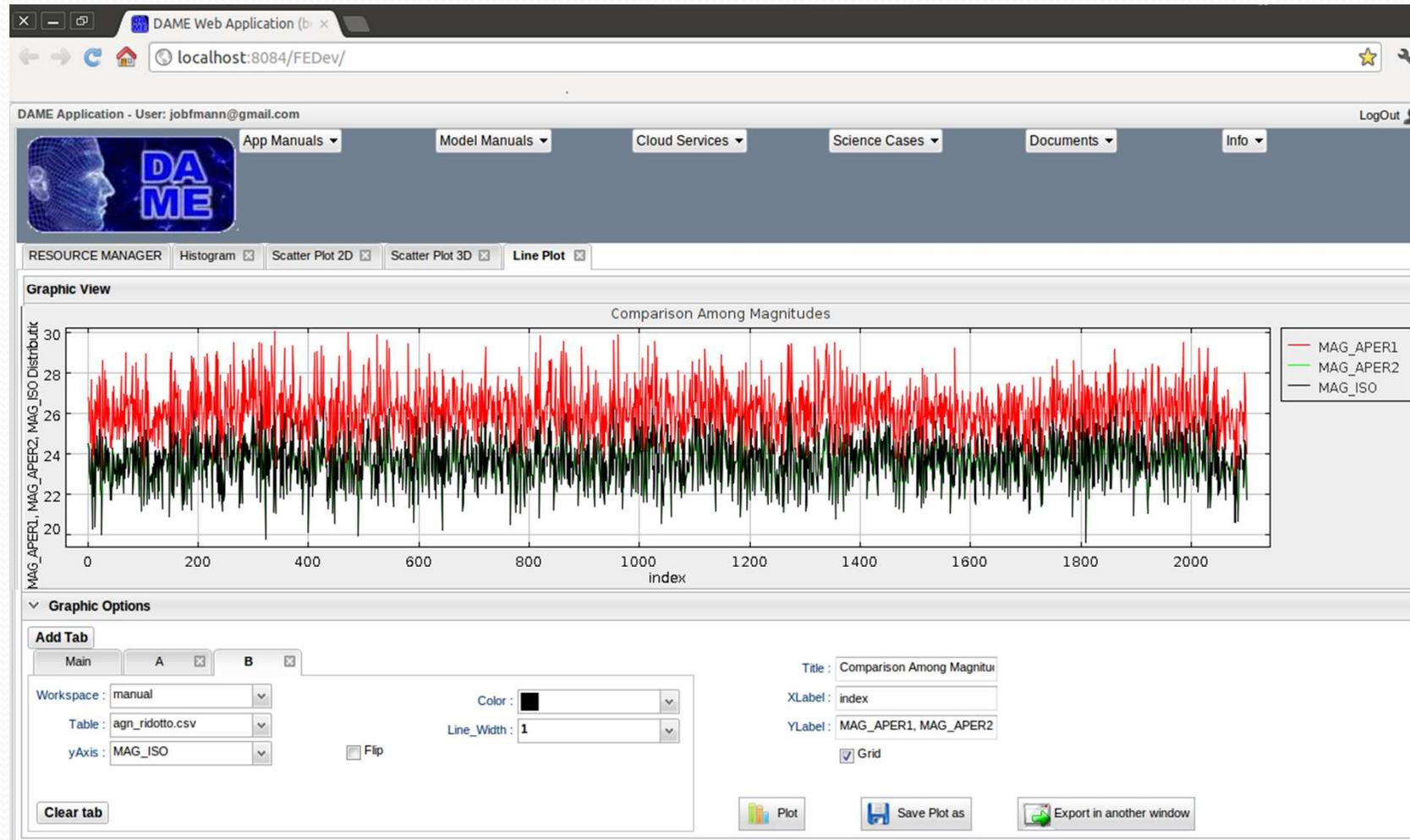
TEST DELLE FUNZIONALITÀ



TEST DELLE FUNZIONALITÀ



TEST DELLE FUNZIONALITÀ



CONCLUSIONI E SVILUPPI FUTURI

- Integrazione di nuovi tipi di grafici

