Using Microsoft Analysis Service to analyze graduates' performances and working conditions

<u>Alberto Leone</u>, Loris Cancellieri, Angelo Guerriero and Andrea Cammelli

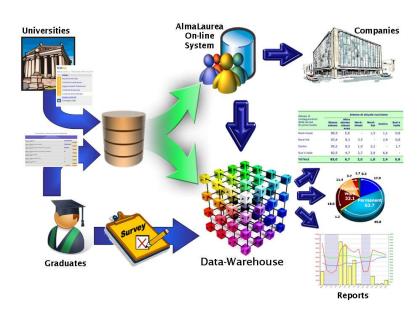
Consorzio Interuniversitario AlmaLaurea, Italy Università di Bologna, Italy





Summary

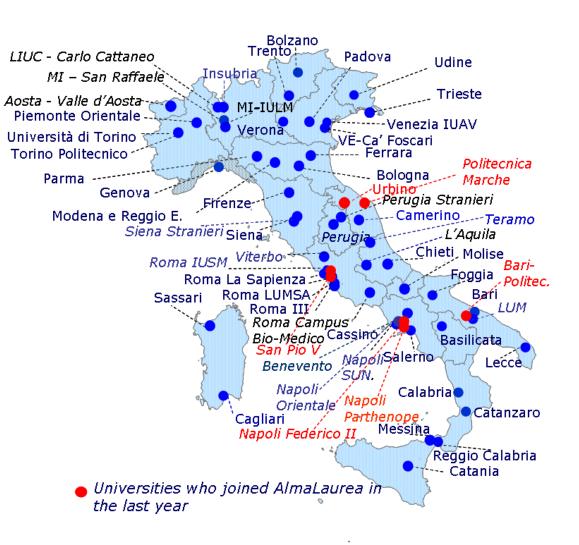
- AlmaLaurea
- Standard reports
- New Data warehouse design and implementation
- Software architecture
- First results
- Conclusions







AlmaLaurea - Consortium of Italian Universities

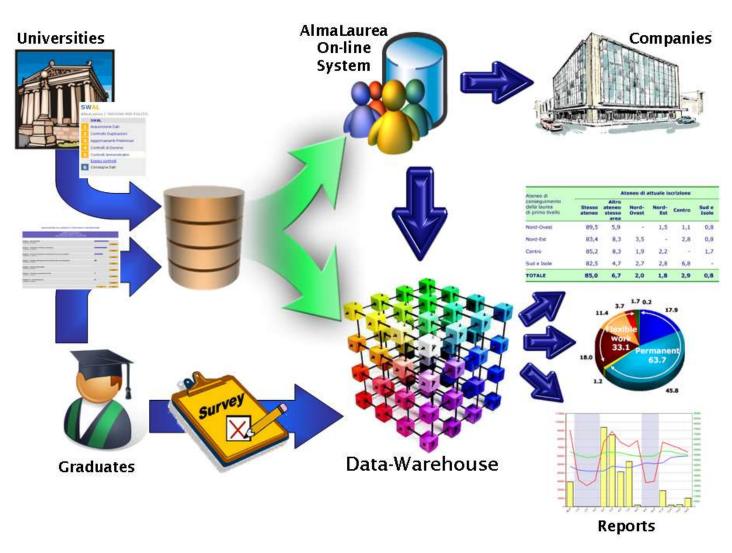


- Started 1994 by Bologna University
- 60 Universities (over 80)
- 1.500.000 grads
- 200.000 new grads/year
- 75% Italian graduates per year





AlmaLaurea



universities institutions press

researchers

students



Standard inputs

- Administrative data from Universities (100%)
- Snapshot and survey at graduation (92%)
- Employment and studying conditions
 - 1 year after graduation (90%)
 - 3 years after graduation (82%)
 - 5 years after graduation (76%)
 - 10 years after graduation ... coming soon



Standard outputs

Graduates' profile

- June 2010: results about 2009

Graduates' working and studing conditions

- March 2010: results about 2009 survey
 - Graduates in 2008, 1 year after graduation
 - Graduates in 2006, 3 years after graduation
 - Graduates in 2004, 5 years after graduation

Documentation:

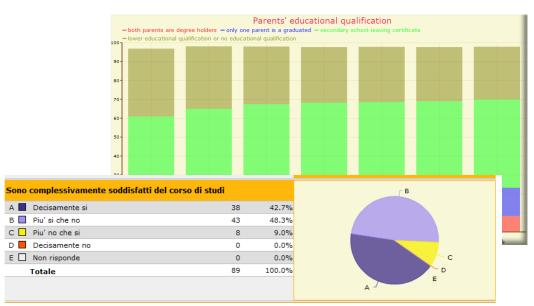
- Annual reports, general considerations
- Detailed reports, at course level (available online)
- Custom reports (e.g. transparency information)
- Microdata (transferred back to universities' statistical bureau)





Summary

- AlmaLaurea
- Standard reports
- New Data warehouse design and implementation
- Software architecture
- First results
- Conclusions



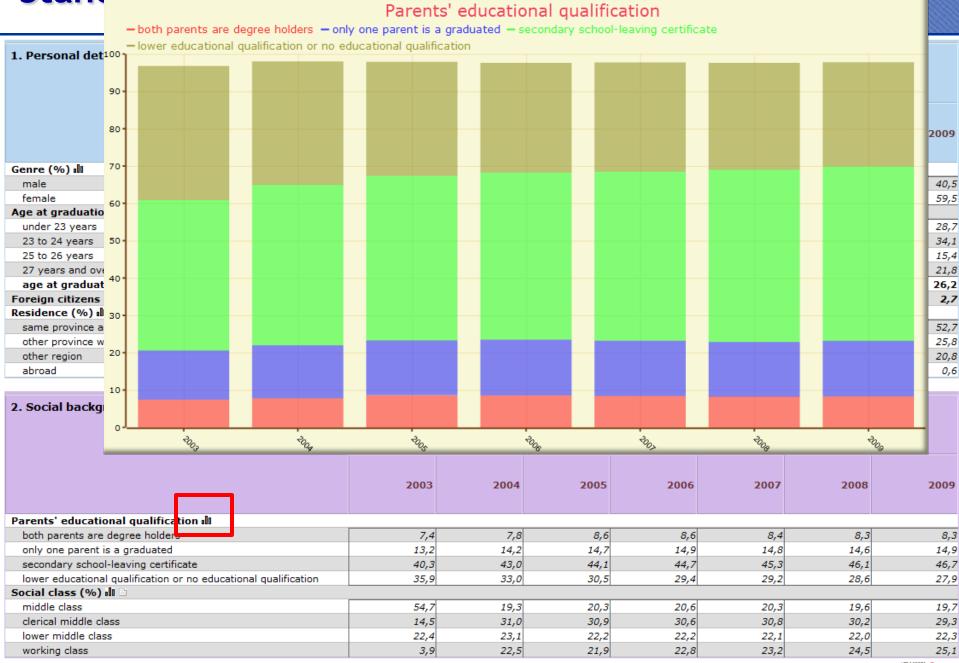




Standard report - sample

Parents' educational qualification III both parents are degree holders 7,4 7, only one parent is a graduated 13,2 14, secondary school-leaving certificate 40,3 43,
both parents are degree holders 7,4 7, only one parent is a graduated 13,2 14,
only one parent is a graduated 13,2 14,
only one parent is a graduated 13,2 14,
lower educational qualification or no educational qualification 35,9 33,
Social class (%) III 🖹
uner provinde Hamm the same region 5272 2072 2072 2072 2072 2072
other region 24,7 23,7 22,5 21,8 20,7 21,0
abroad 0,4 0,4 0,5 0,6 0,6
2. Social background Selected cohort (compared by graduation year) 1-5 6-7 1-7
2003 2004 2005 2006 2007 2008 2
Parents' educational qualification ill
both parents are degree holders 7,4 7,8 8,6 8,6 8,4 8,3
only one parent is a graduated 13,2 14,2 14,7 14,9 14,8 14,6
only one parent is a graduated 13,2 14,2 14,7 14,9 14,8 14,6 secondary school-leaving certificate 40,3 43,0 44,1 44,7 45,3 46,1
only one parent is a graduated 13,2 14,2 14,7 14,9 14,8 14,6 secondary school-leaving certificate 40,3 43,0 44,1 44,7 45,3 46,1 lower educational qualification or no educational qualification 35,9 33,0 30,5 29,4 29,2 28,6
only one parent is a graduated 13,2 14,2 14,7 14,9 14,8 14,6 secondary school-leaving certificate 40,3 43,0 44,1 44,7 45,3 46,1 lower educational qualification or no educational qualification 35,9 33,0 30,5 29,4 29,2 28,6 Social class (%) ill :
only one parent is a graduated 13,2 14,2 14,7 14,9 14,8 14,6 secondary school-leaving certificate 40,3 43,0 44,1 44,7 45,3 46,1 lower educational qualification or no educational qualification 35,9 33,0 30,5 29,4 29,2 28,6 Social class (%) ill ill ill ill ill ill ill ill ill il
only one parent is a graduated 13,2 14,2 14,7 14,9 14,8 14,6 secondary school-leaving certificate 40,3 43,0 44,1 44,7 45,3 46,1 lower educational qualification or no educational qualification 35,9 33,0 30,5 29,4 29,2 28,6 Social class (%) ill :

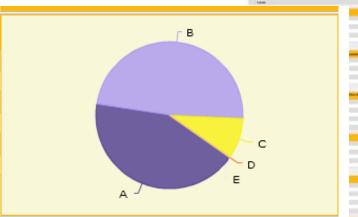
Standard report - sample



Transparency info

- Detailed info about (requested by Ministry for orientation purposes)
 - Employability
 - Degree effectiveness
 - Assessment of university experience

Sono	complessivamente soddisfatti del corso di st	udi	
A	Decisamente si	38	42.7%
В	Piu' si che no	43	48.3%
С	Piu' no che si	8	9.0%
D 📙	Decisamente no	0	0.0%
E \square	Non risponde	0	0.0%
	Totale	89	100.0%





Bologna process results: ten years after

Some statistical result:

- Number of students obtaining the degree within the legal duration:
 - 9.5% before the Bologna process (2001)
 - **39.2%** after (2009)
- Age at graduation
 - 27.2 years before
 - 25.0 years after (23.9 1° level, 25.0 2° level, 26.1 2° level European degree)
- Number of graduates under 23 years
 - None before
 - 17% after
- Stage
 - 17.9% before
 - **54,5%** after





Summary

- AlmaLaurea
- Standard reports
- New Data warehouse design and implementation
- Software architecture
- First results
- Conclusions



Available data Sources

Standard sources

- Data from universities
- Survey data

AlmaLaurea "events"

- Activity on the CV database
 - Browse, search and selection by employers
 - CV update by graduates
- Activity on the website
 - Newsletters and alert response
 - Website navigation
 - Browse , search and apply to Job offers
 - Services usage
 - Help, faq





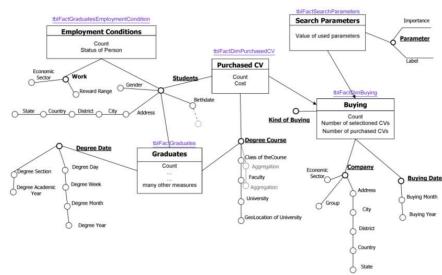
Data collection process

- Data are collected by several and heterogeneus sources
 - Online surveys (CAWI)
 - Telephone surveys (CATI)
 - Relational databases
 - Local database
 - Remote database (administrative data from Universities)
 - Log files



Summary

- AlmaLaurea
- Standard reports
- New Data warehouse design and implementation
- Software architecture
- First results
- Conclusions





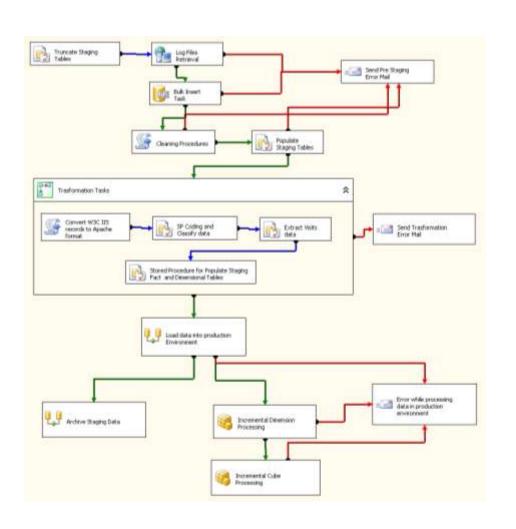
Data Warehouse software architecture

Analysis tools browser Report service custom reports datasets report server data mart (mdx) sql Analysis service data warehouse (data and metadata) Sql server staging (cleaning area) Sql server Sql server extract & transform tools integration services data sources





ETL modules

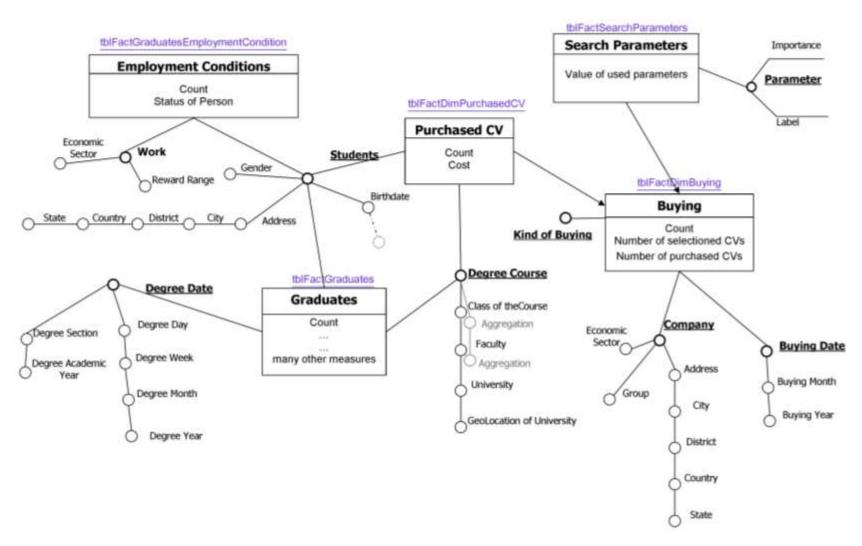


- ETL modules have been implemented by Sql
 Server Integration
 Services package
- Remote databases:
 - Web services (REST and SOAP)
 - Manual Data transfer
- Custom interactive application for administrative data from Universities (SWAL)





Core data warehouse structure - Logical design

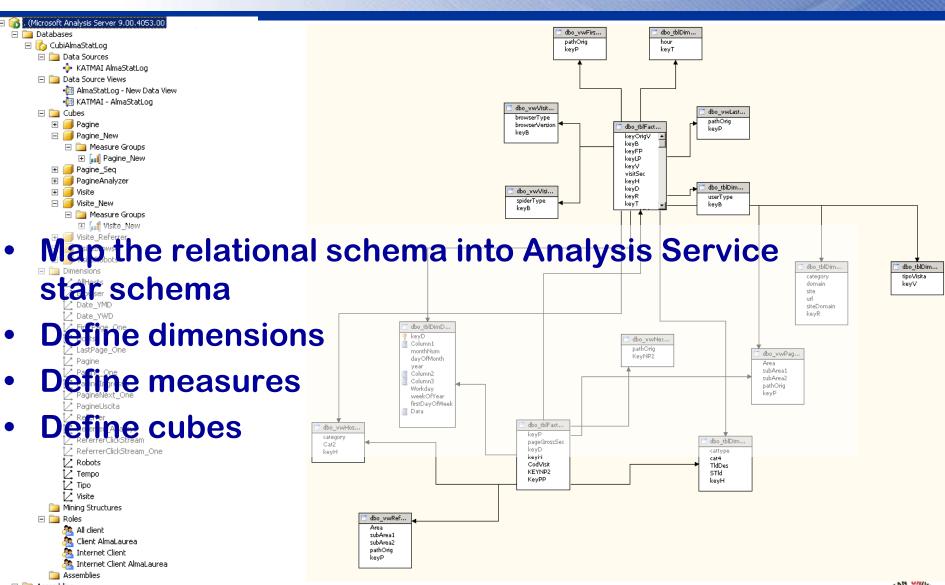






Analysis Service: mapping the data

LMA AUREA



Metadata

- Technical (internal) metadata
 - Mapping and transformations
 - Update status
- Business (external) metadata
 - Data description
 - Source, availability, reliability
 - Rights
 - Dictionary
 - Type





Roles, reports: users vs report type

		us	er						
	anonymous	researcher	University staff	Ministry staff					
Dynamic report (pre- built)	YES (*)	YES	YES	YES					
Static report	NO	YES	YES	YES					
Report builder	NO	YES	NO	NO					

Roles, reports: users vs data sets

		us	er						
	anonymous	researcher	University staff (*)	Ministry staff					
Graduation	YES	YES	YES	YES					
Working Conditions	YES	YES	YES	YES					
Company Activities	NO	YES							
Graduates activities	NO	YES	YES	YES					

(*) limited to single university data



Summary

- AlmaLaurea
- Standard reports
- New Data warehouse design and implementation
- Software architecture
- First results
- Conclusions

Profilo - Indice di Ritardo

Gruppo disciplinare - Titoli dei genitori

		<u>Titoli dei genitori</u>														
Gruppo Statistico agrario		Solo uno dei due con diploma superiore		Solo uno due con l		Titoli infe nessun t		Tutti e du diploma superiore		Tutti e du Laurea	ie con	Undefine	d	Totale		
		0.40	708	0.41	482	0.53	1034	0.40	699	0.31	278	0.72	317	0.46	3518	
architettura		0.47	1772	0.42	1524	0.61	2086	0.45	2019	0.36	1037	0.81	865	0.51	9303	
chimico-farmaceutico		0.35	891	0.39	841	0.49	1267	0.34	1047	0.32	545	0.55	375	0.40	4966	
economico-statistico		0.41	4918	0.38	3342	0.54	6396	0.38	5267	0.30	1926	0.79	2052	0.46	23901	
educazione fisica		0.33	531	0.31	292	0.33	556	0.30	581	0.33	101	0.30	483	0.32	2544	
geo-biologico		0.35	1755	0.32	1289	0.48	1977	0.33	2017	0.30	785	0.53	976	0.39	8799	
giuridico		0.67	2233	0.56	2439	0.86	3106	0.63	2827	0.41	1713	1.04	3110	0.73	15428	
ingegneria		0.45	4086	0.40	3615	0.60	4382	0.43	4929	0.34	2387	0.74	960	0.47	20359	
insegnamento		0.46	1992	0.44	765	0.53	3602	0.42	1510	0.49	250	0.54	1763	0.49	9882	
letterario		0.52	3193	0.47	2853	0.72	3796	0.48	3887	0.39	1821	0.81	2737	0.57	18287	
linguistico		0.40	2205	0.38	1315	0.59	2556	0.38	2311	0.32	739	0.80	1058	0.48	10184	
medico		0.16	3467	0.21	2278	0.18	5617	0.18	3064	0.15	1594	0.27	3258	0.19	19278	
politico-sociale		0.38	4872	0.38	3421	0.43	6824	0.36	5435	0.33	1901	0.50	2951	0.40	25404	
psicologico		0.34	2129	0.29	1423	0.43	2711	0.30	2330	0.27	826	0.46	586	0.35	10005	
scientifico		0.46	1049	0.44	870	0.62	1243	0.40	1206	0.27	631	0.63	502	0.47	5501	
Totale		0.41	35801	0.39	26749	0.52	47153	0.39	39129	0.32	16534	0.65	21993	0.45	187359	

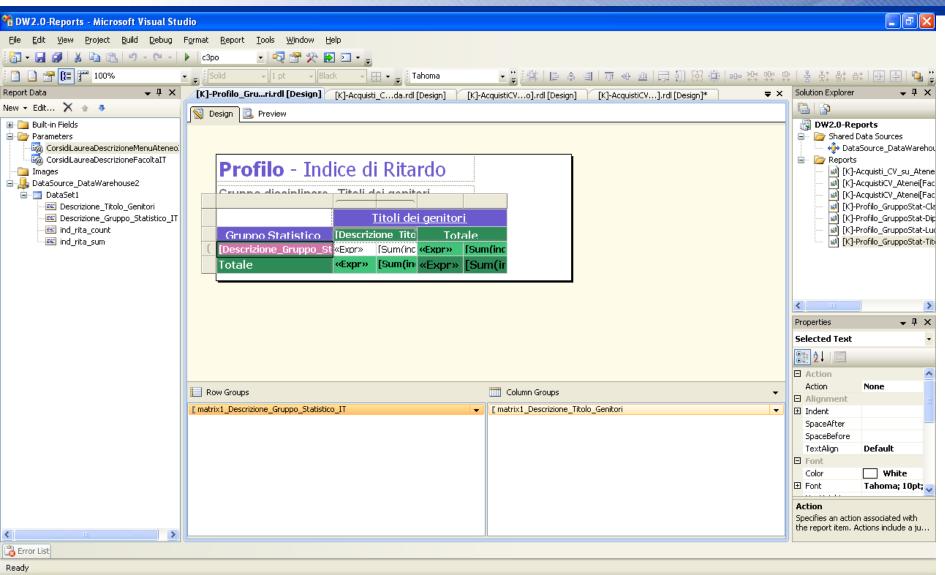




Reporting service development

- Implementation of some test reports to verify
 - Flexibility
 - Usability
 - Compatibility
- Test reports:
 - Data selection (by university, ...). Fixed report structure
 - Data selection. Drill down over rows or columns

Report designer





Sample one: delay in degree completion times study

Profilo - Indice di Ritardo

Gruppo disciplinare - Titoli dei genitori

All the universities

						titolo diploma superiore Laurea 3 1034 0.40 699 0.31 278 0.72 317 0.46 3518 4 2086 0.45 2019 0.36 1037 0.81 865 0.51 9303 9 1267 0.34 1047 0.32 545 0.55 375 0.40 4966 4 6396 0.38 5267 0.30 1926 0.79 2052 0.46 23901 3 556 0.30 581 0.33 101 0.30 483 0.32 2544 3 1977 0.33 2017 0.30 785 0.53 976 0.39 8799 3 3106 0.63 2827 0.41 1713 1.04 3110 0.73 15428 3 3602 0.42 1510 0.49 250 0.54 1763 0.49 9882 2 3796 0.4										
Gruppo Statistico	Solo uno due con superio	ı diploma	Solo uno due con l		Titoli infe nessun t		diploma			ie con	Undefine	d	Tot	ale		
agrario	0.40	708	0.41	482	0.53	1034	0.40	699	0.31	278	0.72	317	0.46	3518		
architettura	0.47	7 1772	0.42	1524	0.61	2086	0.45	2019	0.36	1037	0.81	865	0.51	9303		
chimico-farmaceutico	0.35	5 891	0.39	841	0.49	1267	0.34	1047	0.32	545	0.55	375	0.40	4966		
economico-statistico	0.41	1 4918	0.38	3342	0.54	6396	0.38	5267	0.30	1926	0.79	2052	0.46	23901		
educazione fisica	0.33	3 531	0.31	292	0.33	556	0.30	581	0.33	101	0.30	483	0.32	2544		
geo-biologico	0.35	1755	0.32	1289	0.48	1977	0.33	2017	0.30	785	0.53	976	0.39	8799		
giuridico	0.67	7 2233	0.56	2439	0.86	3106	0.63	2827	0.41	1713	1.04	3110	0.73	15428		
ingegneria	0.45	4086	0.40	3615	0.60	4382	0.43	4929	0.34	2387	0.74	960	0.47	20359		
insegnamento	0.46	1992	0.44	765	0.53	3602	0.42	1510	0.49	250	0.54	1763	0.49	9882		
letterario	0.52	3193	0.47	2853	0.72	3796	0.48	3887	0.39	1821	0.81	2737	0.57	18287		
linguistico	0.40	2205	0.38	1315	0.59	2556	0.38	2311	0.32	739	0.80	1058	0.48	10184		
medico	0.16	3467	0.21	2278	0.18	5617	0.18	3064	0.15	1594	0.27	3258	0.19	19278		
politico-sociale	0.38	4872	0.38	3421	0.43	6824	0.36	5435	0.33	1901	0.50	2951	0.40	25404		
psicologico	0.34	1 2129	0.29	1423	0.43	2711	0.30	2330	0.27	826	0.46	586	0.35	10005		
scientifico	0.46	1049	0.44	870	0.62	1243	0.40	1206	0.27	631	0.63	502	0.47	5501		
Totale	0.41	35801	0.39	26749	0.52	47153	0.39	39129	0.32	16534	0.65	21993	0.45	187359		





Sample one: delay in degree completion times study

Profilo - Indice di Ritardo

Gruppo disciplinare - Titoli dei genitori

Bologna university

			ploma due con laurea nessun titolo diploma superiore Laurea 94 0.26 67 0.53 87 0.37 90 0.32 49 0.85 32 0.38 58 0.38 41 0.38 67 0.30 58 0.41 34 0.63 8 0.36 63 0.33 102 0.47 95 0.31 81 0.35 68 0.43 19 0.36 454 0.31 382 0.39 500 0.30 489 0.22 221 0.85 155 0.36 47 0.24 32 0.38 47 0.31 52 0.08 12 0.55 39 0.36 119 0.25 131 0.31 96 0.22 165 0.23 74 0.43 54 0.26 163 0.48 225 0.80 173 0.59 231 0.34 177 0.98 161												
Gruppo Statistico	Solo ur due co superi	n diploma			ı		diploma			ie con	Undefine	d	Totale		
agrario	0.2	1 94	0.26	67	0.53	87	0.37	90	0.32	49	0.85	32	0.38	419	
architettura	0.3	1 58	0.38	41	0.38	67	0.30	58	0.41	34	0.63	8	0.36	266	
chimico-farmaceutico	0.2	8 63	0.33	102	0.47	95	0.31	81	0.35	68	0.43	19	0.36	428	
economico-statistico	0.3	5 454	0.31	382	0.39	500	0.30	489	0.22	221	0.85	155	0.36	2201	
educazione fisica	0.3	7 47	0.24	32	0.38	47	0.31	52	0.08	12	0.55	39	0.36	229	
geo-biologico	0.2	2 119	0.25	131	0.31	96	0.22	165	0.23	74	0.43	54	0.26	639	
giuridico	0.6	6 163	0.48	225	0.80	173	0.59	231	0.34	177	0.98	161	0.63	1130	
ingegneria	0.3	9 372	0.35	370	0.51	364	0.40	459	0.34	287	0.97	31	0.41	1883	
insegnamento	0.3	0 199	0.49	102	0.38	307	0.31	174	0.51	40	0.54	90	0.38	912	
letterario	0.4	6 424	0.40	459	0.56	394	0.40	492	0.40	293	0.75	627	0.51	2689	
linguistico	0.3	1 169	0.28	125	0.50	163	0.30	177	0.27	89	0.76	54	0.37	777	
medico	0.1	7 195	0.16	168	0.20	242	0.17	186	0.17	136	0.24	112	0.18	1039	
politico-sociale	0.3	0 442	0.26	422	0.47	429	0.30	511	0.26	294	0.51	170	0.34	2268	
psicologico	0.1	2 116	0.19	73	0.28	117	0.15	102	0.13	36	0.66	8	0.19	452	
scientifico	0.3	5 106	0.46	113	0.53	132	0.29	131	0.33	80	0.62	28	0.41	590	
Totale	0.3	4 3021	0.33	2812	0.45	3213	0.33	3398	0.30	1890	0.69	1588	0.39	15922	





Sample two: relationship between field of study and company type

Numero CV Acquistati

5/3/2010 1:25:55 PM

											1	ipo Azie	enda										
Facoltà	Classe di Laurea	Corso di Laurea	Assicurazioni	ioni Associazioni Azienda Azi Sconosciuta o di Sistema		Azie	ende			Enti	Enti forma:		Istituto Superiore		Liberi professionist	Ricerca i personale		Università		Total			
⊞ Agraria		Totale	14	14 2		20 192				9		869	∍			70		263		1439			
■ Architettura		Totale					14		96			35		86									231
⊞ Chimica industria	e	Totale	1	1		1		1	114			15		i			3					139	
□ Conservazione de	i beni culturali	Totale	12	2					12		10			74					32	38			180
	⊞ finanza (19/5)	Totale	10		19	4 7		73		84	6		202	2			166		31			595	
I	⊞ scienze dell'economia (64/S)	Totale	29	3		7		105		167		5			5		1	381		69			1477
	□ scienze dell'economia e della gestione	economia aziendale	33 1.07 %	8 0.	.26 %	4	0.13 %	139	4.51 %	259	8.40 %	24 0.78 %	1179	38.24 %	6	0.19 %	0.00 %	1379	44.73 %	52	1.69 %	3083	100.00 %
	aziendale (17)	economia e commercio - BOLOGNA	57 15.57 %		.55 %		0.00 %		3.55 %	32	8.74 %	1 0.27 %	94	25.68 %		0.00 %	0.00 %	167	45.63 %		0.00 %	366	100.00 %
		economia e gestione delle imprese	3 0.84 %	1 0.	.28 %	1	0.28 %	24	6.72 %	67	18.77 %	0.00 %		55.46 %		0.00 %	0.00 %	47	13.17 %	16	4.48 %	357	100.00 %
		economia e management dei servizi turist	11 20.37 %	0.	.00 %		0.00 %		1.85 %	3	5,56 %	0.00 %		51.85 %	2	3.70 %	0.00 %	7	12.96 %	2	3.70 %	54	100.00 %
		economia e marketing	42 2.67 %	1 0.	.06 %	4	0.25 %	99	6.29 %	134	8,52 %	12 0.76 %		56.45 %	1	0.06 %	0.00 %	361	22.95 %	31	1.97 %	1573	100.00 %
		economia e professione	13 1.37 %	3 0.	.32 %	1	0.11 %		6.41 %	83	8.73 %	5 0.53 %		42.27 %	2	0.21 %	0.00 %	366	38.49 %	15	1.58 %	951	100.00 %
		Totale	159		15		10		337		578	42	:	2789		11			2327		116		6384
	⊞ scienze economiche (28)	Totale	39		1		7		91		156	21		916					655		56		1942
6 € '	⊞ scienze economiche per l'ambiente e la cultura (83/5)	Totale	12		1				25		24	1		263					48		5		379





Summary

- AlmaLaurea
- Standard reports
- New Data warehouse design and implementation
- Software architecture
- First results
- Conclusions



Conclusions

- We are experimenting Microsoft Business Intelligence tools
- The data warehouse (still under construction) is a relational database (SQL server)
- The old reporting architecture provides:
 - Data mart extraction (SPSS analysis)
 - Custom reports
- The new reporting architecture (still under design) should provide
 - Flexible reports
 - Compound data mart





Conclusions ... the end

We are experimenting Microsoft Business Intelligence tools

- Analysis service integration
- Reporting service

First results

- Easy to design, but more investment and training is mandatory
- With respect to our custom reports architecture (sql based) we have:
 - Shorter development times
 - Better performances (even on the development architecture)

Open issues

Performance and hardware dimensioning





Questions?

To learn more about almalaurea www.almalaurea.net

1.370.000 CVs available also in English

More than 10 years of data on graduates'

profile and employability

Yearly reports



