

The impact of the Bologna Process on the labour market: a national data warehouse regarding universities and graduates

Alberto Leone¹, Loris Cancellieri², Angelo Guerriero³, Andrea Cammelli⁴

Consorzio Interuniversitario AlmaLaurea, Viale Masini 36, 40126 Bologna (BO), Italy

¹alberto.leone@almalaurea.it, ²loris.cancellieri@almalaurea.it, ³angelo.guerriero@almalaurea.it

⁴andrea.cammelli@almalaurea.it

Keywords

Graduates database, surveys, data warehouse, recruitment, Bologna Process

1. EXECUTIVE SUMMARY

The university reform has been started in Italy in 2001, consistently to the Bologna Process. Since that time many questions which needed an answer arose. Firstly, there was the necessity to examine the performances of university both in terms of teaching and of services provided, in order to make it possible to evaluate them. Secondly, it was important to understand whether the time needed to complete the academic studies had diminished or not. Thirdly, it was necessary to verify if the percentage of Italian graduates had arisen. Lastly, another significant question regarded how much the labour market has been appreciating the quality of the human capital.

1.1. Background

Since 1994 AlmaLaurea has been collecting a huge quantity of data regarding graduates coming from 52 of the 87 Italian universities. The amount of data collected is equal to the 70% of the total number of graduates who obtained their degree at Italian universities. What AlmaLaurea obtains is a picture of the graduates' condition after their studies, in which their performances and the social and cultural conditions of their families are examined. These data are collected through a questionnaire which also includes a detailed evaluation of the academic experience made.

Answering to the questionnaire automatically creates a CV for each graduate. If graduates agree, these CVs are collected in an on-line databank which can be accessed by companies, so that recruiting and hiring procedure can be fostered. These databank enables AlmaLaurea to create a kind of "graduates' register" and to update it. The databank also makes it possible to monitor how do graduates' situation evolve in terms of academic studies and employment condition. These information are collected through on-line and telephone interviews whose answer rate ranges between 70% and 95% or even more. Other information are collected through the use that both companies and graduates make of the databank, such as: posting job offers, sending back job applications and selection of the applicants for a job through the databank. This huge amount of information constitutes a national data warehouse which enables researchers to monitor the complex situation of the academic work.

1.2. Findings and Results

Some of those information are available on-line. They enable AlmaLaurea to analyse graduates' profile and employment condition. These analyses are so in-depth that they are gradually including an increasing number of information, even the degree course. The real "information heritage" of this huge data collection is stored in the central DW, whose extension and consolidation is still in an ongoing phase. For example, activities like examining the characteristics of the CVs which are frequently read by companies and analysing how do these characteristics vary on the basis of the different kinds of companies which read them are made possible by this information centre.

The aim of this work is, therefore, to deal with the issue of the technologies and processes involved in the creation of the data warehouse. Some examples of the results obtained are also given together with the guidelines for the consultation of the huge quantity of information available in English in the web site of AlmaLaurea.

2. The AlmaLaurea Interuniversity Consortium

AlmaLaurea is an innovative service that puts graduate CVs and résumés online, thus serving as a meeting point between graduates, universities and companies. Founded in 1994 on the initiative of the Statistical Observatory run by the University of Bologna, AlmaLaurea has grown exponentially and today covers every year 70% of Italian graduates.

Managed by a consortium of Italian universities with the support of the Ministry of Education, University and Research, AlmaLaurea was set up with the aim of putting businesses and graduates in contact and establishing itself as a reference point within the university system for the subjects involved (scholars, operators, etc.) in university education, employment and the development of young people in general.

Every year AlmaLaurea publishes two analyses on Italian graduates:

Annual Graduate Profile Report.

The Report examines all the graduates of the year, considering their characteristics and performances in the light of a multitude of variables including age at graduation, continuity of studies and attendance, parents' education, social background, study abroad, apprenticeships or internships, foreign languages and IT skills, etc.

Annual Report on the Occupational Conditions of Graduates.

The Report provides in-depth information about the employment conditions of young graduates at one, three, and five years from completion of studies, the prospects of the labour market and the relationships between university studies and employment opportunities.

These two reports are available in both paper and digital version [5, 6]

3. Overview of available Data Sources

The central data warehouse of Almalaurea contains information originating from many different data sources and from different subjects. The principal source is the relational database of AlmaLaurea, which manages and helps personnel in the process of verifying, validating and cleaning data, which come from two subjects:

- The universities taking part in the Consortium which -besides population data regarding the graduate community- send data concerning the title achieved, the university career, the thesis which has been carried out and previous educational qualifications (including the high school diploma)
- The graduate community which, through the filling in of a detailed on-line questionnaire, provide data pertinent to linguistic and computer related expertise, possible working experience during studies and type of employment which is sought. Besides this, they provide an accurate assessment of the university experience which has been carried out

After that the aforementioned data have been checked and cleaned [1], they lead to the creation of the CV of the graduate and goes to augment anonymously, the central DW. If graduates give their consent, their CVs are published on-line and rendered available for consultation to any company, so as to favour their access to the labour market. Starting from the moment in which the CV is made available to companies, graduates can modify it, enriching it with information on any future post-graduation training courses and on any future work experience.

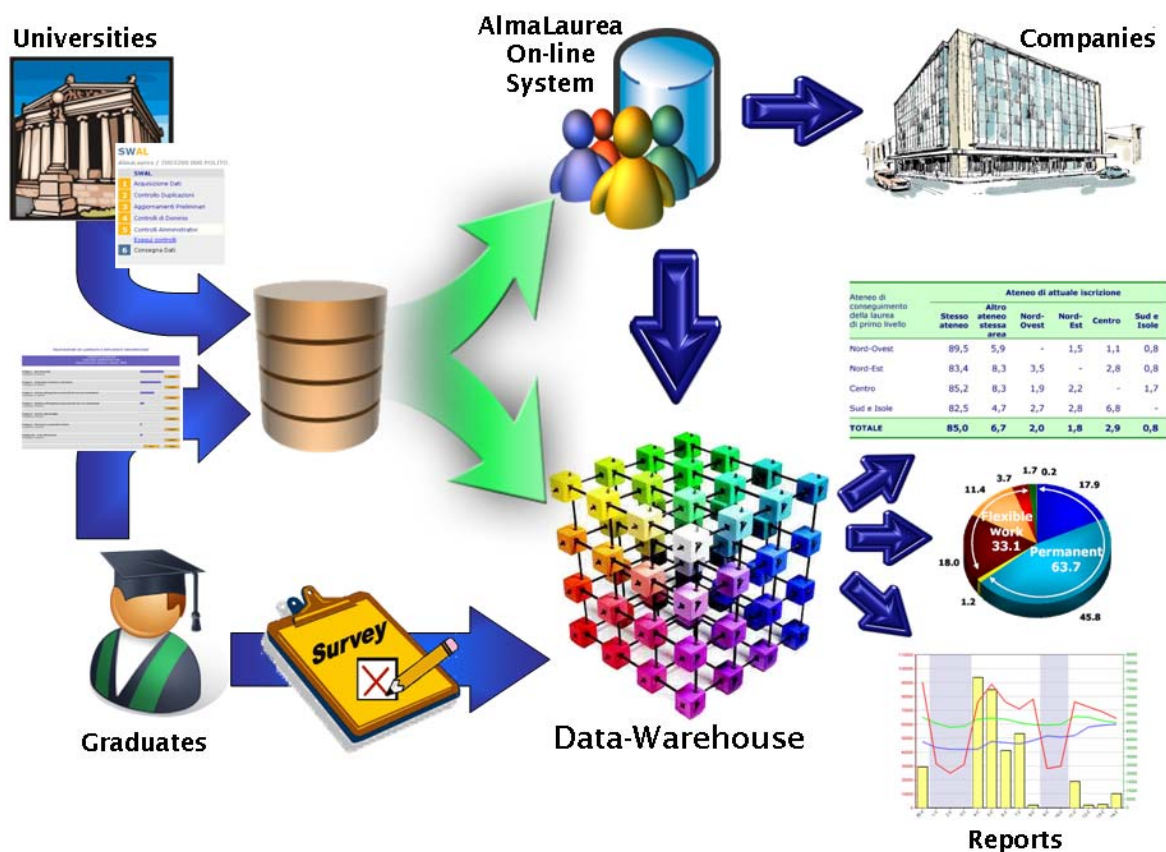
Besides this data provided by universities and by the graduate community, the data-warehouse is enriched by data provided by various interviews regarding the employment conditions, which are conducted in a systematic way in the years following the degree completion [2,5]. These interviews are conducted through CAWI (Computer Aided Web Interview) and CATI (Computer Aided Telephone Interview) methodologies.

Other interesting information sources that are conveyed in the DW derive from activities related to the on-line CV done by companies and graduates themselves. Among these we can include data originating from the distribution of CVs to companies, from the publishing of work offers on the notice-board of the website and from the spontaneous application of graduates to these offers.

Data present in the data-warehouse are examined and certified by a team of experts in statistical sciences collaborating with a team of programmers in order to create a set of data-marts and produce reports.

The following report material is made available, with different levels of details, to universities, to the Ministry of university, to the press and to all the visitors of the AlmaLaurea website.

Over the years the number of reports produced and of variables analyzed has grown considerably, in order to satisfy the demands of very different subjects requiring statistics with very different levels of granularity. Indeed, while the press is interested only in comparisons between universities or, at the most, between various disciplinary groups, the governing bodies of universities are, for the most part, interested in analyzing and comparing the trends of the parameters between faculties and between courses present in their university.



Picture 1: Data Sources of AlmaLaurea DataWarehouse

To give an idea of the number of the different analyzable combinations one can consider that the cube of the analysis of the graduates of the solar year of 2008 from 49 different universities taking part in the consortium (about 190.000 graduates) consists of 13 dimensions and, if all significant aggregations were calculated, one would have more than 6.000.000, each of which would count more than 100 calculated members.

The same kind of analysis and report is done yearly to analyze the employment condition of graduates one, three and five years on after graduation. The acquisition of the data necessary to this kind of analysis is done through CAWI and CATI surveys.

Also in this case reports having different levels of granularity - up to the level of the single degree courses-are produced. The possibility to supply significative statistics with a very fine level of granularity is guaranteed by the high number of graduates that we try to contact for the survey and also by the fact that the answer rate is always over 70% , with a maximum of 95% for the cohorts that are the most willing to answer.

In addition to these data-marts, which can also be consulted publicly, there are another ones which are currently available only to the internal researchers. As for the reports - which are still being elaborated and validated- they will be available to the universities which will apply for them through an extranet system. Through these reports it will be possible to analyse several factors related to the quality of the human capital produced by universities, such as:

- The sectors which are mostly required by companies in a given period
- The possibility to be selected on the basis of the age/ age at graduation
- The importance of factors such as foreign languages skills and diplomas certifying them.
- The characteristics of the CVs which are the most read by companies and their dependence from factors such as: the company economical sector, the main office and the dimension of the company.

4. Privacy issues

Processing this huge quantity of data made it necessary to elaborate and realise a project for the protection of the users' privacy -in particular their personal data- within the meaning of Article 13 of the Italian Personal Data Protection Code [4].

While filling in the questionnaire, graduates must tick off the box related to the authorisation to process personal data for the purpose of selecting personnel and give their consent for the publication of their CVs in our on-line databank. In addition to the check on the website, it is also necessary that graduates print the receipt which certifies that they filled in the questionnaire, sign up the authorisation to process personal data and deliver it personally to the student registry of their faculty.

Graduates can ask to be removed form the databank whenever they want. They just have to make an application to the AlmaLaurea help desk with a copy of the identity card in attachment, which certifies that the application was really made by the person concerned. The effect of this application is the cancellation from the on-line databank of their CVs and of all the data which have been supplied by graduates both through the questionnaire and through any updating made to the CVs in the website.

If graduates are not interested even in being contacted for interviews and/or other initiatives, they can make an application to the AlmaLaurea help desk.

The CAWI and CATI surveys are anonymous and the record with answers given does not contain personal data and/or any other information which may lead to associate the record to the person who gave the answers. Every record of answers has an univocal identifier, which is not used for purposes other than connecting the record to a set of statistic variables (e.g., university, faculty, degree course, degree course grouping, gender, duration of studies, age, graduation mark and so on) with the aim of elaborating them.

All the reports done and made available are anonymous. To avoid the isolation of the answers given by the single subjects we chose, in case of disaggregation into very small cohorts, to do not display the disaggregated data produced by a cohort composed by no more than 5 units.

5. Obtained Results

The availability of this huge quantity of data and reports allows us to meet the requirements of different subjects, such as: rectors, deans, presidents of degree courses, students and journalists. While the academic governing bodies are mostly interested in in-depth analysis concerning several factors, the requirements of the press are definitely different: analysis of the current trend of macro variables are required, since the press needs news having a significant impact.

5.1 The effects of the economic crisis in the graduates labour market

The existence of the data-warehouse and the consolidation of all the data it contains has also enabled us to analyse precisely the effects of the current economic crisis in the labour market. The last survey on graduates' economic condition one year on from graduation has been carried out on the basis of interviews made in October 2008. It has been presented to the public in February 2009 after a phase of validation of data and realisation of data-marts. When analysing the data of this survey, the effects of the economic crisis on the labour market seemed to be very mild. However, the perception is totally different.

By crossing the data of the previous data-mart with those related to the distribution of CVs to the companies for the purpose of selecting personnel, the effects of the economic crisis have arisen completely. These effects became visible because during the first 3 months of 2009 the number of CVs supplied to companies has fallen up to 35 % compared to the first 3 months of 2008.

This kind of research highlight the importance of a centralised data-warehouse for the analysis of the labour market trend. The co-existence of data-marts based on different sources allows us both to carry out a study with a fine level of granularity and to analyse - in almost real time-the trends of some macro-variables related to the entry into the labour market.

5.2 Explore employment surveys' results in the English Site

The results of the yearly surveys on graduates' profile [5] and those on the employment conditions [6] are available and can be freely consulted on the website www.almalaurea.net, which is written in English.

The tool for the consultation of the surveys puts an interface at the user's disposal. This interface has been simplified for the choice of dimensions and - on their inside- of the hierarchical level on which the operations of slicing and dicing on the cube are to be done in order to choose the characteristics of the cohort to be analysed. The mask has also another dropdown list that is used for choosing the dimension on which the user wants to disaggregate in order to compare different cohorts. The report that are publicly available allow the user to select cohorts up to the level degree course grouping/ statistic group, whereas the registered users can even analyse and compare data up to the level of single degree courses.

The mask for the selection of the cohort to be analysed always shows to the user the number of graduates who are part of the selected cohort. In this way it supplies a parameter which allows the users to evaluate the importance of the data that they are going to analyse.



Graduates' employment conditions

Selection of cohort [1600 graduates]

year of degree
2008

type of degree course [4]:
all

years since graduation [3]:
1

University [1]:
Bologna

Faculty [23]:
Engineering - Bologna

degree subject grouping [2]:
all

Comparison columns:

Have you selected the relevant cohort? Do you wish to break it down? If so, please choose your variable
type of degree course

reset **view**

Picture 2: Mask for cohort selection

By clicking on the button *view*, a query is made to the cube and the user receives a report having so many columns as the members of the dimension that has been chosen for disaggregation, a column with the totals and about one hundred rows which are grouped into 10 sections and the values of the variables which are calculated on the selected portion of the cube. The sections are the following ones:

- Investigated cohort
- Post-degree studies
- Employment conditions
- Access to the labour market
- Characteristics of the current work
- Characteristics of the company
- Earnings
- Application and demand for a degree within the current job
- Effectiveness of the degree within the current job
- Job search

For example, a part of the report (3 sections) on the employment conditions of the graduates of the Engineering faculty of the University of Bologna one year on from graduation, who have been disaggregated on the basis of the kind of degree course, so that a comparison between the degree courses set up after the application of the directives of the Bologna Process [14], [15] and the pre-existent courses (old second cycle degree) can be made.

1. Investigated cohort	Selected cohort	Selected Cohort (disaggregated by type of degree course)			
		first cycle degree	European second cycle degree	second cycle degree	old second cycle degree
Number of graduates	1.600	920	100	472	108
Number of interviewees	1.490	855	86	448	101
Response rate	93,1	92,9	86,0	94,9	93,5
Gender of the interviewees (%)					
Male	73,9	76,4	44,0	74,8	76,9
Female	26,1	23,6	56,0	25,2	23,1
Age at graduation (average)	24,9	24,0	26,9	25,4	28,9
Graduation mark out of 110 (average)	100,4	96,8	104,3	107,2	97,0
Duration of studies in years (average)	4,4	4,3	7,4	2,7	9,6

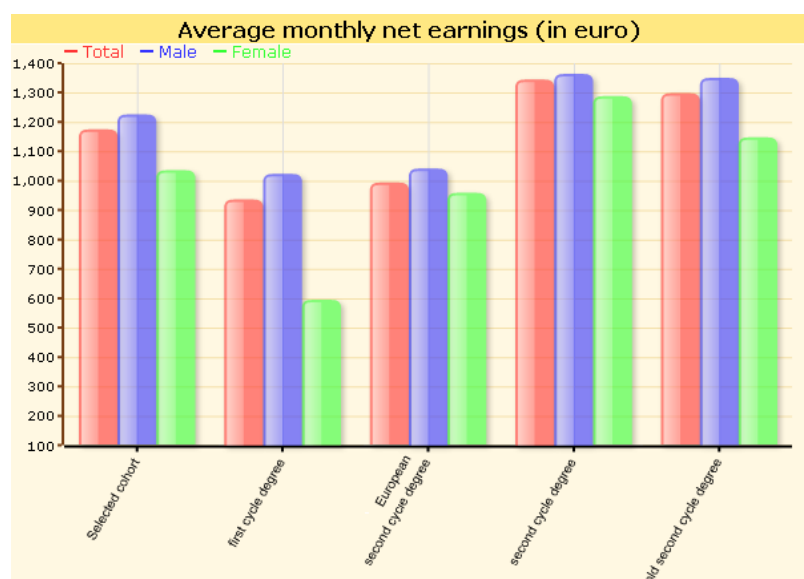
3. Employment condition	Selected cohort	Selected Cohort (disaggregated by type of degree course)			
		first cycle degree	European second cycle degree	second cycle degree	old second cycle degree
Employment condition (%)					
is working	52,3	29,2	87,2	81,9	87,1
doesn't work and doesn't look for a job	39,9	63,4	5,8	10,3	2,0
doesn't work, but looks for a job	7,7	7,4	7,0	7,8	10,9
Percentage of graduates who is not looking for a job, but is attending a degree course or is doing a training activity	38,1	62,3	2,3	6,9	1,0
Percentage of those who work, divided per gender					
Male	52,1	30,6	86,5	82,6	85,7
Female	52,9	24,9	87,8	80,0	91,7
Post-graduation working experiences (%)					
is not working but has worked after graduation	8,3	9,8	4,7	6,7	5,9
has never worked after graduation	39,3	60,9	8,1	11,4	6,9
Employment rate (ISTAT def. Workforce)	58,5	31,7	91,9	95,3	93,1
Unemployment rate (ISTAT def. Workforce)	5,2	10,0	4,8	2,3	4,1

7. Earnings	Selected cohort	Selected Cohort (disaggregated by type of degree course)			
		first cycle degree	European second cycle degree	second cycle degree	old second cycle degree
Average monthly net earnings (in euro)					
Male	1.226	1.024	1.042	1.364	1.350
Female	1.036	595	959	1.288	1.148
Total	1.175	937	994	1.345	1.298

Picture 3: Employment Conditions of Engineering Graduates of Bologna

In case the number of columns is higher than 5 units, a special function allows the user to put in increasing or decreasing order the columns on the basis of the values of a row just by clicking with the mouse on the selected row.

The system also permits to make a graphic representation of the data by clicking on the symbol that can be found on the columns in bolt. For example, the earnings section:



Picture 4: Avg monthly net earnings by kind of degree course

6 Conclusions

The data that have been collected since 1994 by the AlmaLaurea Consortium enable us to analyse the evolution of the Italian academic system. The Consortium yearly obtains a picture of graduates after degree completion, it also follows them during the years on after graduation by monitoring and studying their evolution both at the educational and occupational level.

Using the databank also as an intermediation instrument on the labour market between companies and graduates produces positive effects on employability [12] but it has also the effect of keeping the contents of the databank up-to-date. Having up-to-date e-mail addresses and telephone numbers allows us to obtain response rates over 85% in case of interviews, even if they have been carried out several years on after graduation.

The evolution of the Italian academic system toward the model established within the Bologna Process has been realised. However, no evaluation of the actual benefits is possible without an objective approach that can be realised only through a complete, up-to-date, systematic and well-timed documentation, like the one that is collected and elaborated by AlmaLaurea.

7 REFERENCES

- [1] A. Leone, L. Cancellieri et al (2008) *Merging graduates data from remote and non-homogeneous information systems* EUNIS 2008 Conference Proceedings. <http://eunis.dk/papers/p94.pdf>
- [2] A. Leone et al (2008) *AlmaLaurea Surveys* EUNIS 2008 Conference Proceedings. <http://eunis.dk/papers/p63.pdf>
- [3] A. Cammelli et al. (2008) *AlmaLaurea.net: the database for International graduates* EUNIS 2008 Conference Proceedings <http://eunis.dk/papers/p59.pdf>
- [4] Guarantor of Personal Data Protection (2003): *Standard Italian and community rules on personal data protection* <http://www.garanteprivacy.it/garante/navig/jsp/index.jsp?folderpath=Normativa> (in Italian)
- [5] AlmaLaurea (2009). *Graduate Profile - 2009 Report, Consorzio Interuniversitario AlmaLaurea*, Bologna (see also previous works 1994 - 2008), <http://www.almalaurea.it/eng/universita/profilo/>
- [6] AlmaLaurea (2009). *Employment Conditions of Graduates - 2009 Report, Consorzio Interuniversitario AlmaLaurea*, Bologna (see also previous reports 1996-2008), <http://www.almalaurea.it/eng/universita/occupazione/>
- [7] A. Cammelli et al (2007). *ALMALAUREA: from University courses evaluation to graduates placement*. EUNIS 2007 Conference Proceedings. Abstract online: <http://www.eunis.org/events/congresses/eunis2007/CD/pdf/abstracts/p111.pdf>
- [8] MIUR (2000). *Ministerial Decree of 24 October 2000*. Available from: http://www.miur.it/atti/2000/dm001004_01.htm
- [9] Ralph Kimball, Joe Caserta (2004). *The Data Warehouse ETL Toolkit: Practical Techniques for Extracting, Cleaning, Conforming, and Delivering Data*. John Wiley & Sons
- [10] Ralph Kimball and Margy Ross (2002) *The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling, 2nd Edition*. John Wiley & Sons
- [11] Ralph Kimball, Margy Ross, Warren Thornthwaite, Joy Mundy, Bob Becker (2008) *The Data Warehouse Lifecycle Toolkit, 2nd Edition: Practical Techniques for Building Data Warehouse and Business Intelligence Systems* John Wiley & Sons
- [12] Bagues, M., Sylos-Labini M. (2007) *Do On-Line Labor Market Intermediaries Matter? The impact of AlmaLaurea on University-to-Work Transition* Presented in Boston (MA) "Labour Market Intermediation Conference", May 17th-18th 2007. Will appear on National Bureau of Economics Research Working paper series

[13] Ralph Kimball, Warren Thornthwaite and Joy Mundy (2006) *The Microsoft Data Warehouse Toolkit: With SQL Server 2005 and the Microsoft Business Intelligence Toolset* John Wiley & Sons

[14] MIUR (1999). DM (Ministerial Decree) 509/1999, *Regolamento recante norme concernenti l'autonomia didattica degli atenei*, from:

http://www.miur.it/0006Menu_C/0012Docume/0098Normat/2088Regola.htm

[15] MIUR (2004). DM (Ministerial Decree) 270/2004, *Modifiche al regolamento recante norme concernenti l'autonomia didattica degli atenei, approvato con decreto del Ministro dell'università e della ricerca scientifica e tecnologica 3 novembre 1999, n. 509.*, from:

http://www.miur.it/0006Menu_C/0012Docume/0098Normat/4640Modifi_cf2.htm