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

Alberto Leone, 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

Universities who joined AlmaLaurea in the last year
AlmaLaurea

Universities → AlmaLaurea On-line System → Companies

Graduates → Data-Warehouse → Reports

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 studying 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
## 1. Personal details

<table>
<thead>
<tr>
<th>Parents' educational qualification</th>
<th>2003</th>
<th>2004</th>
</tr>
</thead>
<tbody>
<tr>
<td>both parents are degree holders</td>
<td>7.4</td>
<td>7.8</td>
</tr>
<tr>
<td>only one parent is a graduated</td>
<td>13.2</td>
<td>14.2</td>
</tr>
<tr>
<td>secondary school-leaving certificate</td>
<td>40.3</td>
<td>43.0</td>
</tr>
<tr>
<td>lower educational qualification or no educational qualification</td>
<td>35.9</td>
<td>33.0</td>
</tr>
</tbody>
</table>

### Social class (%)

<table>
<thead>
<tr>
<th>Social class (%)</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
</tr>
</thead>
<tbody>
<tr>
<td>other region</td>
<td>24.7</td>
<td>23.7</td>
<td>22.5</td>
<td>21.8</td>
<td>20.7</td>
<td>21.0</td>
<td>20.8</td>
</tr>
<tr>
<td>abroad</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.5</td>
<td>0.6</td>
<td>0.5</td>
<td>0.6</td>
</tr>
</tbody>
</table>

## 2. Social background

### Selected cohort (compared by graduation year)

<table>
<thead>
<tr>
<th>Parents' educational qualification</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
</tr>
</thead>
<tbody>
<tr>
<td>both parents are degree holders</td>
<td>7.4</td>
<td>7.8</td>
<td>8.6</td>
<td>8.6</td>
<td>8.4</td>
<td>8.3</td>
<td>8.3</td>
</tr>
<tr>
<td>only one parent is a graduated</td>
<td>13.2</td>
<td>14.2</td>
<td>14.7</td>
<td>14.9</td>
<td>14.8</td>
<td>14.6</td>
<td>14.9</td>
</tr>
<tr>
<td>secondary school-leaving certificate</td>
<td>40.3</td>
<td>43.0</td>
<td>44.1</td>
<td>44.7</td>
<td>45.3</td>
<td>45.1</td>
<td>46.7</td>
</tr>
<tr>
<td>lower educational qualification or no educational qualification</td>
<td>35.9</td>
<td>33.0</td>
<td>30.5</td>
<td>29.4</td>
<td>29.2</td>
<td>28.6</td>
<td>27.9</td>
</tr>
</tbody>
</table>

### Social class (%)

<table>
<thead>
<tr>
<th>Social class (%)</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
</tr>
</thead>
<tbody>
<tr>
<td>middle class</td>
<td>54.7</td>
<td>19.3</td>
<td>20.3</td>
<td>20.6</td>
<td>20.3</td>
<td>19.6</td>
<td>19.4</td>
</tr>
<tr>
<td>clerical middle class</td>
<td>14.5</td>
<td>31.0</td>
<td>30.9</td>
<td>30.6</td>
<td>30.6</td>
<td>30.2</td>
<td>29.3</td>
</tr>
<tr>
<td>lower middle class</td>
<td>22.4</td>
<td>23.1</td>
<td>22.2</td>
<td>22.2</td>
<td>22.1</td>
<td>22.0</td>
<td>22.3</td>
</tr>
<tr>
<td>working class</td>
<td>3.9</td>
<td>22.5</td>
<td>21.9</td>
<td>22.8</td>
<td>23.2</td>
<td>24.5</td>
<td>25.1</td>
</tr>
</tbody>
</table>
Transparency info

- Detailed info about (requested by Ministry for orientation purposes)
  - Employability
  - Degree effectiveness
  - Assessment of university experience
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 heterogeneous 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
  - datasets
  - extract & transform tools
  - staging (cleaning area)
- browser
  - custom reports
  - report server
- data sources
  - sql
  - data mart (mdx)
  - data warehouse (data and metadata)

Report service
Analysis service
Sql server
Sql server integration services
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

- Map the relational schema into Analysis Service star schema
- Define dimensions
- Define measures
- Define cubes
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

<table>
<thead>
<tr>
<th></th>
<th>anonymous</th>
<th>researcher</th>
<th>University staff</th>
<th>Ministry staff</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dynamic report (pre-built)</td>
<td>YES (*)</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Static report</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Report builder</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
<td>NO</td>
</tr>
</tbody>
</table>
## Roles, reports: users vs data sets

<table>
<thead>
<tr>
<th>user</th>
<th>anonymous</th>
<th>researcher</th>
<th>University staff (*)</th>
<th>Ministry staff</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graduation</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Working Conditions</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Company Activities</td>
<td>NO</td>
<td>YES</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graduates activities</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

(*) limited to single university data
Summary

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

<table>
<thead>
<tr>
<th>Gruppo Statistico</th>
<th>Solo uno dei due con diploma superiore</th>
<th>Solo uno dei due con laurea</th>
<th>Titoli inferiori o nessun titolo</th>
<th>Tutti e due con diploma superiore</th>
<th>Tutti e due con Laurea</th>
<th>Undefined</th>
<th>Totale</th>
</tr>
</thead>
<tbody>
<tr>
<td>agrario</td>
<td>0.40</td>
<td>0.41</td>
<td>0.53</td>
<td>0.40</td>
<td>0.31</td>
<td>0.72</td>
<td>0.46</td>
</tr>
<tr>
<td>architettura</td>
<td>0.47</td>
<td>0.42</td>
<td>0.61</td>
<td>0.45</td>
<td>0.36</td>
<td>0.81</td>
<td>0.51</td>
</tr>
<tr>
<td>chimico-farmaceutico</td>
<td>0.35</td>
<td>0.39</td>
<td>0.49</td>
<td>0.34</td>
<td>0.32</td>
<td>0.55</td>
<td>0.40</td>
</tr>
<tr>
<td>economico-statistico</td>
<td>0.41</td>
<td>0.38</td>
<td>0.54</td>
<td>0.38</td>
<td>0.30</td>
<td>0.79</td>
<td>0.46</td>
</tr>
<tr>
<td>educazione fisica</td>
<td>0.33</td>
<td>0.31</td>
<td>0.33</td>
<td>0.30</td>
<td>0.33</td>
<td>0.30</td>
<td>0.32</td>
</tr>
<tr>
<td>geo-biologico</td>
<td>0.35</td>
<td>0.32</td>
<td>0.48</td>
<td>0.33</td>
<td>0.30</td>
<td>0.53</td>
<td>0.39</td>
</tr>
<tr>
<td>giuridico</td>
<td>0.67</td>
<td>0.56</td>
<td>0.86</td>
<td>0.63</td>
<td>0.41</td>
<td>1.04</td>
<td>0.73</td>
</tr>
<tr>
<td>ingegneria</td>
<td>0.45</td>
<td>0.40</td>
<td>0.60</td>
<td>0.43</td>
<td>0.34</td>
<td>0.74</td>
<td>0.47</td>
</tr>
<tr>
<td>insegnamento</td>
<td>0.46</td>
<td>0.44</td>
<td>0.53</td>
<td>0.42</td>
<td>0.49</td>
<td>0.54</td>
<td>0.49</td>
</tr>
<tr>
<td>letterario</td>
<td>0.52</td>
<td>0.47</td>
<td>0.72</td>
<td>0.48</td>
<td>0.39</td>
<td>0.81</td>
<td>0.57</td>
</tr>
<tr>
<td>linguistico</td>
<td>0.40</td>
<td>0.38</td>
<td>0.59</td>
<td>0.38</td>
<td>0.32</td>
<td>0.80</td>
<td>0.48</td>
</tr>
<tr>
<td>medico</td>
<td>0.16</td>
<td>0.21</td>
<td>0.18</td>
<td>0.18</td>
<td>0.15</td>
<td>0.27</td>
<td>0.19</td>
</tr>
<tr>
<td>politico-sociale</td>
<td>0.38</td>
<td>0.38</td>
<td>0.43</td>
<td>0.36</td>
<td>0.33</td>
<td>0.50</td>
<td>0.40</td>
</tr>
<tr>
<td>psicologico</td>
<td>0.34</td>
<td>0.29</td>
<td>0.43</td>
<td>0.30</td>
<td>0.27</td>
<td>0.46</td>
<td>0.35</td>
</tr>
<tr>
<td>scientifico</td>
<td>0.46</td>
<td>0.44</td>
<td>0.62</td>
<td>0.40</td>
<td>0.27</td>
<td>0.63</td>
<td>0.47</td>
</tr>
<tr>
<td><strong>Totale</strong></td>
<td><strong>0.41</strong></td>
<td><strong>0.39</strong></td>
<td><strong>0.52</strong></td>
<td><strong>0.39</strong></td>
<td><strong>0.32</strong></td>
<td><strong>0.65</strong></td>
<td><strong>0.45</strong></td>
</tr>
</tbody>
</table>

All the universities
**Sample one: delay in degree completion times study**

### Profilo - Indice di Ritardo

Gruppo disciplinare - Titoli dei genitori

<table>
<thead>
<tr>
<th>Gruppo Statistico</th>
<th>Solo uno dei due con diploma superiore</th>
<th>Solo uno dei due con laurea</th>
<th>Titoli inferiori o nessun titolo</th>
<th>Tutti e due con diploma superiore</th>
<th>Tutti e due con Laurea</th>
<th>Undefined</th>
<th>Totale</th>
</tr>
</thead>
<tbody>
<tr>
<td>agrario</td>
<td>0.21</td>
<td>94</td>
<td>0.26</td>
<td>67</td>
<td>0.37</td>
<td>90</td>
<td>0.32</td>
</tr>
<tr>
<td>architettura</td>
<td>0.31</td>
<td>58</td>
<td>0.38</td>
<td>41</td>
<td>0.38</td>
<td>67</td>
<td>0.30</td>
</tr>
<tr>
<td>chimico-farmaceutico</td>
<td>0.28</td>
<td>63</td>
<td>0.33</td>
<td>102</td>
<td>0.47</td>
<td>95</td>
<td>0.31</td>
</tr>
<tr>
<td>economico-statistico</td>
<td>0.35</td>
<td>454</td>
<td>0.31</td>
<td>382</td>
<td>0.39</td>
<td>500</td>
<td>0.30</td>
</tr>
<tr>
<td>educazione fisica</td>
<td>0.37</td>
<td>47</td>
<td>0.24</td>
<td>32</td>
<td>0.38</td>
<td>47</td>
<td>0.31</td>
</tr>
<tr>
<td>geo-biologico</td>
<td>0.22</td>
<td>119</td>
<td>0.25</td>
<td>131</td>
<td>0.31</td>
<td>96</td>
<td>0.22</td>
</tr>
<tr>
<td>giuridico</td>
<td>0.66</td>
<td>163</td>
<td>0.48</td>
<td>225</td>
<td>0.80</td>
<td>173</td>
<td>0.59</td>
</tr>
<tr>
<td>ingegneria</td>
<td>0.39</td>
<td>372</td>
<td>0.35</td>
<td>370</td>
<td>0.51</td>
<td>364</td>
<td>0.40</td>
</tr>
<tr>
<td>insegnamento</td>
<td>0.30</td>
<td>199</td>
<td>0.49</td>
<td>102</td>
<td>0.38</td>
<td>307</td>
<td>0.31</td>
</tr>
<tr>
<td>letterario</td>
<td>0.46</td>
<td>424</td>
<td>0.40</td>
<td>459</td>
<td>0.56</td>
<td>394</td>
<td>0.40</td>
</tr>
<tr>
<td>linguistico</td>
<td>0.31</td>
<td>169</td>
<td>0.28</td>
<td>125</td>
<td>0.50</td>
<td>163</td>
<td>0.30</td>
</tr>
<tr>
<td>medico</td>
<td>0.17</td>
<td>195</td>
<td>0.16</td>
<td>168</td>
<td>0.20</td>
<td>242</td>
<td>0.17</td>
</tr>
<tr>
<td>politico-sociale</td>
<td>0.30</td>
<td>442</td>
<td>0.26</td>
<td>422</td>
<td>0.47</td>
<td>429</td>
<td>0.30</td>
</tr>
<tr>
<td>psicologico</td>
<td>0.12</td>
<td>116</td>
<td>0.19</td>
<td>73</td>
<td>0.28</td>
<td>117</td>
<td>0.15</td>
</tr>
<tr>
<td>scientifico</td>
<td>0.35</td>
<td>106</td>
<td>0.46</td>
<td>113</td>
<td>0.53</td>
<td>132</td>
<td>0.29</td>
</tr>
</tbody>
</table>

**Totale**

|             | 0.34 | 3021 | 0.33 | 2812 | 0.45 | 3213 | 0.33 | 3398 | 0.30 | 1890 | 0.69 | 1588 | 0.39 | 15922 |

Bologna university
Sample two: relationship between field of study and company type

<table>
<thead>
<tr>
<th>Tipo Azienda</th>
<th>Assicurazioni</th>
<th>Associazioni</th>
<th>Azienda Sconosciuta</th>
<th>Aziende</th>
<th>Banche/Finanziarie</th>
<th>Enti</th>
<th>Enti di formazione</th>
<th>Istituto Superiore</th>
<th>Liberi professionisti</th>
<th>Ricerca personale</th>
<th>Università</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agraria</td>
<td>14</td>
<td>2</td>
<td>20</td>
<td>192</td>
<td>9</td>
<td>869</td>
<td>35</td>
<td>86</td>
<td>15</td>
<td>5</td>
<td>70</td>
<td>263</td>
</tr>
<tr>
<td>Architettura</td>
<td>12</td>
<td>2</td>
<td>14</td>
<td>96</td>
<td>1</td>
<td>114</td>
<td>15</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>32</td>
<td>38</td>
</tr>
<tr>
<td>Chimica industriale</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>114</td>
<td>15</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>32</td>
<td>38</td>
</tr>
<tr>
<td>Conservazione dei beni culturali</td>
<td>10</td>
<td>19</td>
<td>4</td>
<td>73</td>
<td>6</td>
<td>202</td>
<td>166</td>
<td>31</td>
<td>595</td>
<td>31</td>
<td>595</td>
<td>31</td>
</tr>
<tr>
<td>Economia - sede di Bologna</td>
<td>29</td>
<td>3</td>
<td>7</td>
<td>189</td>
<td>167</td>
<td>5</td>
<td>626</td>
<td>1</td>
<td>381</td>
<td>69</td>
<td>1477</td>
<td>1477</td>
</tr>
<tr>
<td>scienze dell'economia (64/5)</td>
<td>economia aziendale</td>
<td>33 1.07 %</td>
<td>8 0.26 %</td>
<td>13 0.41 %</td>
<td>139 4.61 %</td>
<td>259 8.40 %</td>
<td>24 0.78 %</td>
<td>1179 38.24 %</td>
<td>6 0.19 %</td>
<td>0.00 %</td>
<td>1379 44.73 %</td>
<td>52 1.69 %</td>
</tr>
<tr>
<td>scienze dell'economia e della gestione aziendale (17)</td>
<td>economia aziendale</td>
<td>57 15.57 %</td>
<td>2 0.55 %</td>
<td>0.00 %</td>
<td>13 3.55 %</td>
<td>22 8.74 %</td>
<td>1 0.27 %</td>
<td>94 25.68 %</td>
<td>0.00 %</td>
<td>0.00 %</td>
<td>167 45.63 %</td>
<td>0.00 %</td>
</tr>
<tr>
<td>scienze economiche (28)</td>
<td>economia aziendale</td>
<td>3 0.84 %</td>
<td>1 0.28 %</td>
<td>1 0.28 %</td>
<td>24 6.72 %</td>
<td>67 18.77 %</td>
<td>0.00 %</td>
<td>198 55.46 %</td>
<td>0.00 %</td>
<td>0.00 %</td>
<td>47 13.17 %</td>
<td>16 4.48 %</td>
</tr>
<tr>
<td>scienze economiche per l'ambiente e la cultura (83/5)</td>
<td>economia aziendale</td>
<td>11 20.37 %</td>
<td>1 0.00 %</td>
<td>0.00 %</td>
<td>1 1.65 %</td>
<td>3 5.65 %</td>
<td>0.00 %</td>
<td>28 51.85 %</td>
<td>2 3.70 %</td>
<td>0.00 %</td>
<td>7 12.96 %</td>
<td>2 3.70 %</td>
</tr>
<tr>
<td>scienze economiche per l'ambiente e la cultura (83/5)</td>
<td>economia aziendale</td>
<td>42 26.7 %</td>
<td>1 0.06 %</td>
<td>4 0.25 %</td>
<td>99 6.29 %</td>
<td>134 8.52 %</td>
<td>12 0.76 %</td>
<td>886 55.45 %</td>
<td>1 0.06 %</td>
<td>0.00 %</td>
<td>361 23.92 %</td>
<td>31 1.97 %</td>
</tr>
<tr>
<td>scienze economiche per l'ambiente e la cultura (83/5)</td>
<td>economia aziendale</td>
<td>13 1.37 %</td>
<td>3 0.32 %</td>
<td>1 0.11 %</td>
<td>61 4.41 %</td>
<td>63 4.83 %</td>
<td>5 0.53 %</td>
<td>402 42.27 %</td>
<td>2 0.21 %</td>
<td>0.00 %</td>
<td>366 38.49 %</td>
<td>15 1.58 %</td>
</tr>
<tr>
<td>scienze economiche (28)</td>
<td>economia aziendale</td>
<td>159 15.10</td>
<td>337 578 42 2789 11 2327 116 6384</td>
<td>655 56 1942</td>
<td>48 5 379</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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
Thank you!

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