**Table2. Attribute Intersection with Integrated Schema**

|  |  |  |  |
| --- | --- | --- | --- |
| Class Name | Attributes Name | Datasets in which attribute is found | Conflict resolution strategies |
| company | (company)name | dataset 1,2,3 | LongestString, FavourSources, Voting |
| company,location | country | dataset 1,2,3,4 | Voting |
| company | industries | dataset 1,2,3 | Union |
| company | Sales/revenue | dataset 1,2, 3 | FavourSources |
| company | number Of Employees | dataset 2,3 | Average, Growth: Max |
| company | founding year/ date founded | dataset 2,3 | MostComplete (complete date),  MostComplete (sample), Fallback: Random |
| company | Assets/ [assets](http://dbpedia.org/ontology/assets) | dataset 1 ,2 | FavourSources |
| company | Market Value | dataset 1 | / |
| company | profit | dataset 1 | / |
| company | continent | dataset 1 | / |
| company | keyPeople/ leadership | dataset 2,3 | Union,  Intersection? |
| company,location | name | dataset 2,3,4 | Union + FavourSources |
| location | population total | dataset 4 | / |
| location | area total | dataset 4 | / |
| location | elevation | dataset 4 | / |

1, forbes file

2, dbpedia

3, freebase

4, location

Notes for report:

* Conflict resolution strategies: MostRecent would be good, but out of scope for this project

4 EXPLANATIONS WHY ENOUGH ENTITIES ARE LIKELY CONTAINED IN MULTIPLE DATASETS

Take 10 Samples of LES randomly and check the existence of company entities in DBpedia and Freebase

|  |  |  |
| --- | --- | --- |
| company in LES | company in DBPedia | company in Freebase |
| ALDERMORE GROUP PLC | http://dbpedia.org/resource/Aldermore | Aldermore |
| RICARDO | http://dbpedia.org/resource/Ricardo\_plc | RICARDO PLC |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |