**Table2. Attribute Intersection with Integrated Schema**

|  |  |  |
| --- | --- | --- |
| Class Name | Attributes Name | Datasets in which attribute is found |
| company | (company)name | dataset 1,2,3 |
| location | country | Dataset 1, 4 |
| company | industry | dataset 1,2,3 |
| company | Sales/revenue | dataset 1,2 |
| company | number Of Employees | dataset 2,3 |
| company | founding year/ date founded | dataset 2,3 |
| company | headqurters | dataset 2,3 |
| company | Sector/sectors | dataset 1,3 |
| company | Assets/ [assets](http://dbpedia.org/ontology/assets) | dataset 1 ,2 |
| company | Rank | dataset 1 |
| company | Market Value | dataset 1 |
| company | Forbes Webpage | dataset 1 |
| company | profit | dataset 1 |
| company | continent | dataset 1 |
| company | key People/ leadership | dataset 2,3 |
| location | name | dataset 4 |
| location | population total | dataset 4 |
| location | area total | dataset 4 |
| location | elevation | dataset 4 |

1,forbes file

2,dbpedia

3,freebase

4,location

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 |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |