



# KIM PLATFORM

## Knowledge Base Statistics

*A brief overview of the  
KIM knowledge base.*

## 1 Table of Contents

1	Table of Contents.....	2
2	Overview .....	3
3	Geographic Names.....	3
4	Organizations .....	4
5	Precise figures, as of build 0305 for KB: .....	4

## 2 Overview

KIM Ontology contains about 250 concepts and 40 properties grouped into three upper level concepts: *Entity*, *EntitySource* and *LexicalResource*.

The most important concept is *Entity*, which is further specialized into *Abstract*, *Happening* and *Object* concepts.

The knowledge base contains about 77,500 *Entities* with more than 110,000 aliases.

## 3 Geographic Names

Big part of the entities in the knowledge base are geographic features, imported from GNS (GEOnet Names Server) and other sources. The geographic data is organized into instances of *Location* and its subclasses, such as *Continents*, *GlobalRegions*, *Countries*, and so on, connected with *subRegionOf* property.

Currently the following types are imported in KB:

- *Country*
- *Province*
- *County*
- *CountryCapital*
- *LocalCapital*
- *City*
- *Ocean*
- *Sea*
- *Gulf*
- *OilField*
- *Monument*
- *Bridge*
- *Plateau*
- *Mountain*
- *MountainRange*
- *Plain*

The geographic data has about 50,000 *Locations* grouped into *Continents*, *GlobalRegions*, 282 *Countries*, all *CountryCapitals* and 4,700 *Cities* (including all the cities with population over 100,000).

Each location has several aliases (usually including English, French and sometimes the local transcription of the location), geographic coordinates (Long/Lat), the designator (DSG) and Unique Feature Index (UFI), according to GNS.

## 4 Organizations

The knowledge base contains 8,400 biggest world *Organizations*: 7,900 *Companies* (including more than 5,000 *PublicCompanies* with 5,500 management *Positions* within them), 140 *StockExchanges* with 'tradedOn' relations from *PublicCompanies*.

For many organizations there are 'locatedIn' relations to the corresponding countries, reflecting a very general notion of presence. In addition, the following attributes are kept per company:

- *URL/site*
- *short description*
- *Industry sector (flat list of about 20 industry sectors, with activeInSector relation)*
- *index on the stock exchange (for public companies)*
- *Sales last year*
- *Net income last year*
- *Number of employees*
- *key people and positions (e. g. CEO, President, Chairman, etc.)*

The company data came from various sources, mostly per-country lists of registered companies. Only the entities of global importance were imported. The company data is verified to contain all the publicly traded companies listed in Google directory ([http://directory.google.com/Top/Business/Major\\_Companies/Publicly\\_Traded/](http://directory.google.com/Top/Business/Major_Companies/Publicly_Traded/)), Hoovers Online (www.hoovers.com), and is currently being re-evaluated and enhanced with other famous companies, according to the classifications of Forbes (www.forbes.com), Fortune magazines (www.fortune.com), and the European business directory (www.europages.net).

## 5 Precise figures, as of build 0305 for KB:

### 5.1 Figure 1. Number of concepts

Concept	Count
Entities	77,561
Aliases of entities	110,308
Locations	49,348
Cities	4,720
Companies	7,906
Public companies	5,150
Key people (Positions)	5,500
Organizations (including companies)	8,365

## 5.2 Figure 2. Upper level of KIM Ontology

