



KIM PLATFORM

Knowledge Base Statistics

*A brief overview of the
KIM knowledge base.*

Table of Contents

Table of Contents	2
Overview	3
Geographic Names	3
Organizations	4
Precise figures for the KB, as of revision 12406:	4
1.1 Figure 1. Number of concepts 4	
1.2 Figure 2. Upper level of KIM Ontology 5	

Overview

KIM Ontology contains about 350 concepts and 195 properties. The most important are 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 44,800 *Entities* with more than 77,000 aliases.

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.

This is a short excerpt of location types, existing in the KB:

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

The geographic data has about 12,000 *Locations* grouped into *Continents*, *GlobalRegions*, 260 *Countries*, all *CountryCapitals* and 4,300 *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.

Organizations

The knowledge base contains 8,300 biggest world *Organizations*: 7,800 *Companies* (including more than 5,000 *PublicCompanies* with 5,700 management *Positions* within them).

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/), 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).

Precise figures for the KB, as of revision 12406:

1.1 Figure 1. Number of concepts

Concept	Count
Entities	44,857
Aliases of entities	77,030
Locations	12,424
Cities	4,299
Companies	7,862
Public companies	5,069
Key people (Positions)	5,753
Organizations (including companies)	8,348

1.2 Figure 2. Upper level of KIM Ontology

