The SIMBAD database

Marc Wenger

SIMBAD Content

A database collecting data for more than 5,000,000 astronomical objects

- Stars, galaxies, in general objects outside the solar system
- **Basic data** (coordinates, proper motions, radial velocities, parallaxes, spectral and morphological types, etc …)
- **Identifiers**. Cross identifications among many catalogs. 10500 different acronyms, 745 objects with more than 40 identifiers, 1 object with 103 identifiers
- **References.** 3,425,000 objects share 210,000 references providing 8,175,000 citations
- **Measurements**. 3300000 measurements in 35 catalogs for 2,000,000 objects
- **Notes**. For 47,000 objects
- **Links** to other services (VizieR, Heasarc, NED)
- **Annotations** from users

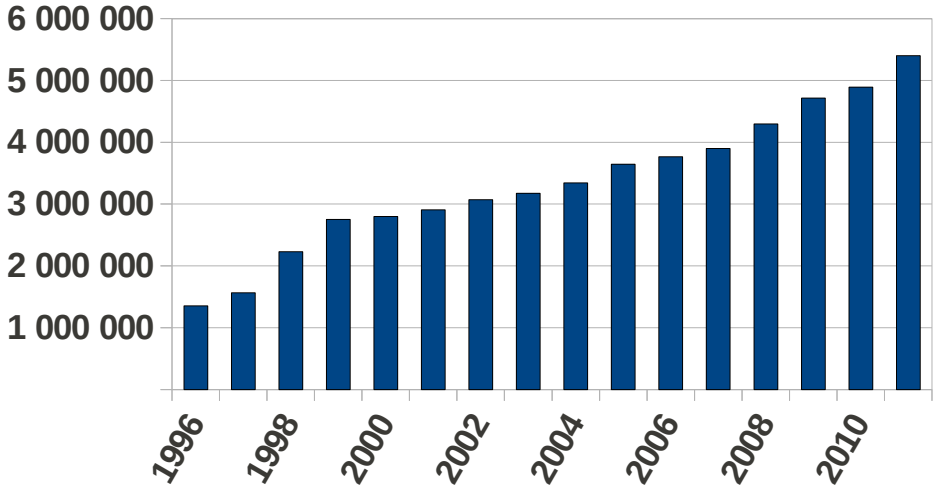
Simbad : a 40 year story

Four major versions :

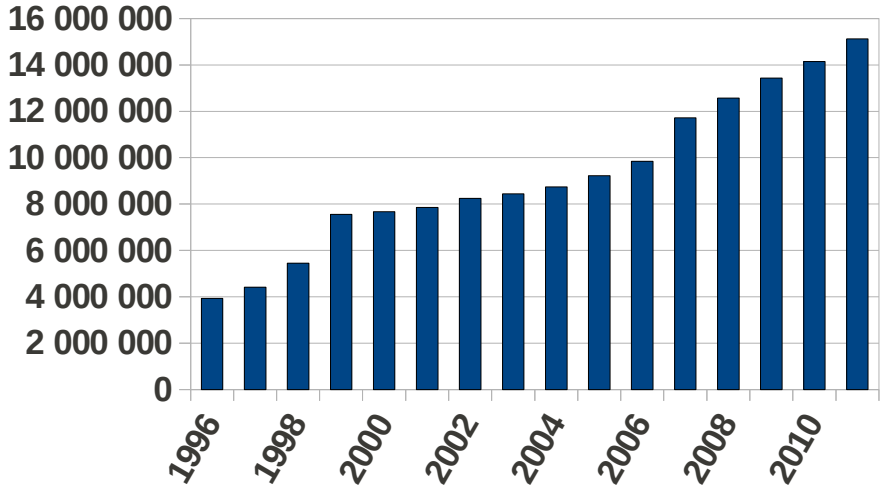
- Born with the CDS in **1972** as the CSI (Catalog of Stellar Identifications). Keypunched cards, IBM 360/65 in Meudon, access through dedicated lines for the French community
- Rewritten in **1981** on a Univac 1110 in Strasbourg. Access in command line mode though packet switching networks.
  In 1988, first usage of Internet in France across the Atlantic for demos in the USA (Lisa I and IAU meeting)
- New release in **1990** (C language) on Unix stations at the Observatory. Access through Internet. Graphical interface, Web. From 50 to 20,000 queries/day in 16 years.
- Today: Simbad4 since **Dec 2006** (Java language, Linux servers). 5,000,000 objects, 265,000 queries/day in 2010.

Simbad : content evolution 1996 - 2011

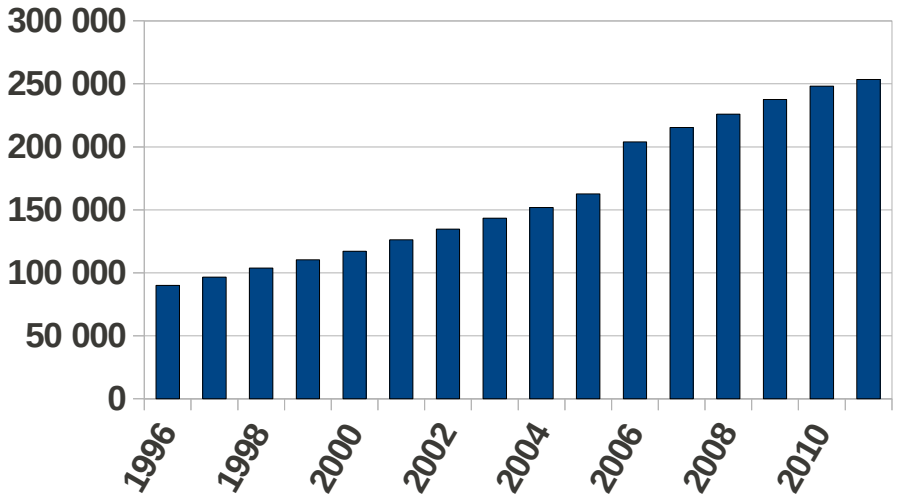
Objects



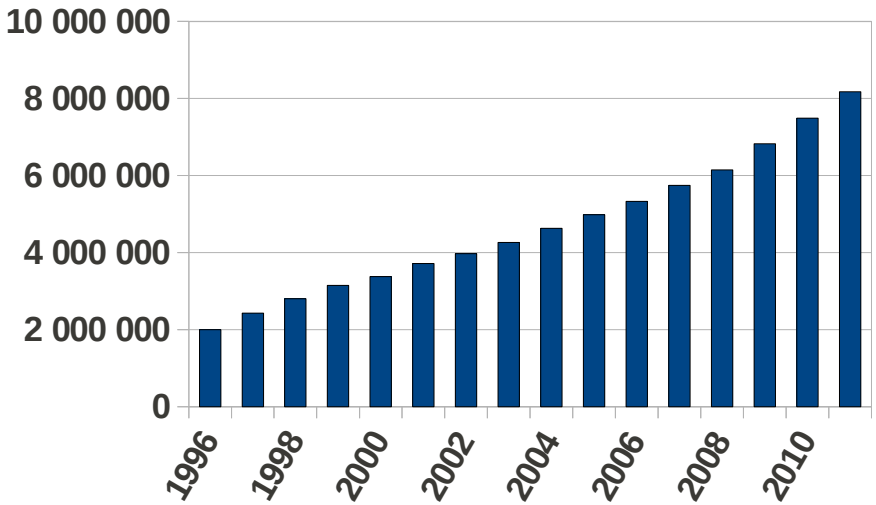
Identifiers



References

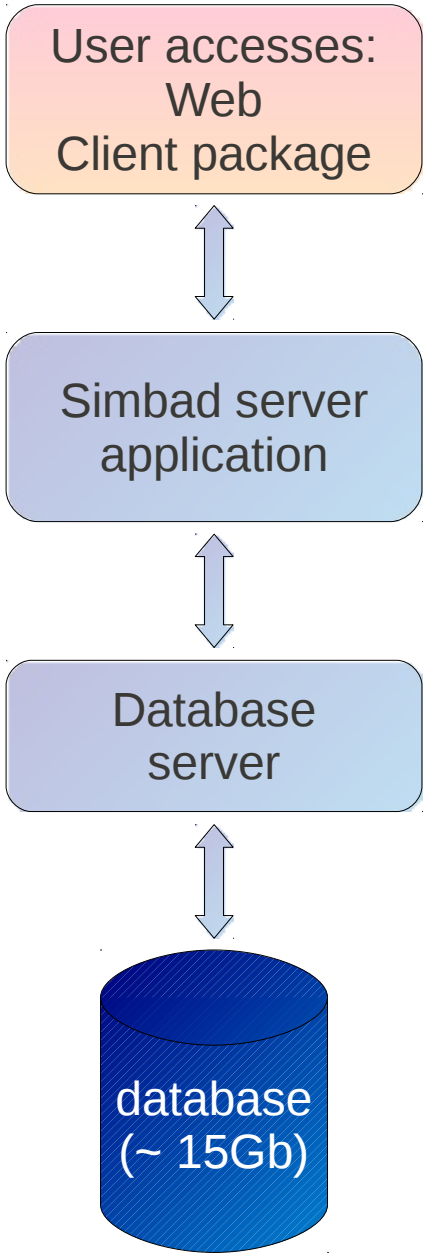


Citations



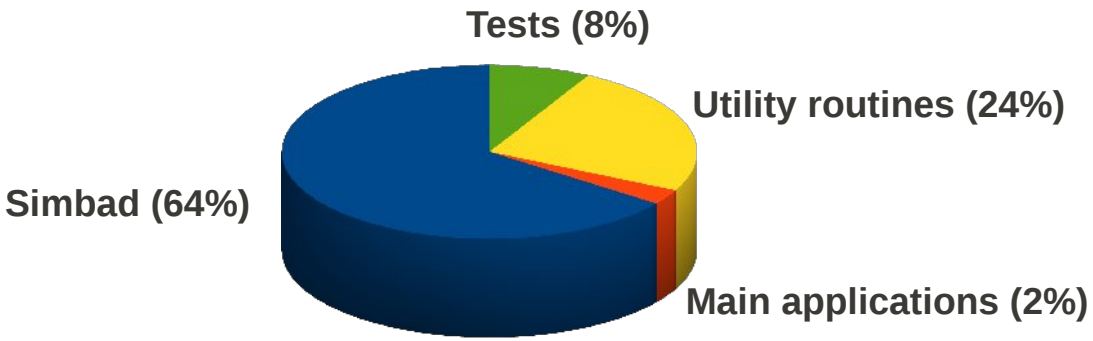
Simbad current environment

- **SIMBAD runs on three Linux Pcs**
- **Security: only one server, managing the user interfaces, is accessible from the outside world**
- **The Database Management System is PostgreSQL, an open source relational DBMS**
- **A mirror, nightly updated, runs at the CfA**

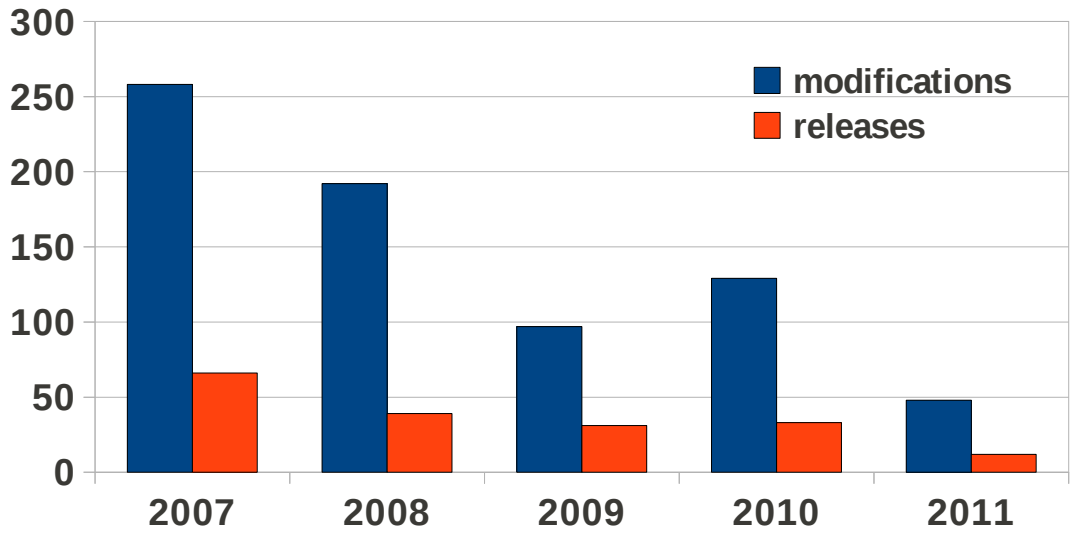


Simbad software

- **Two main developers: Anaïs Oberto, Marc Wenger**
- **The software is 99% written in Java (+1% scripts)**
- **330,000 lines of code in 1170 Java classes**



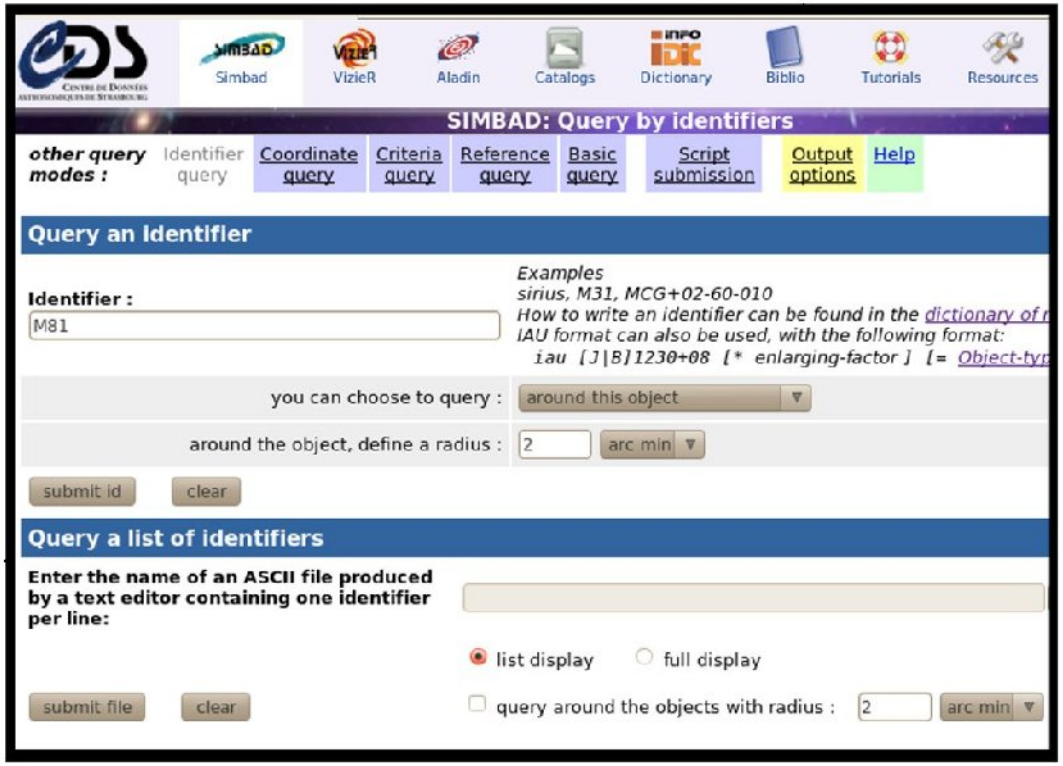
- **the software evolution shows in 4 ½ years 720 modifications in 180 releases.**



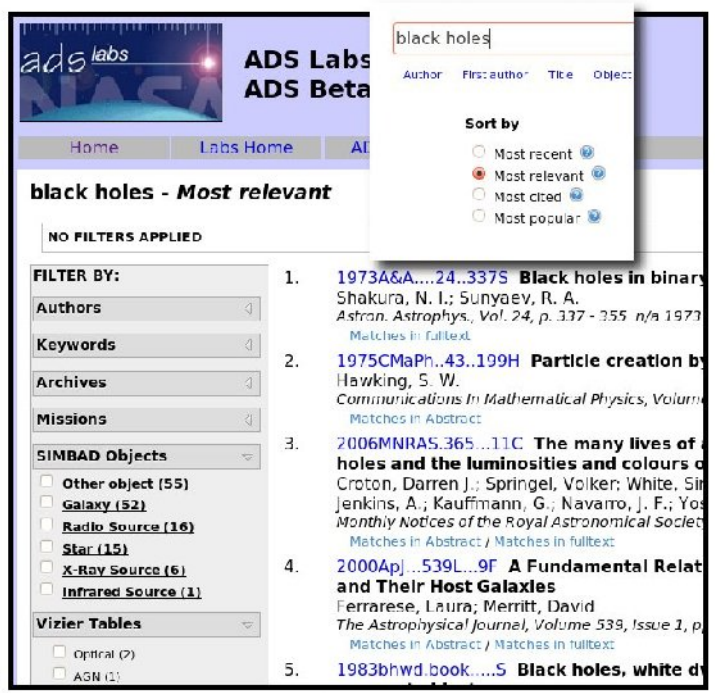
Software usage

- **Queries through web pages**
- **Queries in applications using URL accesses**
- **Queries through the Simbad client (simcli) package for 'legacy' applications in C language, in particular observatory archives including major NASA archives**
- **Sesame : a name resolver facility implemented as a Web service**
- **Queries by identifiers, references, coordinates and regions, criteria**
- **Output formats : HTML, Ascii, VOTable, JSON**

Some Simbad accesses



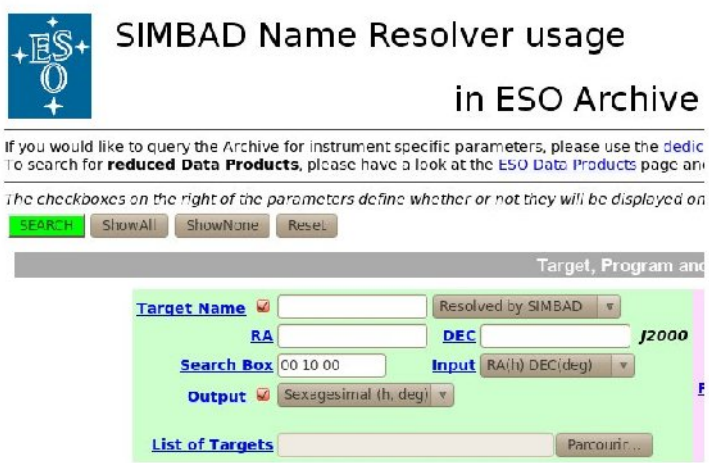
Simbad interface



ADS Labs
SIMBAD object types from a list of references

SIMBAD Name Resolver usage

in ESO Archive

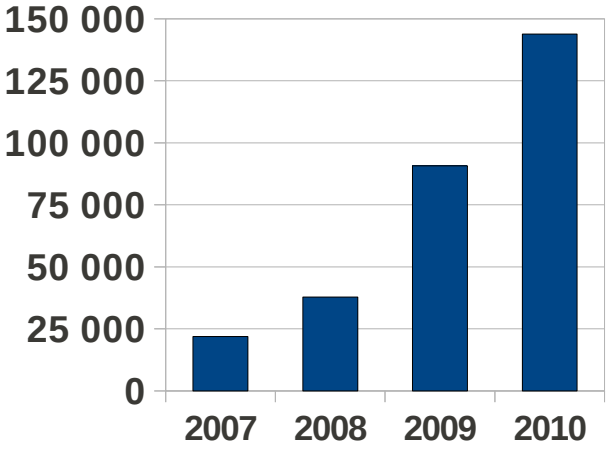


in ESA ISOC Science Data Archive

in STScI MAST Archive

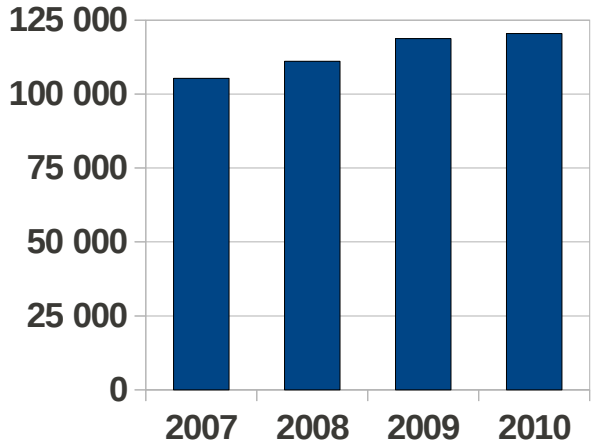
Usage statistics

Web queries / day

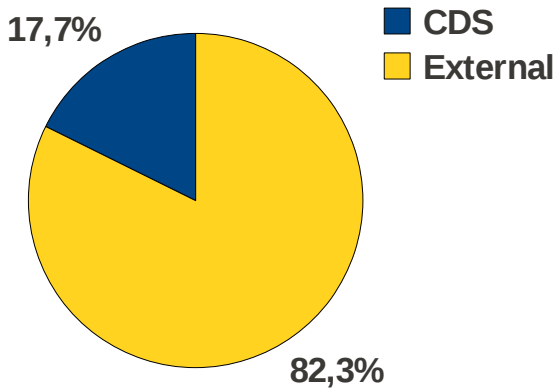


2010:
265,000
queries/day

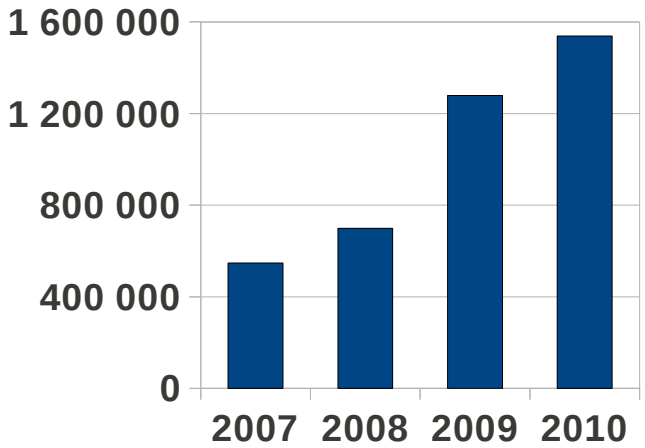
Simcli queries / day



Ratio CDS/external queries

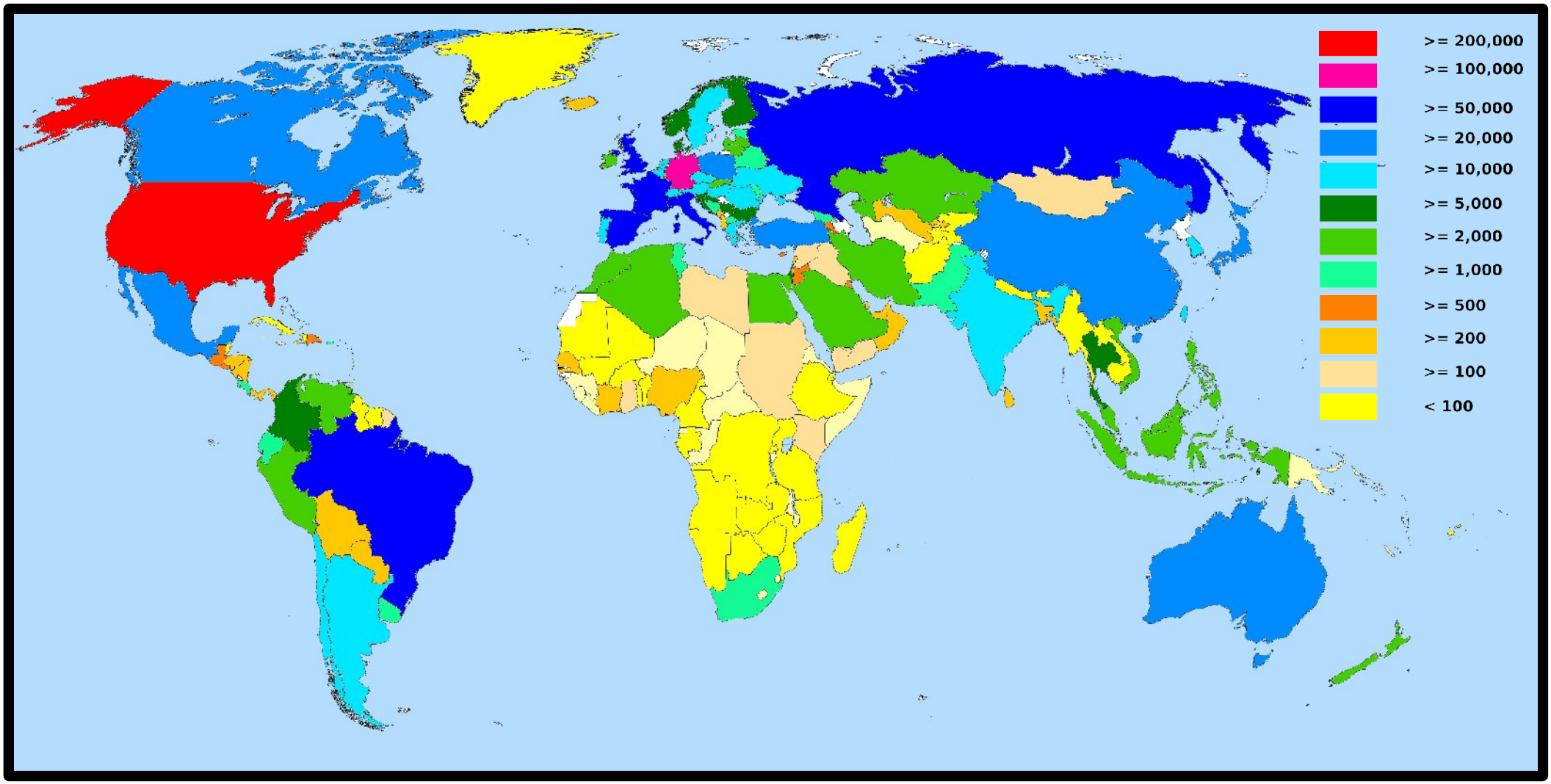


Different IP addresses / year



Users / country

Different users (IP addresses) from Jan 2007 to Jun 2009 : 223 countries



Four countries have never used SIMBAD in these 30 month:
Western Sahara, Kosovo, Azerbaidjan, North Korea

Evolutions

- **Implemented in 2010 : User annotations in objects**
- **Implemented in 2011 : VO Table Access Protocol (TAP) to query SIMBAD using Astronomical Data Query Language (ADQL) and allowing batch/asynchronous queries**
- **To come in 2012 / 2013:**
  - **A new Web interface, more interactive rich web client using up-to-date technologies (HTML5, CSS3 and JQuery/Javascript)**
  - **A better integrated on-line help pages and user's guide**
  - **Query optimisation, already prototyped in the *name resolver* facility**