


# Travel Search Engine

SophiaConf - July 2011



# Agenda

- Context
  - The Problem
  - An OpenSource Solution
  - Future
  - Questions
- 

Context



# Context

- Lots of data related to travel and leisure that are spread all over the internet
  - Airport, Airline, Hotels, Cities, Historical information, Climate information
  - Geonames, Ourairports, Wikipedia, Wikimapia ...
- Each one could have been stored in many different formats
  - RDF, TXT, CVS ...

# The Problem



# The Problem

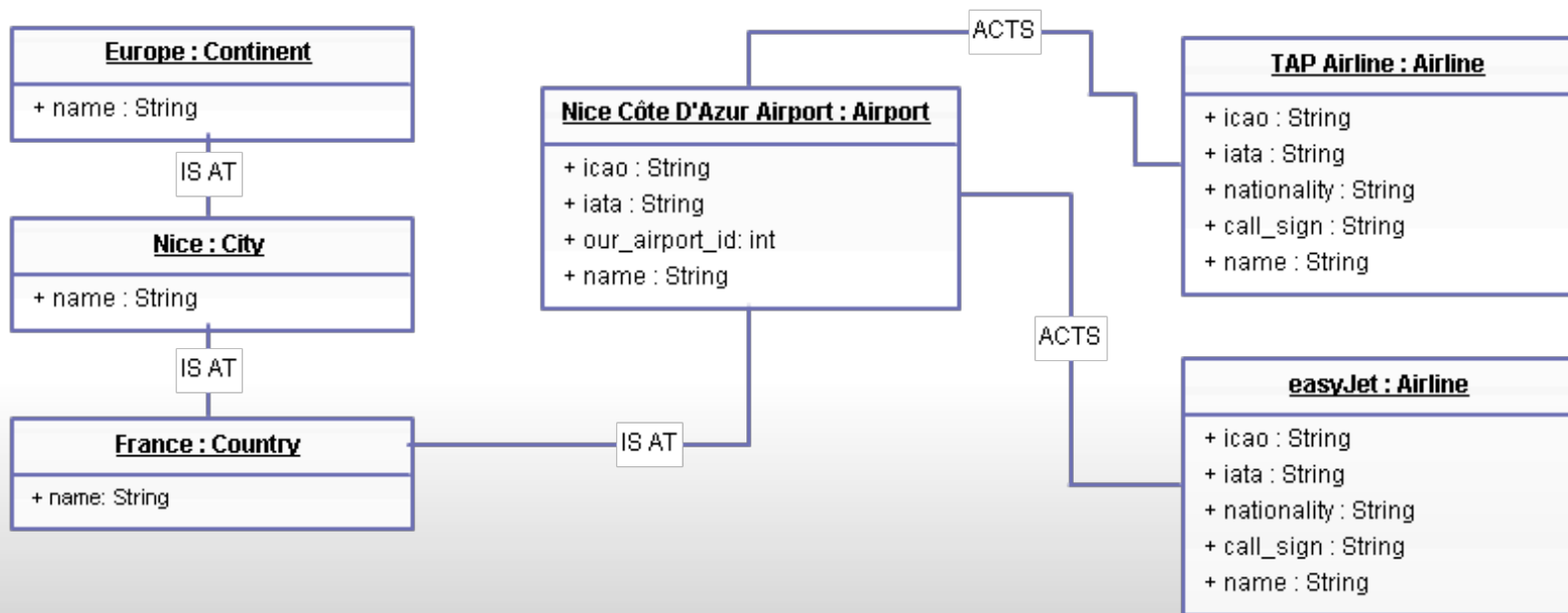
- Merge all the interesting and useful travel and leisure data available in a way that makes sense.
- To build both a human and a machine interface to ease the access to all public data related to travel and leisure.

# An OpenSource Solution



# An OpenSource Solution

- Finding one entity with some characteristics is trivial, but how to find relevant information based on it ?
  - Graph database !

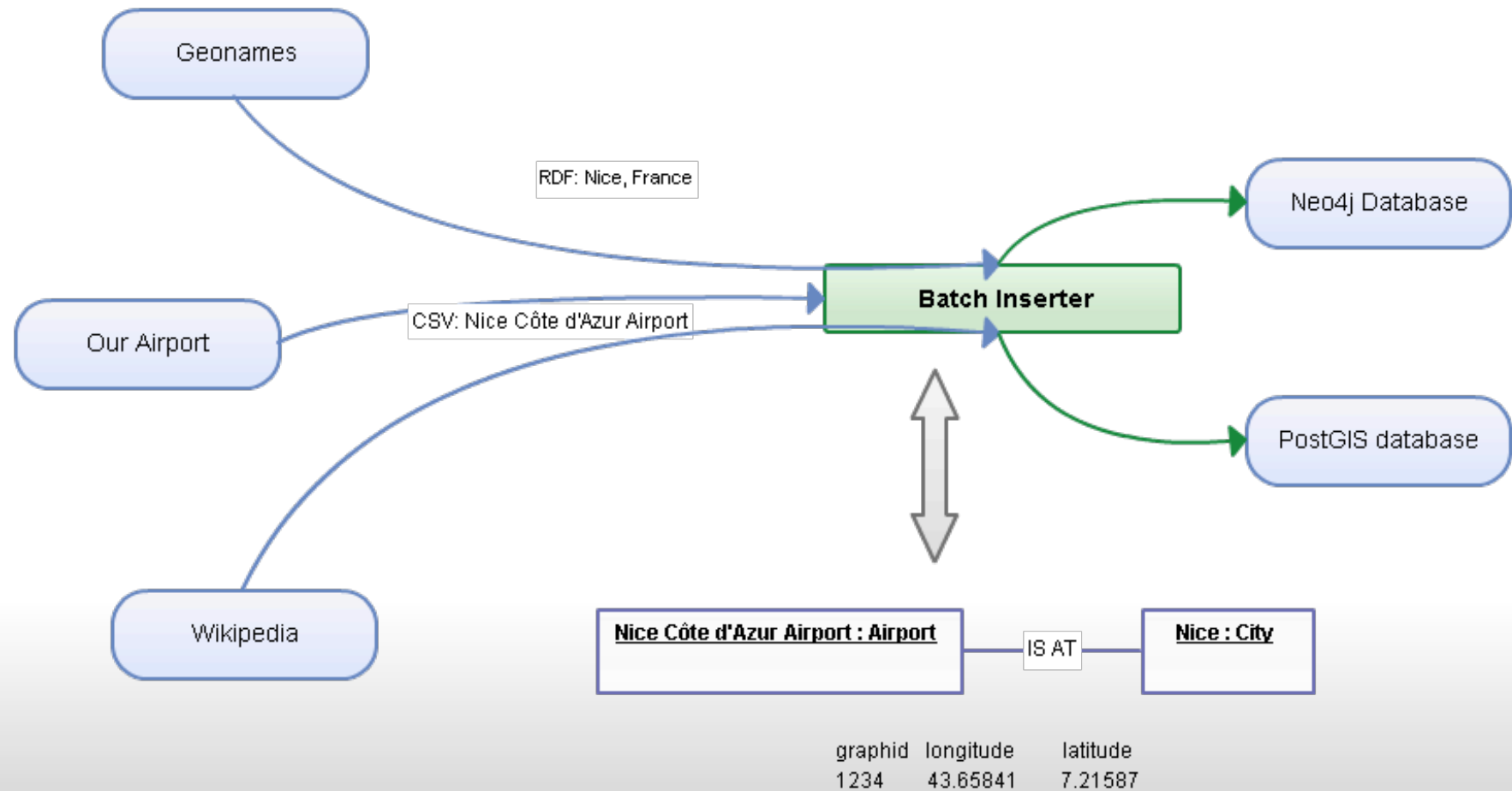




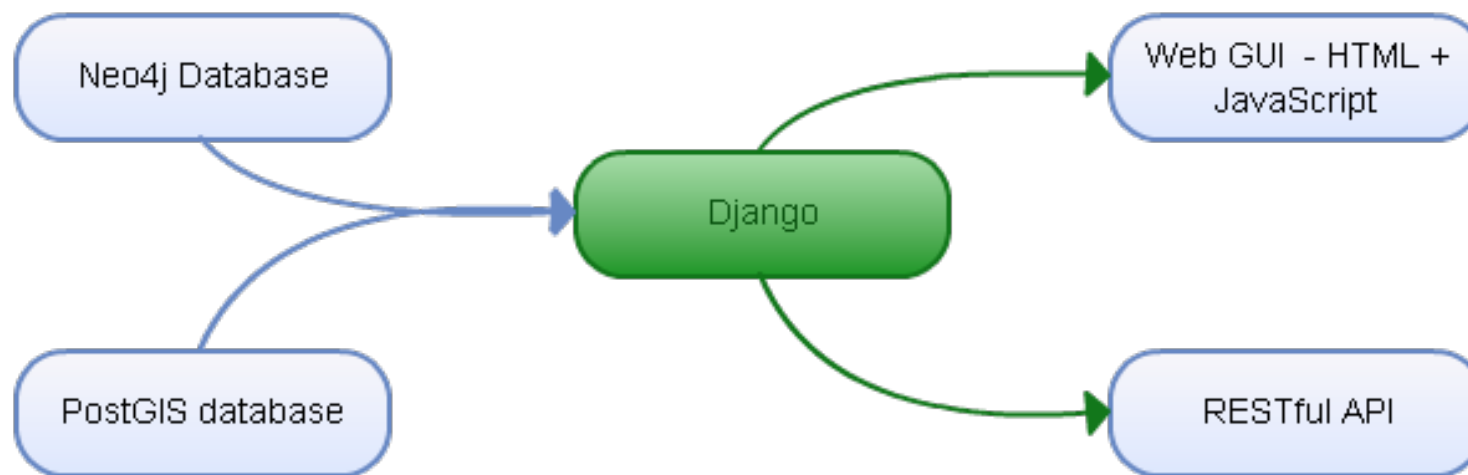
# An OpenSource Solution

- **Neo4j** : the graph database
  - Easy data navigation
  - Ex.: "node.traverse(order=[constants.BREADTH\_FIRST])"
- **Apache Lucene**: the indexer
  - Exact and fulltext indexing
  - Ex.: "name:Nice AND description:beach\* AND description:private\*"
- **Postgres PostGIS**: Geographical database
  - An easy way to represent geographical data.
  - Ex.: "give me all the airports that are 20 km away from a given point".
- **Django**: web framework
  - To ease building the front-end and it's communication with the graph and geographical database.
- **Jquery + HTML**
  - To build a simple and responsive user interface.

# An OpenSource Solution



# An OpenSource Solution



Future



# Future

- **New entities and relationships**
  - Hotels, Aircraft, Parks, Concert halls ...
  - "In New York, where I can find a Park that's called something like 'central' ? "
- **A more natural language processing**
  - For making easier know what the user is looking for
- **Translation for all properties**
- **Properties and indexing for the relationships**
- **Link to the source of the information**
  - The user could change the information directly on the source.
- **RESTful API**
  - Provide the same information so other services/programs could use it.

Questions ??