Capgemini and Cloudera Big Data Solution Helps Global Employment Agency Improve Insight into Localized Job Markets

One of the world's largest HR firms is using Cloudera Enterprise to match job seekers with openings and build new services for its clients

The Situation

The client is one of the world's largest HR service providers, operating in approximately 40 countries. It faced a challenge around the sheer volume of information that is increasingly becoming available in job markets today. Sources include everything from job boards, institutions, LinkedIn, and other social media, to videos and company reports. Even sources such as weather forecasts can be relevant in relation to seasonal and tourism industry job opportunities.

The HR service provider wanted to develop a platform to consolidate all available jobs in a particular geography. They asked Capgemini to demonstrate how an enterprise data hub, based on Cloudera Enterprise, could deliver a Big Data solution that would help them gain faster access to higher quality information.

The initial objective was to find ways to tackle volumes of information for its own business, with the secondary aim of developing entirely new HR services for governments and companies. This would include becoming far more proactive in job markets, anticipating requirements and opportunities, and acting to fulfill recruitment requirements.
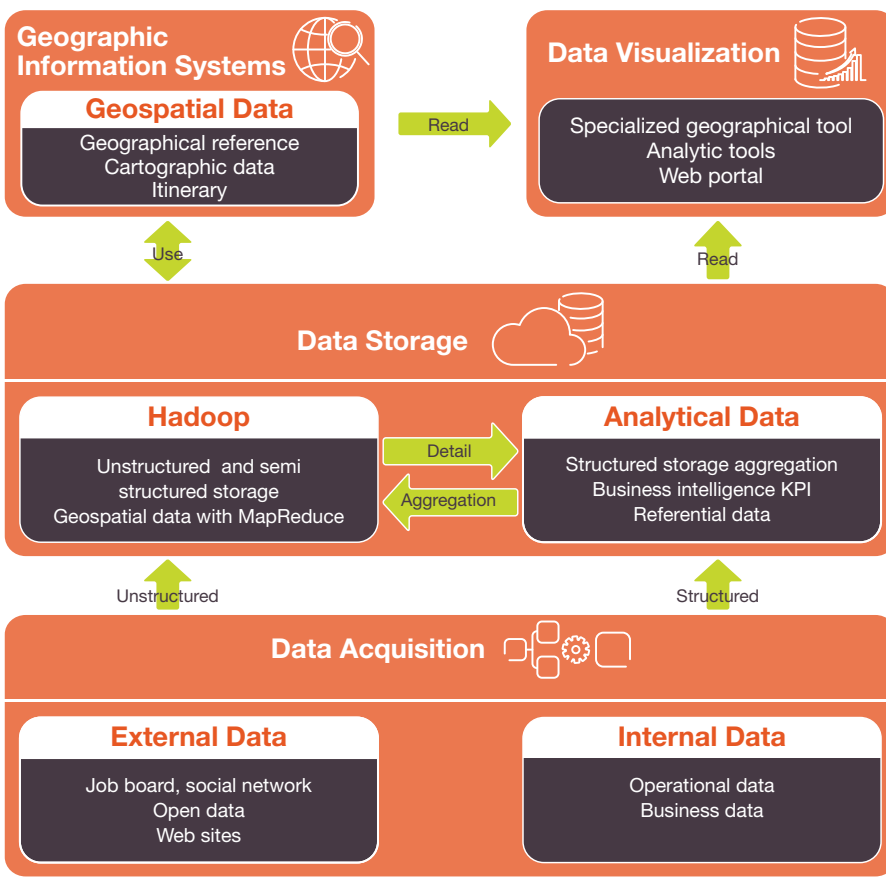
The Solution

A Big Data solution was needed to manage and deliver rapid insight into the sheer volumes of data involved. Capgemini and Cloudera experts worked closely with the client's team to list and prioritize the use-cases relevant for their initial business goals. This led to a proof of concept (POC) project to match job demand and supply in a specific geography - in this case, France.

People matter, results count.

Capgemini's solution involved implementing four major elements not present in the existing system:

- Cloudera Enterprise to store all data and run the data modeling engine
- An analytical SQL database running in conjunction with Cloudera Enterprise
- A data visualization product
- A solution to geo-code the geographical data

As part of the solution, a learning algorithm processes the text within CVs and job openings to identify and flag similarities. The solution also delivers data from the visualization product to the existing reporting tools which are already familiar to business users.

Figure 1: Architecture

Geographic Information Systems		Data Visualization	
Geospatial Data Geographical reference Cartographic data Itinerary		**Specialized geographical tool** Analytic tools Web portal	

Read

Use / Read

Data Storage

Hadoop	Analytical Data
Unstructured and semi structured storage Geospatial data with MapReduce	Structured storage aggregation Business intelligence KPI Referential data

Detail / Aggregation

Unstructured / Structured

Data Acquisition

External Data	Internal Data
Job board, social network Open data Web sites	Operational data Business data

The Result

The initial project has successfully demonstrated the viability of the Big Data solution to deliver improved insight into job markets in a specific geography. This enables users to more proactively match job seekers with relevant openings.

The system automated the matching of offers and CVs. For instance it was able to analyze 200,000 documents (including offers and CVs) in only two hours using a handful of computers. The client has verified the results of the analysis, with the feedback that "users have been amazed" by the accuracy of the system.

After a successful and low-cost POC, the project is now being expanded to explore around 15 additional business use-cases. These include features such as the ability to have the amount of employment contracts signed per month, per activity, and per work area or the ability to analyze the travel time between the job seeker and the job offer.

About Capgemini

With more than 130,000 people in over 40 countries, Capgemini is one of the world's foremost providers of consulting, technology and outsourcing services. The Group reported 2013 global revenues of EUR 10.1 billion.

Together with its clients, Capgemini creates and delivers business and technology solutions that fit their needs and drive the results they want. A deeply multicultural organization, Capgemini has developed its own way of working, the Collaborative Business Experience™, and draws on Rightshore®, its worldwide delivery model.

Learn more about us at
www.capgemini.com/bim

In collaboration with

cloudera®

Cloudera is revolutionizing enterprise data management by offering the first unified Platform for Big Data, an enterprise data hub built on Apache Hadoop™. Cloudera offers enterprises one place to store, process and analyze all their data, empowering them to extend the value of existing investments while enabling fundamental new ways to derive value from their data. Only Cloudera offers everything needed on a journey to an enterprise data hub, including software for business critical data challenges such as storage, access, management, analysis, security and search. As the leading educator of Hadoop professionals, Cloudera has trained over 20,000 individuals worldwide. Over 1,000 partners and a seasoned professional services team help deliver greater time to value. Finally, only Cloudera provides proactive and predictive support to run an enterprise data hub with confidence. Leading organizations in every industry plus top public sector organizations globally run Cloudera in production.

More information is available at:
www.cloudera.com

MCOS_GI_AH_20140703

People matter, results count.