

23.07.2008

Matching of Job Descriptions with CVs – A Turnkey InfoCodex Solution

1. Scenario

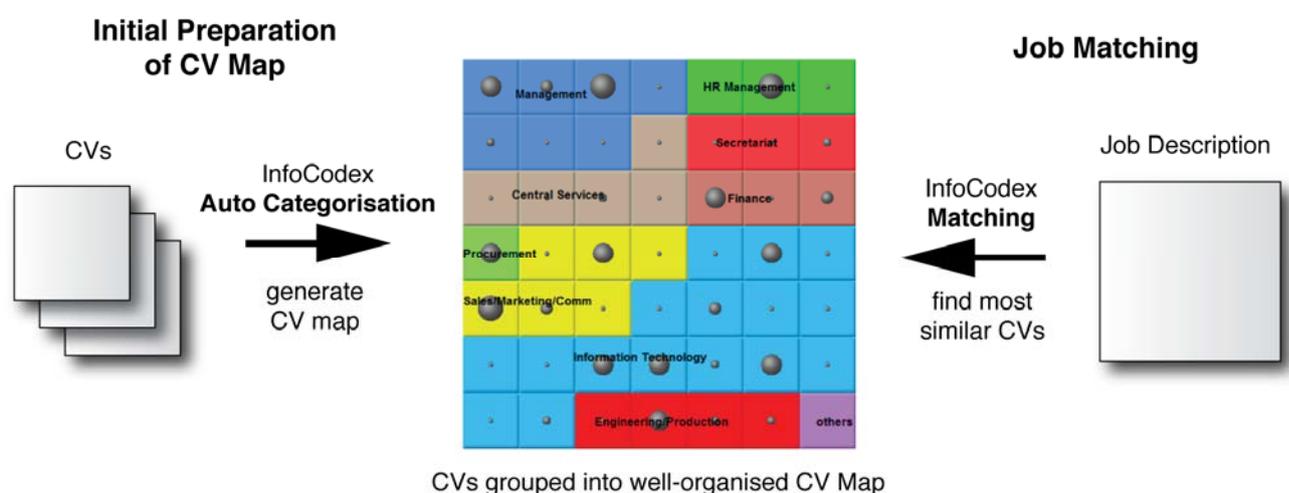
Efficient and effective personnel recruitment and HR succession planning requires an automatic content-specific matching of job descriptions with a pool of curricula vitae (CVs). The CVs and the job descriptions are often structured differently, use different terminologies, and may even exist in different languages.

Finding the best suited candidates for a particular job is a time consuming and resource intense task but is essential in human resource management. Very often, the number of potential candidates (i.e. the number of existing CVs) for a particular job is large, and vice versa, the number of potential jobs for a particular candidate is large. Hence, a fully automated and integrated solution would yield substantial benefits.

2. The Solution

InfoCodex fully automatically analyses the available CVs, i.e. it recognizes the content and extracts the relevant facts of each individual CV. Using this information, the CVs are then organised into a “CV map”. The structure of the CV map is determined automatically by the contents of the various CVs and is optimised continuously as the pool of CVs changes. In order to automate the matching of job descriptions with CVs, InfoCodex analyses then the job descriptions of the positions to be filled in the same way as the CV and positions each of them into the CV map. This process is comparable to an automatic grouping of books (job descriptions) into a well-organized book-shelf (CV map). The CVs belonging to the compartment of the CV map where the job description was positioned are the best matches.

This mapping procedure is truly *cross-lingual* and takes into account the *effective content* of the CVs. For example, an English, French, German, Italian, and Spanish CV with equivalent content are recognized by the system as very similar CVs and hence are located at the same position in the CV map.



The matching of a specific job description with the available CVs is done by InfoCodex’s similarity search engine that compares the effective content of the job description (= search text) with all CVs and returns a list of the most relevant CVs.

Example output of the InfoCodex solution:

Retrieval by Text Search

Human Resources Generalist Job Description
 Position Description:
 The HR Generalist manages the day-to-day operations of the Human Resource office. The HR Generalist manages the administration of the human resources policies, procedures and programs. The HR

[Choose words occurring in current collection](#)

Thematic Grouping by: Clustering | Heat Map

Management	HR-Management		
Central Services	Secretariat	Finance	
Procurement			
Sales/Marketing/Comm			
	Information Technology		
	Engineering/Production		others

Total found: 292 doc. of 577
Field: Gestion RH 2
 55 doc. / 1 hit
 Cd1080-2
 Ressources humaines (Interviewer Gestion des carrières Aspect légal du recrutement)

Current Selection (click title for abstracts, commenting etc.) 1-10 of total 55

Relev	Date	WordsOrigin	N/A Concepts	Fav. Comments
hi		responsable formation; responsable ressources humaine		
	17.03.08 370	C:\xfer\icvs	human resources, management, recruitment	<input type="checkbox"/>
hi		responsable ressources humaine; responsable emploi / formation; responsable développement rh		
	17.03.08 1002	C:\xfer\icvs	human resources, management, software	<input type="checkbox"/>

3. Application Options

A. Interactive Matching

The job descriptions are entered into the search mask (i.e. by copy/paste) of the interactive InfoCodex application. The results can be displayed in list format or exported into a file.

B. Integration into an Existing System

An http request containing the job description is sent to the server (using the POST or the GET method), and the results are returned as an XML stream that can be read and processed by the calling system.

Basic Example (crude form, just for explanation purposes)

Minimum required relevance

Maximum number of CVs returned

Text

Human Resources Generalist Job Description

Position Description:
 The HR Generalist manages the day-to-day operations of the Human Resource office. The HR Generalist manages the administration of the human resources policies, procedures and programs. The HR Generalist carries out responsibilities in the following functional areas: departmental development, Human Resource Information Systems (HRIS), employee relations, training and development, benefits, compensation, organizational development, and employment.

The HR generalist is responsible for all or part of these areas:

- recruiting and staffing logistics;
- organizational and space planning;
- performance management and improvement systems;
- organization development;
- employment and compliance to regulatory concerns and reporting;
- employee orientation, development, and training;
- policy development and documentation;
- employee relations;

Returned answer (XML stream)

```

<?xml version="1.0" encoding="ISO-8859-1" ?>
<SJ>
  <Update>-1</Update>
  <version>SJWE 3.36 XP</version>
  <Count>10</Count>
  - <CLst>
    - <Record>
      <Cid>168</Cid>
      <relev>100%</relev>
      <cvtit>responsable formation; responsable ressources humaine</cvtit>
      <Email>magail.rofidal@voila.fr</Email>
      <KeyFilter>9#20000000400000001#0-07F-2F-1FF-1F-2-000000C000002000-000F-08-2-1</KeyFilter>
      <Words>ressource humaine, gestion/direction, recruter, droit du travail, budget, bilan, restructuration d'organisation, relation syndicats, cycle de recrutement,
      organisation, formation, législation</Words>
      <Domain>Gestion RH - Ressources humaines</Domain>
    </Record>
    - <Record>
      <Cid>373</Cid>
      <relev>92%</relev>
      <cvtit>responsable ressources humaine; chargée des ressources humaine; chargée de recrutement</cvtit>
      <Email>stefaridon@yahoo.fr</Email>
      <KeyFilter>3#F80048#18-0FF-0E-1FC-1F-2-0000008000000000-0003-7F-2-1</KeyFilter>
      <Words>ressource humaine, gestion/direction, recruter, gestion des carrières, organisation, synthèse orale et écrite, rédaction d'annonces, planification de ressources,
      maîtrise outils de Sourcing, droit du travail, entretien semi directif, entretien directif</Words>
      <Domain>Gestion RH - Ressources humaines</Domain>
    </Record>
  etc.

```

C. Web Service

The system can also be implemented as “Software as a Service (SaaS)”

D. Functional Enhancement

Instead of matching specific jobs to a pool of CVs, the system can be configured the other way around: Job offers from a variety of sources are arranged into a job map or landscape and CVs are matched with the open jobs.

This is particularly useful for publishers and job-market-search providers.

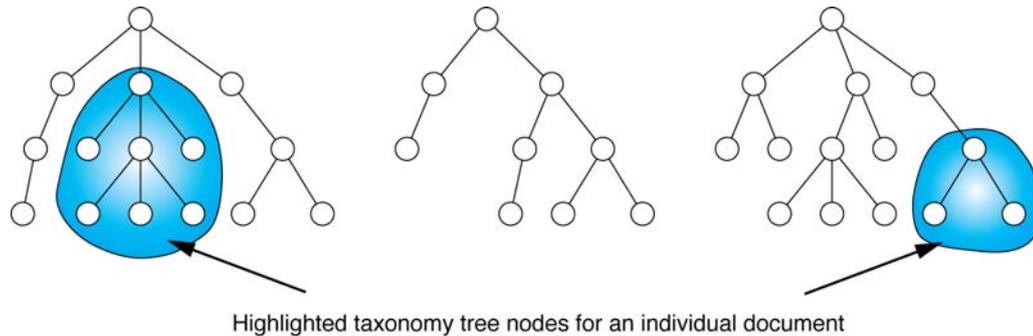
E. Customizing for User-specified Job Category Schemes (Job Taxonomy)

The user can stipulate a job taxonomy (fixed job map) which is tailored to the relevant thematic areas. The optional job taxonomy can be supplied by an Excel table.

	A	B	C	D
1	Job Categories			
2				
3	Code	Lv	Category	Keywords
4				
5	Cd1010	1	Management	General mangement, business administration
6				consulting, business organization
7				management consultant
8	Cd1020	2	Executive Management	general director, CEO
9				member of the board
10				corporate executive, top management
11	Cd1030	2	Middle Management	Product development
12				planning and logistics
13				project managemement, organization of work
14				quality management

4. Technical Foundation of the Job Matching Procedure

The words or expressions (groups of words such as "European Union", "Enterprise Search Engine" etc.) present in the text are matched against InfoCodex's linguistic database comprising more than 3 Million words/expressions which are structured into synonym groups. These synonym groups point to a node in a universal taxonomy tree characterizing the meaning of the matched synonym. During the analysis of a document, all nodes in the taxonomy tree which are addressed by the matching process are highlighted, and the ensemble of highlighted nodes is a measure for the thematic areas that are covered by the present document.



The addressed thematic areas of each document are then projected into a **100-dimensional content-space**, and finally a categorisation of the documents is achieved by means of a **self-organising neural network** (Kohonen-Map). The categorization of a document by InfoCodex is not a simple assignment to a single node in the taxonomy tree, but rather a multidimensional projection.

This neural network provides a **well-founded similarity measure** based on information-theoretical principals which allows the comparison of documents according to their content. A small distance between documents corresponds to a high similarity and vice versa. In mathematical terms, the similarity measure is given by the weighted scalar product of the two vectors, corrected by the Kullback-Leibler distance from the main themes, combined with the weighted score-sum of the matching keywords and their nodes in the taxonomy tree, respectively.

The similarity measure is independent of the language of the document and is only weakly dependent on the exact wording. The described processes are patented in the EU and USA.