

¿Qué leen los usuarios de Facebook? Análisis de la red social

Eva Cepeda García

Resumen—El propósito de este trabajo es el análisis y la implementación de un sistema de recopilación de datos públicos o accesibles de la red social LinkedIn a través del desarrollo de una aplicación Java. La aplicación analiza y correlaciona estos datos para implementar dos tipos de aplicaciones, en primer lugar la selección del mejor candidato o candidatos para una empresa y puesto de trabajo a partir de la obtención de un mapa de aptitudes de los actuales empleados en el mismo puesto o puestos similares y la comparativa de este modelo de aptitudes con las aptitudes de los candidatos, y en segundo lugar, la obtención de datos estadísticos de alumnos y ex alumnos de una universidad, como en qué sector trabajan, en qué zona geográfica, como se definen (aptitudes), etc....El análisis intenta contemplar todas las problemáticas derivadas de la minería de datos sobre sistemas con información tan heterogénea y tantas restricciones de obtención de datos según el volumen, gestionado a través de distintos tipos de licenciamiento.

Palabras clave—LinkedIn, red, social, aptitudes, empleados, alumnos, candidatos, minería de datos.

Abstract—The purpose of this work is the analysis and implementation of a system for collecting data accessible or public off social network LinkedIn by developing a Java application. The application analyzes and correlates this data to implement two types of applications, first selecting the best candidate or candidates for a job and company from obtaining a map of skills of current employees in the same job or similar positions and comparing this skill model with the skills of the candidates, and secondly, the collection of statistical data on students and alumni of a university, and in what sector they work, in what geographical area, as defined (skills), etcThe analysis attempts to cover all the problematic a rising from data mining on heterogeneous information systems with so many constraints and data collection according to the volume, handled through different types of licensing.

Index Terms—LinkedIn, networking, social, skills, staff, students, candidates, data mining.

1 INTRODUCCIÓN

EN la sociedad actual, el auge de las redes sociales es indiscutible, la revolución de los dispositivos móviles ha potenciado en gran parte este auge. Existen varias redes sociales de gran repercusión como Facebook con 1.150 millones de usuarios, orientada a la actividad personal y social, Google+ con 500 millones de usuarios, Twitter con 500 millones de usuarios orientada al microblogging, LinkedIn con 238 millones de usuarios orientada exclusivamente al plano profesional y con gran presencia de marcas, 3 millones de ellas ya han registrado su perfil de empresa, Pinterest con 70 millones de usuarios, Instagram con 150 millones de usuarios.....

Esta es la realidad actual, la presencia en la red de personas y empresas se ha convertido en algo habitual y casi necesario... la pregunta es, ¿Están nuestros datos? ¿Cuanto se preocupan los usuarios de proteger sus datos de

forma adecuada? ¿Qué fácil es acceder a ellos?

La respuesta que hemos obtenido a través del desarrollo del presente proyecto de fin de grado, es que los usuarios no cuidan la privacidad de sus datos, es sorprendente la cantidad de información que se puede obtener de una persona a la que ni siquiera conoces y que por supuesto, no forma parte de tu red social. Afortunadamente las empresas responsables de estas redes sociales, ante el auge de sus productos y la facilidad de acceso a sus datos que han demostrado algunos expertos [1], han restringido sus API's de acceso a los datos para minimizar el número de datos que un proceso automático puede obtener.

El uso de las redes sociales puede ser muy positivo y provocar satisfacción a nivel social a un gran número de usuarios, pero los peligros que entraña la difusión pública de determinados datos, es algo que debería preocuparnos, sobre todo en ciertos colectivos más vulnerables como niños y adolescentes, que sufren en algunos casos problemas nada virtuales debido al mal uso de las redes sociales.

-
- E-mail de contacto: eva.cepeda@e-campus.uab.cat
 - Mención realizada: Computación.
 - Trabajo tutorizado por: Ramon Baldrich (CVC)
 - Curso 2013/14

Solo con educación y concienciación de los usuarios es posible garantizar la seguridad virtual de las personas, ya que como he podido comprobar, siempre hay una forma de conseguir la información sobre alguien si se tienen los medios y conocimientos necesarios, y lo que se haga de esa información depende principalmente de nuestra conciencia....

2 ESTADO DEL ARTE

Este proyecto está inspirado en la lectura del artículo "Personality, Gender, and Age in the Language of Social Media: The Open-Vocabulary Approach" que detalla los resultados de un proyecto de investigación sobre análisis de datos en la red social Facebook en el que participaron H. Andrew Schwartz, Johannes C. Eichstaedt, Margaret L. Kern, Lukasz Dziurzynski, Stephanie M. Ramones, Megha Agrawal, Achal Shah, Michal Kosinski, David Stillwell, Martin E. P. Seligman, Lyle H. Ungar [2].

La primera idea para el desarrollo del proyecto, fue realizar una minería de datos sobre la red social Facebook, para obtener datos relativos a las preferencias literarias del máximo número posible de usuarios cuyo perfil público incluyera esa información, poder analizarla según el género, la edad y/o la zona geográfica y confeccionar un mapa de preferencias literarias para un determinado perfil de usuario con el objetivo de recomendar unos títulos u otros en función del usuario conectado.

Después de algunas pruebas lo que inicialmente parecía factible no lo fue, ya que en los últimos años y debido a la facilidad existente hasta no hace mucho tiempo para obtener datos públicos en cantidades masivas de esta red social [2], Facebook ha restringido mucho el acceso a los datos públicos a través de su API de desarrollo. A pesar de que una gran cantidad de usuarios deja sus perfiles completos totalmente públicos, la API de desarrollo solo permite acceder a los datos que Facebook considera públicos (nombre, genero, URL del perfil público y localización geográfica), y solo permite acceder a los perfiles completos de los contactos del usuario conectado, lo cual desvirtúa mucho los resultados del estudio ya que la muestra que obtenemos es muy reducida.

Por todo esto, reorientamos el proyecto hacia la red social LinkedIn cuya orientación más profesional nos inspira otro tipo de análisis y las API's de desarrollo disponibles nos permiten acceder más fácilmente a cierto tipo de información [3].

- ✓ **Análisis 1:** Selección del mejor candidato o candidatas para una empresa y puesto de trabajo a partir de la obtención de un mapa de aptitudes de los actuales empleados en el mismo puesto o puestos similares y la comparativa de este modelo de aptitudes con las aptitudes de los candidatos.

- ✓ **Análisis 2:** Obtención de datos estadísticos de alumnos y ex alumnos de una universidad y estudios concretos, como en qué sector trabajan, en que zona geográfica, como se definen (aptitudes), etc....

3 OBJETIVOS

A través de la recopilación de datos públicos o accesibles de la red social LinkedIn, la aplicación analizará y correlacionará estos datos para implementar dos tipos de aplicaciones:

- **Aplicación 1:** Selección del mejor candidato o candidatas para una empresa y puesto de trabajo a partir de la obtención de un mapa de aptitudes de los actuales empleados en el mismo puesto o puestos similares y la comparativa de este modelo de aptitudes con las aptitudes de los candidatos.
- **Aplicación 2:** Obtención de datos estadísticos de alumnos y ex alumnos de una universidad, como en qué sector trabajan, en que zona geográfica, como se definen (aptitudes), etc....

4 METODOLOGÍA

4.1 Estructura de la aplicación

El sistema constará de tres módulos:

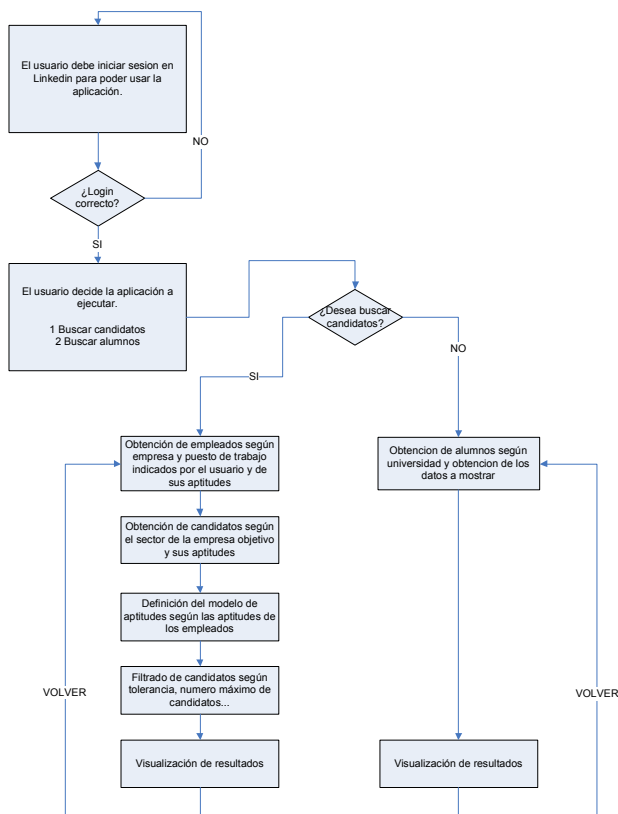
- ❖ Recopilación de datos: Recopilación de datos públicos de perfiles de LinkedIn [4].
- ❖ Análisis de datos: Análisis estadístico de los datos en función de unas restricciones o filtros.
- ❖ Visualización de datos: Generación de un gráfico que muestre de forma visual el resultado del análisis.

4.2 Plataforma de desarrollo

La aplicación se desarrollará en jsp/java sobre un servidor de aplicaciones open source (Liferay v6.1 CE Tomcat 7 GA3). La aplicación no almacenará datos.

El acceso a la aplicación de forma anónima se realizará a través del protocolo HTTP, aunque será necesario el registro en LinkedIn [5] [6] para poder obtener resultados ya que las búsquedas se realizan con los permisos del usuario conectado lo que permite obtener más o menos resultados en función de su tipo de licencia [7].

4.3 Diagrama de flujo de la aplicación



5 DESARROLLO DE HERRAMIENTAS COMUNES

5.1 Registro de la aplicación

LinkedIn quiere asegurarse de que las aplicaciones que usan sus API's son aplicaciones que aportan valor a la red social, es por esto que para utilizar determinadas API's de búsqueda de personas se debe registrar la aplicación y realizar una solicitud que debe ser previamente aprobada antes de poder usar la API desde la aplicación [8].

Además la aplicación debe ser debidamente registrada en su sistema.

5.2 Autenticación del usuario

Independientemente de los permisos de que disponga la aplicación respecto al uso de las API's o el número de consultas diarias que puede realizar, las llamadas a las API's de búsqueda se deben realizar con los permisos de un usuario registrado en LinkedIn con lo cual es necesario que el usuario que va a ejecutar la aplicación realice un proceso de login sobre LinkedIn.

El proceso de login está definido en la documentación de la API de LinkedIn sobre el proceso de autenticación [5].

5.3 Llamadas a las API'S de búsqueda

Una vez obtenido el token de acceso podemos realizar las llamadas a las API's de búsquedas necesarias para obtener los resultados deseados.

Los parámetros de los que disponemos a la hora de realizar búsquedas de personas son los siguientes:

```

http://api.linkedin.com/v1/people-search?keywords=[space delimited keywords]&first-name=[firstname]&last-name=[lastname]&company-name=[companyname]&current-company=[true | false]&title=[title]&current-title=[true | false]&school-name=[schoolname]&current-school=[true | false]&country-code=[country code]&postal-code=[postal code]&distance=[miles]&start=[number]&count=[1-25]&facet=[facetcode, values]&facets=[facetcodes]&sort=[connections | recommenders | distance | relevance]
keywords=[spacedelimitedkeywords]&first-name=[firstname]&last-name=[lastname]&company-name=[companyname]&current-company=[true | false]&title=[title]&current-title=[true | false]&school-name=[schoolname]&current-school=[true | false]&country-code=[country code] &postal-code=[postal code] &distance=[miles] &start=[number]&count=[1-25]&facet=[facetcode, values] &facets=[facetcodes]&sort=[connections | recommenders | distance | relevance]
  
```

Como soporte en el desarrollo y para probar las API's y sus parámetros hemos utilizado la "LinkedIn API Console" [9].

5.4 Obtención de resultados con varias búsquedas

Uno de los problemas que nos encontramos a la hora de realizar las búsquedas, es que por cada llamada a la API solo podemos obtener un máximo de 20 resultados, con lo cual debemos realizar sucesivas llamadas a la API variando los parámetros &start=X&count=Y para poder obtener el total de resultados de la búsqueda. Este total de resultados dependerá de la cuenta con la que nos conectemos. En el caso de una cuenta gratuita los resultados accesibles serán unos 100 aproximadamente [7]. El resultado de la búsqueda te da el total de resultados encontrados que puede ser mucho mayor que el número de resultados accesibles por el usuario.

6 DESARROLLO DE LA APLICACIÓN 1: BÚSQUEDA DE CANDIDATOS

6.1 Obtención de empleados

El objetivo es encontrar los empleados de una empresa concreta que pertenecen a un determinado puesto de

trabajo y realizar un mapa de aptitudes que nos sirva para encontrar los mejores candidatos para una empresa y puesto de trabajo concreto.

Para filtrar el puesto utilizamos el campo “headline” que define el puesto de trabajo actual de un usuario.

En estos resultados podemos observar la problemática que supone lo heterogéneo de las definiciones de los puestos de trabajo:

- ❖ Diferentes idiomas de denominación.
- ❖ Diferentes denominaciones para el mismo puesto de trabajo.

Solucionar este problema para que la búsqueda que realiza el usuario en su idioma y a partir de una descripción de puesto de trabajo determinada pudiera traducirse de forma que abarcara todas las definiciones de un mismo puesto de trabajo independientemente del idioma, implica un volumen de trabajo considerable, no previsto inicialmente y que realiza una aportación significativa a la consecución de los objetivos del proyecto, con lo cual queda fuera del alcance del mismo. La búsqueda se realizará en base a los términos introducidos por el usuario en la búsqueda de forma literal.

6.2 Obtención de candidatos

El objetivo es encontrar una lista de usuarios que inicialmente pudieran ser afines de alguna forma a la empresa para la que se busca cubrir el puesto de trabajo y refinar esa lista en base a sus aptitudes respecto al mapa de aptitudes confeccionado a partir de los empleados encontrados.

Para realizar la lista inicial de usuarios y con el objetivo de acotar de alguna forma esta primera búsqueda, utilizamos el campo “Industry” de la empresa objetivo. De esta forma en la lista inicial de candidatos realizamos un primer filtro que descarta usuarios no relacionados con el sector de la empresa.

6.3 Obtención de las aptitudes de un usuario

Por restricciones de la API los datos sobre aptitudes solo son accesibles si se tiene pleno acceso al perfil, y ese acceso depende del tipo de licencia que se disponga. Con la licencia gratuita no podemos acceder a esos datos a través de la API con lo cual se hemos tenido que buscar una alternativa. La alternativa consiste en parsear el HTML que proporciona la URL pública del perfil del usuario en cuestión, que incluye los datos sobre los SKILLS del usuario.

Vemos que el código fuente de la página está bastante bien estructurado y la sección de aptitudes tiene un formato fácilmente parseable. Esto nos permite obtener la lista de aptitudes (skills) de cada usuario.

```

<div class="header">
  <div class="content">
    <ol id="skills-list" class="skills">
      <li class="competency show-bean">
        <span class="jellybean"> Lotus Connections </span>
      </li>
      <li class="competency show-bean">
        <span class="jellybean"> WebSphere Portal </span>
      </li>
      <li class="competency show-bean">
      <li class="competency show-bean">
      <li class="competency show-bean">
      <li class="competency show-bean">
      <li class="competency show-bean">
      <li class="competency show-bean">
      <li class="competency show-bean">
      <li class="competency show-bean">
      <li class="competency show-bean">
      <li class="competency show-bean">
    </ol>
  </div>
</div>

```

En este punto nos encontramos con el problema de que si en algún momento LinkedIn cambia la estructura de la página, y la sección de aptitudes queda modificada a nivel de código HTML, nuestra aplicación puede no ser capaz de encontrar las aptitudes de los usuarios con lo que la aplicación perdería su funcionalidad. Se podría plantear la forma de minimizar el problema implementando un parseador más flexible y adaptable a cambios, pero no se ha realizado esta implementación ya que esta fuera de los objetivos del proyecto.

6.4 Definición del modelo de aptitudes

A partir de los empleados encontrados para una empresa y puesto de trabajo, generamos un modelo de aptitudes (skill model) a partir del cual seleccionaremos los candidatos más apropiados a partir de la lista inicial.

Para cada empleado encontrado para una empresa y puesto de trabajo, obtenemos su lista de aptitudes tal y como se ha visto en el punto anterior. Para cada aptitud de la que disponga el usuario, si ya existe en el modelo de aptitudes, incrementamos en uno la ocurrencia para esa aptitud y si no existe, se incluye como nueva aptitud, de esta forma obtenemos una lista de aptitudes por ocurrencias dentro de los empleados.

La lista de aptitudes en el modelo puede ser muy grande, con el objetivo de acotar el modelo, le pediremos al usuario que indique el número mínimo de ocurrencias que debe tener una aptitud para aparecer en el modelo, de esa forma podemos descartar aptitudes con muy pocas ocurrencias a favor de aptitudes más relevantes.

Una vez acotado el modelo de aptitudes, la aplicación lo muestra como una nube de palabras o nube de etiquetas para que el usuario pueda ver de forma visual el modelo de aptitudes obtenido. Gracias a esta representación, el usuario puede ver de forma rápida cuales son las aptitudes más relevantes, es decir, las que aparecen con más frecuencia en nuestro modelo de aptitudes, ya que aparecen resaltadas con colores distintos y tamaños proporcionales a su relevancia.

La nube de palabras está generada a través de un código

go HTML y una hoja de estilos CSS que aplica estilos proporcionales a la relevancia del término.



La definición de las aptitudes de los usuarios se realiza de forma libre, es decir, a la hora de definir el modelo nos encontramos de nuevo con el problema de la definición heterogénea de aptitudes equivalentes como nos paso también en la definición de los puestos de trabajo, ya sea por diferencia de idioma o por divergencia en la denominación de una misma aptitud. Resolver este problema o minimizar al máximo sus efectos supone un esfuerzo en desarrollo que esta fuera del alcance de este proyecto.

6.5 Filtrado de candidatos

A partir de la lista de inicial candidatos afines al puesto de trabajo que queremos cubrir, a partir de la búsqueda por el sector de la empresa objetivo, empieza un proceso de refinamiento a partir del modelo de aptitudes confeccionado, del que resultará una lista de candidatos cuyas aptitudes estén lo mas alineadas posible con respecto a las aptitudes de los actuales empleados de la empresa que cubren el puesto de trabajo demandado.

Para poder refinar esta lista de candidatos primero debemos decidir qué tipo de coincidencia debe tener con el modelo de aptitudes, es decir, queremos darle más importancia a los candidatos que dispongan de un mayor número de aptitudes coincidentes con el modelo, o por el contrario queremos los candidatos que dispongan de las aptitudes más relevantes (con mayor ocurrencia) del modelo.

Para indicar este aspecto del filtrado, el usuario debe seleccionar la tolerancia. El campo tolerancia tiene tres opciones:

- ❖ **Tolerancia estricta:** se da más importancia a la relevancia de las aptitudes que a la cantidad de aptitudes coincidentes. Es decir, para calcular la puntuación final del candidato respecto al resto, se suma el número de ocurrencias de cada aptitud coincidente con el modelo de aptitudes. Un candidato con pocas aptitudes coincidentes, si estas son de mayor relevancia, conseguirá mejor puntuación que otro candidato con un mayor número de aptitudes coincidentes, pero que tengan menor relevancia.
- ❖ **Tolerancia laxa:** buscamos candidatos con coinciden-

cia en el mayor número de aptitudes. Para calcular la puntuación final del candidato respecto al resto, se suma uno por cada aptitud coincidente con el modelo de aptitudes. Un candidato con un mayor número de coincidencias obtendrá mejor puntuación que el resto, aunque estas aptitudes sean las menos relevantes del modelo.

- ❖ **Tolerancia continua:** método de tolerancia continuo. Se da más importancia a la relevancia de las aptitudes que a la cantidad de aptitudes coincidentes matizando esa importancia en función de parámetro corrector. Es decir, para calcular la puntuación final del candidato respecto al resto, se suma el número de ocurrencias de cada aptitud coincidente con el modelo de aptitudes elevado a 1/alfa, donde alfa es el factor de tolerancia continua.

Para proporcionar un criterio más de filtrado, el usuario debe indicar el número máximo de candidatos que debe devolver el proceso de filtrado.

Los aptitudes de los candidatos que no se encuentren en el modelo de aptitudes construido a partir de los empleados encontrados, no afecta al resultado final ya que su valor es 0, es decir no se tienen en cuenta a la hora de calcular la puntuación de los empleados.

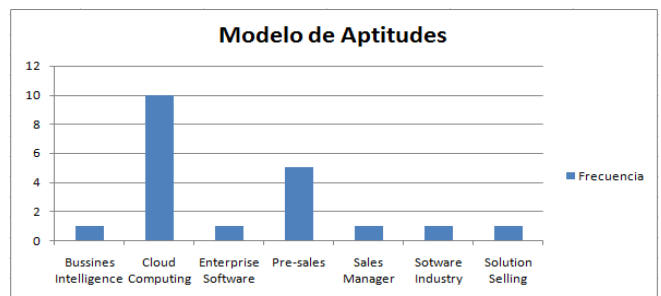
6.6 Cálculo del histograma de aptitudes

El modelo de aptitudes construido a partir de los empleados encontrados según los criterios de búsqueda indicados por el usuario, es una lista de aptitudes con unas frecuencias asociadas, es decir es una distribución de frecuencias en base a unas aptitudes, y esta distribución de frecuencias define nuestro modelo de aptitudes.

“Un histograma es una representación gráfica de una distribución de frecuencias por medio de rectángulos, cuyas anchuras representan intervalos de la clasificación y cuyas alturas representan las correspondientes frecuencias.”[10]

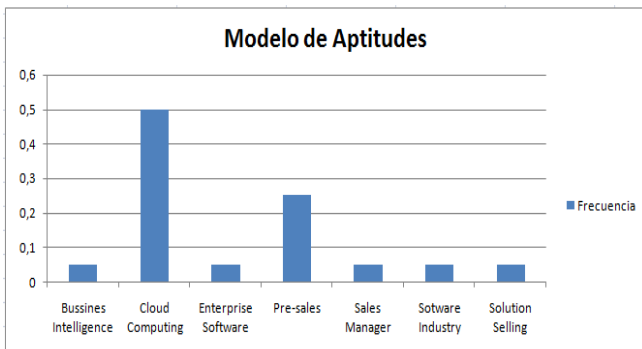
En base a esta definición, podemos representar nuestro modelo de aptitudes mediante un histograma o distribución de frecuencias de este tipo.

Por ejemplo:

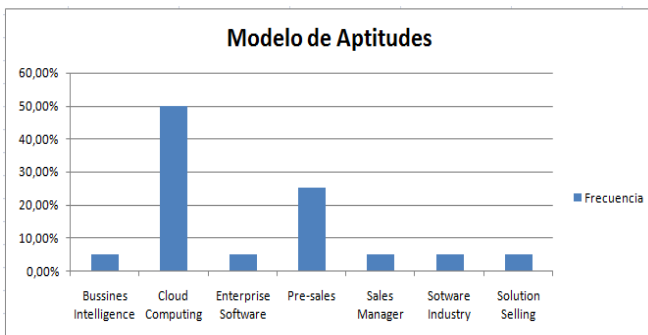


En base a esto hablaremos de histogramas de aptitudes de los empleados para definir los modelos de aptitudes obtenidos en las búsquedas, definiremos $h_E(A)$ como el histograma de aptitudes de los empleados (o modelo de aptitudes). Los valores que puede tomar la función para cada aptitud individual, $h_E(A_j)$, va de 1 a infinito. Ninguna aptitud del modelo podrá tener un valor inferior a 1, ya que el modelo se construye en base a las aptitudes de los empleados encontrados en la búsqueda realizada por el usuario.

Para normalizar los resultados, convertiremos los valores de frecuencia de las aptitudes del modelo en valores entre 0 y 1 de forma que la suma total de ocurrencias sea 1.



Para que la representación sea más clara, podemos mostrar los porcentajes.



Por otro lado definimos $h_{Ei}(A)$ como el modelo de aptitudes del empleado i -ésimo de la búsqueda. Este modelo es una lista de aptitudes de las que dispone con frecuencia 1.

$$\text{Por lo tanto } h_E(A) = \sum_i (h_{Ei}(A))$$

Para poder definir el histograma de aptitudes de los candidatos $h_C(A)$ debemos decir que es la lista de aptitudes de las que dispone el candidato respecto al modelo.

Es decir, a cada aptitud individual del candidato $h_{Ci}(A_j)$ que pertenezca a nuestro modelo de actitudes (es decir que el valor en el modelo de datos para ese aptitud

sea > 0) le damos el valor 1, en caso contrario su valor será 0.

$$h_{Ci}(A_j) = 1 \text{ si } h_E(A_j) > 0 \text{ y } 0 \text{ si } h_E(A_j) = 0$$

Una vez obtenidos los valores para todos los candidatos de $h_C(A)$, los candidatos se ordenan en función de $S(h_E(A), h_{Ci}(A))$.

El valor de esta función depende del tipo de tolerancia que apliquemos a la hora de calcular los candidatos resultantes.

Si aplicamos la tolerancia estricta, lo que haremos es dar más importancia a la relevancia de las aptitudes que a la cantidad de aptitudes coincidentes. Es decir, para calcular la puntuación final del candidato respecto al resto de ellos, se suma el número de ocurrencias de cada aptitud coincidente con el modelo de aptitudes. Un candidato con pocas aptitudes coincidentes, si estas son de mayor relevancia, conseguirá mejor puntuación que otro candidato con un mayor número de aptitudes coincidentes, pero que tengan menor relevancia.

$$S(h_E, h_{Ci}) = \sum_j (h_E(A_j) * h_{Ci}(A_j))$$

Si aplicamos la tolerancia laxa, le daremos más importancia al número de aptitudes coincidentes que a su relevancia. Para calcular la puntuación final del candidato respecto al resto, se suma uno por cada aptitud coincidente con el modelo de aptitudes. Un candidato con un mayor número de coincidencias obtendrá mejor puntuación que el resto, aunque estas aptitudes sean las menos relevantes del modelo.

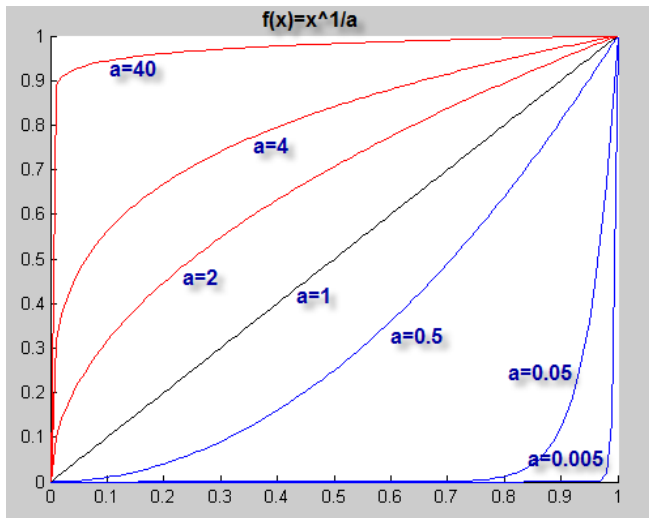
$$S(h_E, h_{Ci}) = \sum_j (h_E(A_j) * h_{Ci}(A_j))$$

donde $h_E(A_j) = 1 \forall j$.

Por último, si aplicamos la tolerancia continua, le daremos más importancia a la relevancia de las aptitudes que a la cantidad de aptitudes coincidentes, matizando esa importancia en función de parámetro corrector. Es decir, para calcular la puntuación final del candidato respecto al resto, se suma el número de ocurrencias de cada aptitud coincidente con el modelo de aptitudes elevado a $1/\alpha$, donde alfa es el factor de tolerancia continua. Los resultados para este tipo de tolerancia, cuando el factor de corrección sea 1, debe coincidir exactamente con los resultados que obtendríamos si aplicáramos tolerancia estricta.

$$S(h_E, h_{Ci}) = \sum_j (h_E(A_j)^{(1/\alpha)} * h_{Ci}(A_j))$$

Si analizamos el valor de la función $h_E(A_j)^{(1/\alpha)}$ podemos observar cómo se matiza al alza o a la baja, la relevancia de la aptitud del empleado en función del factor de corrección.



Si aplicamos valores de corrección a la tolerancia continua mayores a 1

$$V = h_E(A_j)^{(1/a)} \text{ cuando } a > 1$$

observamos que se amplifica la relevancia de la aptitud, se hace más estricta (líneas rojas).

Si aplicamos valores de corrección a la tolerancia continua menores a 1

$$V = h_E(A_j)^{(1/a)} \text{ cuando } a < 1$$

observamos que se reduce de forma gradual, la relevancia de la aptitud, se hace más laxa (líneas azules).

Si aplicamos valores de corrección a la tolerancia continua iguales a 1

$$V = h_E(A_j)^{(1/a)} \text{ cuando } a = 1$$

observamos que el valor de la aptitud no sufre ninguna corrección, obtenemos el mismo resultado que en la tolerancia estricta (línea negra).

7 DESARROLLO DE LA APLICACIÓN 2: BÚSQUEDA DE ALUMNOS

7.1 Obtención de alumnos

El objetivo es encontrar los alumnos de una universidad o centro de estudios concreto y mostrar unos indicadores que resuman algunos de los datos recogidos de la búsqueda como cuáles son sus aptitudes, cual es la zona geográfica en la que residen o en qué sector trabaja.

Para filtrar hemos utilizado el parámetro de búsqueda `school-name=` que permite buscar una universidad o centro de estudios dentro de la sección "Educación" del usuario aunque este dato concreto no sea accesible direc-

tamente desde la API.

En estos resultados podemos observar la problemática que supone lo heterogéneo de las definiciones de los centros de estudios o universidades:

- ❖ Diferentes idiomas de denominación.
- ❖ Diferentes denominaciones para el mismo dentro de estudios.

Solucionar este problema para que la búsqueda que realiza el usuario en su idioma y a partir de una descripción de un centro de estudios o universidad determinado pudiera traducirse de forma que abarcara todas las definiciones de un mismo centro de estudios o universidad independientemente del idioma, implica un volumen de trabajo considerable, no previsto inicialmente y que realiza una aportación significativa a la consecución de los objetivos del proyecto, con lo cual queda fuera del alcance del mismo. La búsqueda se realizará en base a los términos introducidos por el usuario en la búsqueda de forma literal.

7.2 Obtención de alumnos

El objetivo es encontrar una lista de usuarios que estudian o han estudiado en una universidad o centro de estudios concretos y mostrar algunos indicadores de datos significativos obtenidos a través de la API como el sector o la localización geográfica, o a través del parseo del código HTML de su página pública de perfil como el caso de las aptitudes.

Debemos recordar que uno de los problemas que nos encontramos a la hora de realizar las búsquedas es que por cada llamada a la API solo podemos obtener un máximo de 20 resultados, con lo cual debemos realizar sucesivas llamadas a la api variando los parámetros `&start=X&count=Y` para poder obtener el total de resultados de la búsqueda.

7.3 Obtención de las aptitudes (skills) de un alumno

Por restricciones de la API los datos sobre aptitudes solo son accesibles si se tiene pleno acceso al perfil, y ese acceso depende del tipo de licencia que se disponga. Con la licencia gratuita no podemos acceder a esos datos a través de la API con lo cual se hemos tenido que buscar una alternativa. La alternativa consiste en parsear el HTML que proporciona la URL pública del perfil del usuario en cuestión, que incluye los datos sobre los SKILLS del usuario.

Vemos que el código fuente de la pagina está bastante bien estructurado y la sección de aptitudes tiene un formato fácilmente parseable. Esto nos permite obtener la lista de aptitudes (skills) de cada usuario.

```

<div class="header">
  <div class="content">
    <ol id="skills-list" class="skills">
      <li class="competency show-bean">
        <span class="jellybean"> Lotus Connections </span>
      </li>
      <li class="competency show-bean">
        <span class="jellybean"> WebSphere Portal </span>
      </li>
      <li class="competency show-bean">
      <li class="competency show-bean">
      <li class="competency show-bean">
      <li class="competency show-bean">
      <li class="competency show-bean">
      <li class="competency show-bean">
      <li class="competency show-bean">
      <li class="competency show-bean">
      <li class="competency show-bean">
      <li class="competency show-bean">
      <li class="competency show-bean">
      <li class="competency show-bean">
      <li class="competency show-bean">
      <li class="competency show-bean">
      <li class="competency show-bean">
      <li class="competency show-bean">
      <li class="competency show-bean">
    </ol>
  </div>
</div>

```

En este punto nos encontramos con el problema de que si en algún momento LinkedIn cambia la estructura de la página, y la sección de aptitudes queda modificada a nivel de código HTML, nuestra aplicación puede no ser capaz de encontrar las aptitudes de los usuarios con lo que la aplicación perdería su funcionalidad. Se podría plantear la forma de minimizar el problema implementando un parseador más flexible y adaptable a cambios, pero no se ha realizado esta implementación ya que esta fuera de los objetivos del proyecto.

7.4 Definición del modelo de indicadores

A partir de los alumnos encontrados para un centro de estudios o universidad, generamos un modelo de aptitudes (skill model), un modelo de localización geográfica y un modelo de sector en el que trabajan, a partir de los cuales construiremos y visualizaremos los indicadores correspondientes.



La definición de los atributos de los usuarios se realiza de forma libre, es decir a la hora de definir los modelos nos encontramos de nuevo con el problema de la definición heterogénea de aptitudes, localizaciones, sectores, etc.... equivalentes como nos paso también en la definición de los puestos de trabajo, ya sea por diferencia de idioma o por divergencia en la denominación de un mismo concepto. Resolver este problema o minimizar al máximo sus efectos supone un esfuerzo en desarrollo que esta fuera del alcance de este proyecto.

4 RESULTADOS

Para presentar el funcionamiento de la aplicación y justificar los resultados obtenidos hemos realizado una

serie de búsquedas con diferentes parámetros a través de las cuales podremos ver representadas las explicaciones realizadas en apartados anteriores.

La primera secuencia de búsquedas realizada, corresponde a la aplicación de selección de candidatos y refleja las diferencias obtenidas en los resultados en función de los parámetros de tolerancia introducidos.

La búsqueda se basa en los siguientes parámetros:

Empresa: IBM, Puesto: Ventas, Ocurrencias mínimas: 1, N° candidatos: 100

Variantes según la tolerancia

Tolerancia: estricta

Tolerancia: laxa

Tolerancia: continua, Factor de tolerancia continua: 2.0

Tolerancia: continua, Factor de tolerancia continua: 1.0

Tolerancia: continua, Factor de tolerancia continua: 0.5

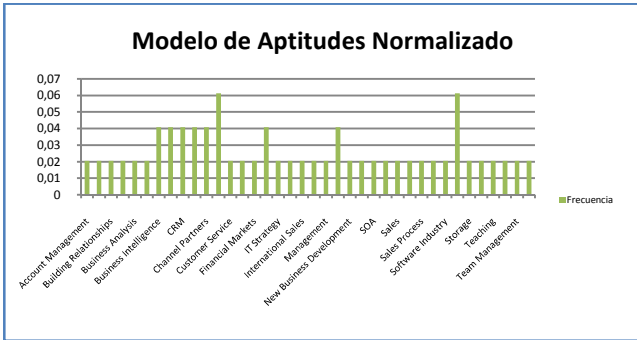
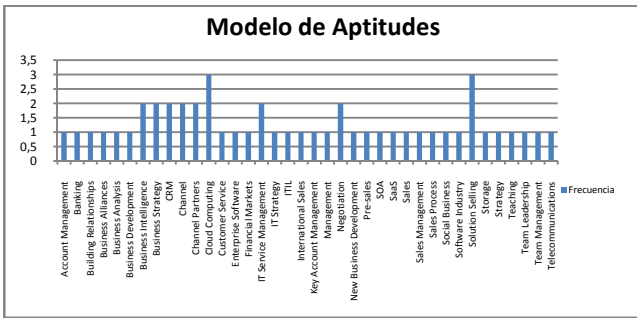
En todos los casos obtenemos el mismo modelo de aptitudes:

Aptitud	Frecuencia	Frecuencia Normalizada
Account Management	1	0,02040816
Banking	1	0,02040816
BuildingRelationships	1	0,02040816
Business Alliances	1	0,02040816
Business Analysis	1	0,02040816
Business Development	1	0,02040816
Business Intelligence	2	0,04081633
Business Strategy	2	0,04081633
CRM	2	0,04081633
Channel	2	0,04081633
ChannelPartners	2	0,04081633
Cloud Computing	3	0,06122449
CustomerService	1	0,02040816
Enterprise Software	1	0,02040816
FinancialMarkets	1	0,02040816
IT Service Management	2	0,04081633
IT Strategy	1	0,02040816
ITIL	1	0,02040816
International Sales	1	0,02040816
Key Account Management	1	0,02040816
Management	1	0,02040816
Negotiation	2	0,04081633
New Business Development	1	0,02040816
Pre-sales	1	0,02040816
SOA	1	0,02040816
SaaS	1	0,02040816
Sales	1	0,02040816
Sales Management	1	0,02040816
Sales Process	1	0,02040816
Social Business	1	0,02040816
Software Industry	1	0,02040816
SolutionSelling	3	0,06122449
Storage	1	0,02040816
Strategy	1	0,02040816
Teaching	1	0,02040816
TeamLeadership	1	0,02040816
Team Management	1	0,02040816
Telecommunications	1	0,02040816

49

1

Cuya representación gráfica a través de un histograma sería la siguiente:



Para facilitar al usuario la visión de este modelo de aptitudes, la aplicación la muestra como una nube de palabras.



A partir de este punto se calcula la lista de candidatos y su puntuación en función de la tolerancia seleccionada.

En la tabla del *Anexo 1* podemos ver la comparativa de todas las tolerancias seleccionadas.

Como podemos ver en la tabla, las tolerancias estrictas potencian a las aptitudes con mayor relevancia, mientras que las laxas, potencian el número de aptitudes coincidentes con el modelo.

También podemos observar, según justificamos anteriormente (Ver 6.6 Cálculo del histograma de aptitudes) que los resultados obtenidos con la tolerancia estricta y con la tolerancia continua aplicando un valor de corrección de 1, son equivalentes.

Para ver el efecto que tiene el problema de heterogeneidad de datos, hemos realizado una búsqueda traduciendo a Inglés el término Ventas, es decir, buscaremos para la misma empresa pero para el puesto Sales.

Empresa: IBM, Puesto: Sales, Ocurrencias mínimas: 1, Nº candidatos: 100, Tolerancia: estricta

A diferencia de la búsqueda en castellano, hemos obtenido un modelo de aptitudes mucho más grande. Esto nos indica que es mucho más frecuente la definición de los puestos de trabajo en Inglés que en castellano en cuanto a perfiles de LinkedIn se refiere.

Si el modelo de aptitudes obtenido es demasiado grande, sería conveniente restringir la inclusión en el modelo de aquellas aptitudes minoritarias.

Volvemos a realizar la búsqueda con los siguientes parámetros:

Empresa: IBM, Puesto: Sales, Ocurrencias mínimas: 7, Nº candidatos: 10, Tolerancia: estricta

Y obtenemos un modelo de aptitudes mucho más manejable (solo aptitudes con un mínimo de 7 ocurrencias):

Aptitud	Frecuencia	Normalizada
Business Intelligence	11	0,056994819
Business Strategy	9	0,046632124
ChannelPartners	9	0,046632124
Cloud Computing	18	0,093264249
Enterprise Software	11	0,056994819
IT Service Management	8	0,041450777
IT Strategy	9	0,046632124
Management	7	0,03626943
Negotiation	8	0,041450777
Outsourcing	7	0,03626943
Pre-sales	13	0,067357513
Professional Services	8	0,041450777
SaaS	10	0,051813472
Sales	9	0,046632124
Sales Enablement	7	0,03626943
Sales Management	10	0,051813472
Sales Operations	9	0,046632124
Software Industry	11	0,056994819
SolutionSelling	10	0,051813472
TeamLeadership	9	0,046632124
	193	1

Y la lista de posibles candidatos, más concreta (solo 10 candidatos):

Nombre	Puntuación
Miguel Angel Bustelo	141
Mingo Olmos	141
Jose Francisco Ribeiro-Duarte Seijas	136
Xavier Llobera	122
Stan Bernard	93
Ashish Sharma	91
Todd Underhill	82
Cesar Garcia Cabello	82
Mark Tappin	81
Xavier Estruga	79

La segunda secuencia de búsquedas realizada, corresponde a la aplicación de obtención de datos estadísticos de alumnos y ex alumnos de una universidad.

La búsqueda se basa en los siguientes parámetros:

Universidad: UAB

El resultado de la búsqueda nos devuelve los siguientes resultados.

Alumnos encontrados: 92. Lista de resultados ver tabla del *Anexo 2*.

5 CONCLUSIONES

La principal conclusión que me aporta el desarrollo de este trabajo, es que a pesar de los esfuerzos de las empresas que lideran el panorama de redes sociales como Facebook, Google, Pinterest, Instagram... de proteger los datos de sus usuarios, es relativamente fácil acceder a un gran número de datos de forma más o menos sencilla a través de las APIS que las mismas redes sociales nos proporcionan u otros sistemas, como la obtención de datos a través del parseo de las paginas publicas de perfil de los usuarios.

Ambos sistemas nos proporcionan una variedad de datos considerable, aunque ciertas técnicas de obtención de datos como el parseo de HTML, hace que las aplicaciones no sean robustas a posibles cambios de diseño de las paginas parseadas, y el hacerlas robustas impliquen un esfuerzo adicional de diseño e implementación.

Otra de los temas que llama la atención es la problemática que supone que el usuario defina ciertos términos a la hora de rellenar sus perfiles, es decir, la heterogeneidad de la información de la que disponemos que hace difícil la correlación de ciertos datos para obtener conclusiones a nivel estadístico.

Por otro lado, el desarrollo de este proyecto de fin de grado me ha proporcionado la posibilidad de obtener un conocimiento más profundo de ciertas técnicas de análisis y representación de datos, las problemáticas que suponen, de qué forma podemos solventarlas por lo menos a nivel teórico.

Como línea futura a este trabajo se podrían implementar las soluciones analizadas a las problemáticas de la heterogeneidad de datos y a las limitaciones en los resultados obtenidos debido a las restricciones de licenciamiento, obteniendo de esta forma un sistema de obtención de datos más completo y robusto.

AGRADECIMIENTOS

Quiero dar las gracias a mi tutor de proyecto, Ramon-Baldrich, por su excelente labor como docente, por el apoyo, paciencia y ayuda que me ha prestado en la realización de este proyecto y que me ha permitido culminar con éxito mis estudios. Muchas gracias.

BIBLIOGRAFÍA

- [1] Ron Bowes. Publican torrent con datos de 100 millones de usuarios de Facebook [en línea]. Bitelia.com. 28 de julio de 2010. <http://bitelia.com/2010/07/publican-torrent-con-datos-de-100-millones-de-usuarios-de-facebook>
- [2] Schwartz HA, Eichstaedt JC, Kern ML, Dziurzynski L, Ramones SM, et al. (2013) Personality, Gender, and Age in the Language of Social Media: The Open-Vocabulary Approach. PLoS ONE

8(9): e73791. doi:10.1371/journal.pone.0073791 [en línea]. <http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0073791>

- [3] LinkedIn Developers. Información API REST de LinkedIn [en línea]. <http://developer.linkedin.com/rest>
- [4] LinkedIn Developers. People Search API [en línea]. <http://developer.linkedin.com/documents/people-search-api>
- [5] LinkedIn Developers. Authentication [en línea]. <http://developer.linkedin.com/documents/authentication>
- [6] Karan Balkar. Tutorial #57: Connect to LinkedIn using Java and OAuth [en línea]. karanbalkar.com. 18 de septiembre de 2013. <http://karanbalkar.com/2013/09/tutorial-57-connect-to-linkedin-using-java-and-oauth/>
- [7] LinkedIn. Información licenciamientos de LinkedIn [en línea]. <http://www.linkedin.com/mnyfe/subscriptionv2>
- [8] LinkedIn Developers. Vetted API Access [en línea]. <https://developer.linkedin.com/blog/vetted-api-access>
- [9] LinkedIn API Console | Learn, test, and debug the LinkedIn API [en línea]. <https://apigee.com/embed/console/linkedin>
- [10] Real Academia Española <http://lema.rae.es/drae/?val=histograma>

ANEXOS

Anexo 1

	Estricta	Laxa	Continua 2.0	Continua 1.0	Continua 0.5
Nombre	Puntuación	Puntuación	Puntuación	Puntuación	Puntuación
Jorge G. Moreno	32	22	24	32	56
Karen Firkser	30	21	23	30	52
Miguel Angel Bustelo	30	19	21	30	56
Julian Ruiz Gonzalez	28	20	22	28	48
Sathish Chandran Kannoli	28	20	22	28	48
Vivek A.S	27	20	22	27	45
Xavier Llobera	27	19	21	27	47
Pere Porta Lloret	25	17	19	25	45
Shashikant Shetty	23	18	19	23	35
Carles Pérez Amador	21	15	16	21	35
Shane Tunnell	21	15	17	21	37
Sari Traih	20	15	16	20	32
Stan Bernard	20	14	16	20	36
Jose Miguel Arroyo	19	13	14	19	33
Ashish Sharma	18	12	13	18	32
Peter Lin	18	11	13	18	36
Albert Trullols	17	13	14	17	27
Alberto Valero	17	12	13	17	29
Gary Foweraker	17	12	13	17	29
Sergio Sanchez	16	10	12	16	32
Alfredo Moliner	15	10	11	15	27
Daymond Cox	14	10	11	14	24
Joseph Rogel	14	10	11	14	24
Manel Borràs Borràs	14	10	11	14	24
Miquel Soler	14	10	11	14	24
Sthony Amadei	14	11	11	14	20
Christopher Woods	13	10	10	13	19
Harish Radhakrishnan	13	9	10	13	23
Mariano Cuadrelli	13	11	11	13	17
Robert Lawton B.A (Hons)	13	11	11	13	17
Mario Fernandez Dolado, PMP	12	8	9	12	22
Prashanth Upadrashta	12	10	10	12	16
Chhavi M Tandon	11	9	9	11	15
Francis Szumski	11	8	9	11	19
Gilroy Michael	11	9	9	11	15
Griselda Montiel	11	8	8	11	17
Marcelo Schiavo	11	8	9	11	19
Paul Mehrer	11	9	9	11	15
Rafael Moran	11	8	9	11	19
Ricardo Benitez	11	8	8	11	17
Cristóbal Aguirre	10	9	9	10	12
Kritagya Sharma	10	7	7	10	16
Raul Bueno	10	8	8	10	14
Vicente Rodriguez Eguibar	10	7	8	10	18
Carl Grover	9	7	8	9	15
ChiaraArch (LION)	9	8	8	9	11
Cristina Pereira	9	8	8	9	11
Lisa Clontz	9	6	7	9	17
Ian Calvert	8	7	7	8	10
Abdulrahman Shaheen	7	4	5	7	15
Dietmar Tietz	7	6	6	7	9
Federico Garcia Schierbeek	7	6	6	7	9
Francisco Gonçalves	7	4	5	7	15
Phong Vo	6	4	4	6	10
Adrian Santangelo	5	2	3	5	13
Alejandro Rios Blanco	5	4	4	5	7
Josep Tarruella	5	4	4	5	7
Oriol Valls Anton	5	4	4	5	7
Pablo Junco	5	5	5	5	5
Santiago Bonet	5	3	4	5	11
Sergio Almonacid	5	3	4	5	11
Sergio Romero	5	3	4	5	11
Vamsi Krishna	5	4	4	5	7
Claudio Cossio	4	4	4	4	4
Divya Lavanva	4	3	3	4	6
Haidar Shbeeb	4	3	3	4	6
Will Byrd	4	3	3	4	6
Ehikiova Agbakhamen	3	1	2	3	9
Gavin Smith	3	3	3	3	3
Prashanth Pattabiraman	3	3	3	3	3
Roger CARHUATOCIO	3	2	2	3	5
Ronald Heerema	3	3	3	3	3
Alberto Robles	2	2	2	2	2
Pablo Rodriguez Madroño	2	1	1	2	4
Renjith Mathew	2	1	1	2	4
Richard Kujan	2	2	2	2	2
Ruben Prida	2	1	1	2	4
Dani Carballo	1	1	1	1	1
Dharmender Kamat	1	1	1	1	1
Ramon Gonzalez	1	1	1	1	1

Anexo 2

¿Que aptitudes tienen?		¿Donde viven?		¿En que sector?	
Business Strategy	16	Barcelona Area, Spain	80	Information Technology and Services	11
Social Media	12	Spain	6	Marketing and Advertising	6
Spanish	12	Madrid Area, Spain	2	Human Resources	6
Negotiation	12	Alacant Area, Spain	1	Management Consulting	5
Marketing Strategy	10		1	Computer Software	5
English	10	United Kingdom	1	Internet	4
Web 2.0	9	Dallas/Fort Worth Area	1	Utilities	4
Project Management	8			Public Relations and Communications	3
Social Media Marketing	8			Telecommunications	3
Online Marketing	8			Program Development	2
Team Management	8			Computer Hardware	2
Team Leadership	8			Hospital & Health Care	2
Marketing Communications	8			Health, Wellness and Fitness	2
Digital Marketing	7			Financial Services	2
Business Development	6			Pharmaceuticals	2
Marketing	6			Events Services	2
Change Management	6			Education Management	2
Entrepreneurship	6			Government Administration	2
Public Relations	6			Chemicals	2
Microsoft Office	6			Research	2
Pre-sales	5			Wireless	1
Management	5			Law Practice	1
SEO	5			E-Learning	1
Online Advertising	5			Translation and Localization	1
Management Consulting	5			Professional Training & Coaching	1
IT Management	4			Airlines/Aviation	1
ITIL	4			Leisure, Travel & Tourism	1
Outsourcing	4			Accounting	1
IT Strategy	4			Media Production	1
Community Management	4			Real Estate	1
Software Project Management	4			Primary/Secondary Education	1
Strategic Planning	4			International Affairs	1
Sales Management	4			Legal Services	1
Corporate Communications	4			Online Media	1
Recruiting	4			Biotechnology	1
IT Service Management	3			Insurance	1
Cloud Computing	3			Facilities Services	1
Social Networking	3				
SEM	3				
E-Learning	3				
Healthcare	3				
Microsoft SQL Server	3				
Business Planning	3				
New Business Development	3				
E-commerce	3				
Training	3				
Strategic Consulting	2				
Service Delivery	2				
Cross-functional Team Leadership	2				
Public Health	2				
Solution Selling	2				
IT Outsourcing	2				
Virtualization	2				
VMware	2				
Data Center	2				
Managed Services	2				
Blogging	2				
Leadership	2				
XML	2				
Java	2				
JavaScript	2				
Databases	2				
Business Intelligence	2				
ERP	2				
Account Management	2				
Strategy	2				
Team Building	2				
B2B	2				
Start-ups	2				
Persona	2				
Teaching	2				
Healthcare Information Technology	2				
Event Management	2				
Teamwork	2				
Coaching	2				
Customer Service	2				
Export	2				
Market Research	2				
International Sales	2				
Excel	2				
PowerPoint	2				
HR Consulting	2				
Performance Management	2				
Organizational Development	2				
Technical Recruiting	2				

Molecular Biology	2					Data Warehousing	1				
Biotechnology	2					Software Industry	1				
Bioinformatics	2					ASP.NET MVC	1				
Human Resources	2					C#	1				
Public Sector	1					jQuery	1				
IT Transformation	1					HTML	1				
PRINCE2	1					CSS3	1				
Project Delivery	1					WinForms	1				
Sales Process	1					SVG	1				
Result Oriented	1					Python	1				
Customer Satisfaction	1					PHP4/5	1				
Sales Presentations	1					Business Analysis	1				
Management by Objectives	1					Product Management	1				
Storage	1					User Experience	1				
SaaS	1					Usability Engineering	1				
HP	1					Web Development	1				
Enterprise Software	1					Usability	1				
Business Alliances	1					Trading	1				
Channel Partners	1					Trading Systems	1				
WordPress	1					Trading System Development	1				
Multimedia	1					TradeStation	1				
Google Analytics	1					ninjatrader	1				
Web Analytics	1					Trading Strategies	1				
eHealth	1					International Relations	1				
mHealth	1					Venture Capital	1				
Healthcare Management	1					French	1				
Service Management	1					Contract Negotiation	1				
IT Operations	1					Contract Law	1				
Sla	1					Public/private Partnerships	1				
Spring, J2EE	1					Drafting Agreements	1				
Hibernate, JPA	1					Migration Law	1				
Oracle 11g, MySQL	1					NGOs	1				
Sonar, Hudson	1					International NGOs	1				
PMD, Nexus	1					Oncology	1				
Flex, JSP	1					Cancer	1				
XML, XSD	1					Clinical	1				
SQL Server, Access	1					Metodologías participativas en la enseñanza. Formación de equipos de profesores	1				
Windows 7, Vista, XP	1					Moodle	1				
Android 2.3, 4.2	1					Distance Learning	1				
Linux Ubuntu, RedHat, Suse	1					Educational Technology	1				
Apache-Tomcat, Jetty	1					Instructional Design	1				
WebLogic, Oracle	1					Blended Learning	1				
Jonas, Jboss	1					HL7	1				
JBPM, W4	1					Clinical Research	1				
FED, Flow	1					Data Analysis	1				
Sterling, Staffware	1					Hospitals	1				
SCRUM (Agile Methodology)	1					Balanced Scorecard	1				
RUP, BPR	1					Strategic Sales	1				
Alfresco, OpenCMS 7	1					Sales Execution	1				
SharePoint, FatWire, Documentum	1					Communication	1				
TCP/IP, UDP, VPN	1					Project Planning	1				
HTTP, HTTPS, SSL	1					Música, guitarra	1				
SMTP, POP, MIME	1					Television	1				
FTP, TELNET	1					Entertainment	1				
MS Office 2007, 2003	1					Fashion Design	1				
Open Office 3.1, 2.4	1					Website Translation	1				
JBuilder, Kawa, TOAD	1					Software GILT	1				
MVC, Struts, EJB	1					Website Localization	1				
ASP, .NET, C++, PHP	1					Language Services	1				
Perl, JavaScript	1					Translation Services	1				
HTML, DHTML	1					Ethnomusicology	1				
Software Development	1					Music History	1				
Java Enterprise Edition	1					Computer Music	1				
Tomcat	1					International Project Management	1				
Oracle	1					Language Technology	1				
MySQL	1					SAP	1				
JSP	1					E-commerce SEO	1				
Weblogic	1					E-business	1				
Subversion	1					Translation	1				
Govern de TI	1					Localization	1				
Direcció de Sistemes d'Informació	1					Internet Strategy	1				
Governance	1					Glossaries	1				
Knowledge Management	1					Annual Reports	1				
Direcció de projectes	1					Catalogue Production	1				
PMP	1					Accounting	1				
ITIL Certified	1					Media Relations	1				
Gestió del coneixement	1					Strategic Communications	1				
Gestió d'organitzacions públiques	1					Internal Communications	1				
vSphere	1					Journalism	1				
vCenter Server	1					Press Releases	1				
VCP 3	1					Event Planning	1				
Windows Server	1					Catalan	1				
Microsoft Certified Professional	1					Brand Management	1				
Microsoft Exchange	1					Import	1				
Active Directory	1					Strategy Implementation	1				
SharePoint	1					Director Comercial & Expansión	1				
VMware Infrastructure	1					Director de Marketing	1				
VDI	1					Consultor en Estrategias Empresariales	1				
ISO 27001	1					Trade Marketing	1				
Telecommunications	1					Marketing Management	1				
SQL	1					Advertising	1				
Linux	1					Responsable, complicitadora, ordenada, puntual, treballadora	1				
VMware ESX	1					M'agrada aprendre i m'adapto amb facilitat als	1				
NetApp	1										
Sonicwall	1										
Citrix	1										
Netezza	1										

canvis						PCR	1				
Capaç d'assolir nous reptes	1					Cell Culture	1				
M'implico en la feina	1					Genomics	1				
Capacitat de treballar en equip i de forma autònoma	1					Cell Biology	1				
Sóc de caràcter seriós i tracto tothom amb respecte, simpatia i afabilitat	1					MBA EADA 1995-1997	1				
Adhesives	1					Sourcing	1				
Product Development	1					Interviews	1				
Coatings	1					CRM	1				
Powder Coating	1					Talent Acquisition	1				
Paint	1					Networking	1				
Additives	1					Formación en habilidades directivas	1				
CPNL - Certificat oficial nivell C de català	1					Formación en competencias de Coaching	1				
Reporting Systems	1					PMO	1				
Reporting & Analysis	1					Executive Coaching	1				
Business Reporting	1					Personnel Management	1				
SAP BI/BW	1					EADA - Escuela de alta dirección y administración	1				
Creating Processes	1					Formanet - master técnicas de ventas	1				
Purchasing Processes	1					Esama - master	1				
SAP R/3	1					Technical Training	1				
SAP ERP	1					technical training sound during 3 years	1				
SAP MM	1					technical training led technologies	1				
Vendor Management	1					AV	1				
SAP SRM	1					Video Conferencing	1				
Ariba	1					Sales	1				
User Acceptance Testing	1					Electronics	1				
Visio	1					Sales Operations	1				
Word	1					Troubleshooting	1				
Outlook	1					Digital Signage	1				
Lotus Notes	1					Product Marketing	1				
Photoshop	1					Project Coordination	1				
Illustrator	1					Hardware	1				
InDesign	1					Desarrollo Estratégico de Corporaciones	1				
Dreamweaver	1					Desarrollo e implantación Planes de Marketing y Comerciales	1				
Photography	1					Análisis de Eficiencia de Inversiones de Marketing	1				
Mac	1					Desarrollo de Cuadros de Mando Operativos	1				
Windows	1										
Animació i contacontes	1										
Expressió corporal i teatre	1										
Docent de resolució de conflictes	1										
Títol de Monitora de Lleure	1										
Títol de Directora de Lleure	1										
Postgrau en Psicologia de l'Educació (ISEP)	1										
Educadora d'Estudi Assistit	1										
Docent de cursos de monitors i directors	1										
Col·legiada pel COPC	1										
Psicòloga Clínica	1										
Microsoft Excel	1										
TI Recruitment	1										
Facebook	1										
Contract Recruitment	1										
LinkedIn	1										
E-commerce Recruitment	1										
iOS	1										
iOS development	1										
Sistema Red	1										
contrat@	1										
A3	1										
Labor Relations	1										
Corporate Law	1										
Personal Finance	1										
Profile	1										
Carerix	1										
Google+	1										
Twitter	1										
Yammer	1										
Skype	1										
Blogger	1										
Field Hockey	1										
Football	1										
Protein Expression	1										
Protein Purification	1										
Protein Engineering	1										
Bacterial Culture	1										
Membrane Proteins	1										
Fermentation	1										
Western Blotting	1										
FPLC	1										
Structural Bioinformatics	1										
X-ray crystallography	1										
Mutagenesis	1										
Site-directed Mutagenesis	1										
GFP	1										
SPSS	1										
Public Speaking	1										
Budgets	1										
Strategic Human Resources Leadership	1										
Human Resource Planning	1										
Global Human Resources Management	1										
Research	1										
Genetics	1										
Science	1										
Lifesciences	1										
Biochemistry	1										