



UNIVERSIDAD TÉCNICA ESTATAL DE QUEVEDO
FACULTAD DE CIENCIAS DE LA INGENIERÍA
CARRERA INGENIERÍA EN SISTEMAS

Proyecto de Investigación previo
a la obtención del título de
Ingeniero en Sistemas.

Título del Proyecto de Investigación:

**“VISUALIZACIÓN DE LAS HABILIDADES PROFESIONALES MEDIANTE EL
ANÁLISIS DE ANUNCIOS DE SITIOS WEB DE EMPLEOS”**

Autores:

Ariel Oswaldo Fernández Loor

Dayanna Lizzet Velez Araujo

Director de Proyecto de Investigación:

Ing. Ariosto Vicuña Pino

Quevedo – Los Ríos - Ecuador

2022

DECLARACIÓN DE AUTORÍA Y CESIÓN DE DERECHOS

Yo, **Ariel Oswaldo Fernández Loor** y **Dayanna Lizzet Velez Araujo**, declaramos que la investigación aquí descrita es de nuestra autoría; que no ha sido previamente presentado para ningún grado o calificación profesional; y, que he consultado las referencias bibliográficas que se incluyen en este documento.

La Universidad Técnica Estatal de Quevedo, puede hacer uso de los derechos correspondientes a este documento, según lo establecido por la Ley de Propiedad Intelectual, por su Reglamento y por la normatividad institucional vigente.

f.



Ariel Oswaldo Fernández Loor

C.C. 120702810-9

f.

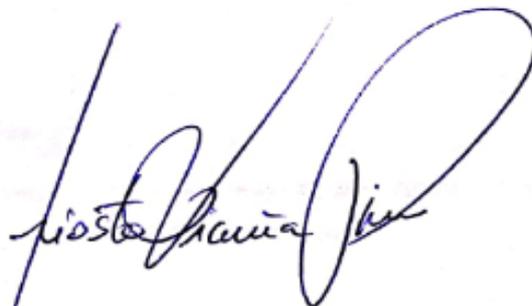


Dayanna Lizzet Velez Araujo

C.C. 120612420-6

CERTIFICACIÓN DE CULMINACIÓN DEL PROYECTO DE INVESTIGACIÓN

El suscrito ING. ARIOSTO VICUÑA PINO, Docente de la Universidad Técnica Estatal de Quevedo, certifica que los estudiantes ARIEL OSWALDO FERNÁNDEZ LOOR y DAYANNA LIZZET VELEZ ARAUJO, realizaron el Proyecto de Investigación de grado titulado “VISUALIZACIÓN DE LAS HABILIDADES PROFESIONALES MEDIANTE EL ANÁLISIS DE ANUNCIOS DE SITIOS WEB DE EMPLEOS”, previo a la obtención del título de INGENIERO EN SISTEMAS, bajo mi dirección, habiendo cumplido con las disposiciones reglamentarias establecidas para el efecto.



ING. Ariosto Vicuña Pino M.Sc.

DIRECTOR DEL PROYECTO DE INVESTIGACIÓN

CERTIFICACIÓN DEL REPORTE DE LA HERRAMIENTA DE PREVENCIÓN DE COINCIDENCIA Y/O PLAGIO ACADÉMICO

INFORME DEL DIRECTOR DE PROYECTO DE INVESTIGACIÓN SOBRE EL SISTEMA URKUND

Yo, Ing. Ariosto Eugenio Vicuña Pino M.Sc., en calidad de director del Proyecto de Investigación titulado: **“VISUALIZACIÓN DE LAS HABILIDADES PROFESIONALES MEDIANTE EL ANÁLISIS DE ANUNCIOS DE SITIOS WEB DE EMPLEOS”**, me permito manifestar a usted y por medio del Consejo Académico de Facultad lo siguiente:

Que, el Sr. FERNÁNDEZ LOOR ARIEL OSWALDO y la Srta. VELEZ ARAUJO DAYANNA LIZZET, egresados de la Facultad de Ciencias de la Ingeniería, han cumplido con las correcciones pertinentes, e ingresado su proyecto de investigación al sistema URKUND, tengo a bien certificar la siguiente información sobre el informe del sistema anti-plagio con un porcentaje del 1 % como se observa en la siguiente figura:



Document Information

Analyzed document	Proyecto de investigación - FERNANDEZ-VELEZ docx (D144821165)
Submitted	2022-09-26 06:21:00
Submitted by	
Submitter email	ariel.fernandez2017@uteq.edu.ec
Similarity	1%
Analysis address	avicuna.uteq@analysis.orkund.com

ING. Ariosto Vicuña Pino M.Sc.
DIRECTOR DEL PROYECTO DE INVESTIGACIÓN



UNIVERSIDAD TÉCNICA ESTATAL DE QUEVEDO
FACULTAD DE CIENCIAS DE LA INGENIERÍA
CARRERA INGENIERÍA EN SISTEMAS

PROYECTO DE INVESTIGACIÓN

Título:

**“VISUALIZACIÓN DE LAS HABILIDADES PROFESIONALES MEDIANTE EL
ANÁLISIS DE ANUNCIOS DE SITIOS WEB DE EMPLEOS”**

Presentado al Consejo Directivo como requisito previo a la obtención del título de Ingeniero en Sistemas.

Aprobado por:

PRESIDENTA DEL TRIBUNAL
Ing. Nancy Rodríguez Gavilanes M.Sc.

MIEMBRO DEL TRIBUNAL
Ing. Giojan Aguirre Pérez M.Sc.

MIEMBRO DEL TRIBUNAL
PhD. Cristian Zambrano Vega

QUEVEDO - LOS RIOS - ECUADOR

AGRADECIMIENTO

Quiero agradecer a Dios por haberme dado la sabiduría necesaria para permitir haber llegado hasta estas instancias de mi vida. Agradezco también a mis padres y hermanos, quienes me han brindado su apoyo incondicional para mi vida.

Además, quiero agradecer a mi familia por el apoyo incondicional que me han brindado. También a mi compañera de tesis, por su entrega y comprensión durante el desarrollo del proyecto.

En general a todos los docentes que han sido partícipes de mi formación profesional, por compartir sus conocimientos académicos y personales que sin duda me servirán en el futuro. Finalmente, quiero agradecer a mis compañeros y amigos con quienes he compartido momentos agradables.

Ariel Oswaldo Fernández Loor

Quiero primero agradecer a Dios y a la Virgen por darme la fuerza para superar cada obstáculo que se ha presentado en el transcurso de mi etapa universitaria. Agradezco también a mi mamá, quien ha sido el pilar fundamental en mi vida, la cual con sus consejos ha sido mi guía y motivación para luchar por mis sueños.

Además, quiero agradecer a mi familia por el apoyo incondicional que me han brindado en el transcurso de mi etapa universitaria. También a mi compañero de tesis, por su entrega y comprensión durante el desarrollo del proyecto.

A todos los docentes que han sido partícipes de mi formación profesional, por compartir sus conocimientos académicos y personales que sin duda me servirán en el futuro. Finalmente, quiero agradecer a mis compañeros y amigos con quienes he compartido momentos agradables.

Dayanna Lizzet Velez Araujo

DEDICATORIA

A mis padres por aconsejarme y apoyarme en cada paso que doy para cumplir cada meta que me propongo. A mis hermanos que me han enseñado a salir adelante. Gracias por preocuparse y por compartir sus vidas; gracias por formar parte de este importante logro.

A todas mis amistades que de una u otra manera me han brindado su apoyo, consejos y palabras de aliento para completar esta meta. Y claro, a Dios por permitirme llegar a este punto, fin de una bonita etapa, pero el comienzo de otra llena de retos y nuevas aspiraciones.

Ariel Oswaldo Fernández Loor

A Dios por guiarme y bendecirme en cada paso que doy para cumplir mis sueños.

A mi mamá María Nela Araujo, por ser mi mejor amiga y apoyarme en cada etapa de mi vida. Por haberme inculcado valores, por cada sonrisa que compartí con ella, por estar ahí para secar mis lágrimas y por siempre confiar en mi potencial. Por todo el esfuerzo que ha hecho por sacarme adelante, nunca me cansaré de agradecerle.

A mi familia materna, por estar siempre para mí en los buenos y malos momentos. A todas las personas que llegaron a mi vida, gracias por sus palabras de aliento y su apoyo incondicional.

Dayanna Lizzet Velez Araujo

RESUMEN Y PALABRAS CLAVES

Los avances tecnológicos han causado el surgimiento de nuevas ocupaciones, por lo tanto, las universidades deben actualizar sus currículos para que sus estudiantes adquieran las habilidades necesarias para enfrentarse al mundo laboral. Este presente estudio combina la técnica del Web Scraping con el Análisis de Redes Sociales para identificar las habilidades con mayor demanda en distintas áreas de conocimiento. Para el efecto, se construyeron siete redes de habilidades diferentes en base a las habilidades profesionales requeridas para cada área de conocimiento según la información disponible en los anuncios de empleo. Con la interpretación y análisis de las métricas (centralidad, cercanía y densidad) de cada una de las redes se determinaron aquellas habilidades con mayor influencia en el mercado laboral actual. El análisis de las métricas redes de habilidades profesionales determinaron que las habilidades con mayor demanda se encuentran relacionadas con el desarrollo de software y manejo de datos, siendo así SQL, Java y Javascript las habilidades profesionales con mayor número de enlaces según el indicador de centralidad. Demostrando que las habilidades con mayor demanda se encuentran relacionadas con el desarrollo de software y el manejo de datos.

Palabras claves: Web Scraping, SNA, habilidades, currículos.

ABSTRACT AND KEYWORDS

Technological advances have caused the emergence of new occupations; therefore, universities must update their curricula so that their students acquire the necessary skills to face the working world. This study combines the Web Scraping technique with Social Network Analysis to identify the skills most in demand in different areas of knowledge. For this purpose, seven different skill networks were constructed based on the professional skills required for each area of knowledge according to the information available in job advertisements. With the interpretation and analysis of the metrics (centrality, closeness and density) of each of the networks, those skills with the greatest influence in the current labor market were determined. The analysis of the metrics of the professional skills networks determined that the skills with the highest demand are related to software development and data management, with SQL, Java and Javascript being the professional skills with the highest number of links according to the centrality indicator. Demonstrating that the skills with the highest demand are related to software development and data management.

Keywords: Web Scraping, SNA, skills, curricula.

TABLA DE CONTENIDO

DECLARACIÓN DE AUTORÍA Y CESIÓN DE DERECHOS	ii
CERTIFICACIÓN DE CULMINACIÓN DEL PROYECTO DE INVESTIGACIÓN.....	iii
CERTIFICACIÓN DEL REPORTE DE LA HERRAMIENTA DE PREVENCIÓN DE COINCIDENCIA Y/O PLAGIO ACADÉMICO	iv
AGRADECIMIENTO	vi
DEDICATORIA.....	vii
RESUMEN Y PALABRAS CLAVES.....	viii
ABSTRACT AND KEYWORDS	ix
TABLA DE CONTENIDO	x
ÍNDICE DE TABLAS.....	xiv
ÍNDICE DE ILUSTRACIONES	xv
INDICE DE ANEXOS	xvii
CÓDIGO DUBLIN	xviii
INTRODUCCIÓN.....	1
CAPÍTULO I.....	3
CONTEXTUALIZACIÓN DE LA INVESTIGACIÓN	3
1.1.Problema de investigación.....	4
1.1.1.Planteamiento del problema	4
1.1.2.Formulación del problema.....	5
1.1.3.Sistematización del problema.....	5
1.2.Objetivos.....	6
1.2.1.Objetivo General.....	6

1.2.2.Objetivos Específicos	6
1.3.Justificación	6
CAPÍTULO II.....	8
FUNDAMENTACIÓN TEÓRICA DE LA INVESTIGACIÓN	8
2.1.Marco Conceptual.....	9
2.1.1.Web Scraping	9
2.1.2.Legalidad y ética del Web Scraping.....	10
2.1.3.Lenguajes de programación populares para Web Scraping.....	11
2.1.3.1.Web Scraping con Java	11
2.1.3.2.Web Scraping con Python	11
2.1.3.3.Web Scraping con PHP	11
2.1.4.Librerías y APIs para Web Scraping	11
2.1.5.Análisis de Redes Sociales (<i>Social Networks Analysis, SNA</i>).....	12
2.1.6.Elementos básicos de una red	13
2.1.7.Métricas en el SNA	13
2.1.7.1.Grado de intermediación (<i>betweenness</i>).....	13
2.1.7.2.Grado de cercanía (<i>closeness</i>)	13
2.1.7.3.Grado de centralidad (<i>centrality</i>).....	14
2.1.8.Software para SNA	14
2.1.9.Definición de habilidades (<i>skills</i>)	15
2.1.10.Clasificación de habilidades	15
2.1.10.1.Habilidades duras	15
2.1.10.2.Habilidades blandas.....	15
2.1.11.E-recruitment	16
2.1.12.Portales de empleo.....	16
2.1.13.Ejemplo de portales de empleo en línea	16
2.1.13.1.Indeed	17

2.1.13.2.LinkedIn	17
2.1.14.Definición de currículum.....	17
2.2.Marco referencial.....	19
2.3.Marco Legal.....	20
 CAPÍTULO III	 22
 METODOLOGÍA DE LA INVESTIGACIÓN	 22
3.1.Localización.....	23
3.2.Tipo de investigación	23
3.3.Método de investigación.....	23
3.3.1.Método analítico	23
3.3.2.Método deductivo	24
3.4.Fuentes de recopilación de información	24
3.5.Diseño de la investigación.....	24
3.5.1.Primer Etapa: Revisión de fuentes de datos.....	25
3.5.2.Segunda Etapa: Recopilación de datos	26
3.5.3.Tercera Etapa: Limpieza y preprocesamiento	27
3.5.4.Cuarta Etapa: Construcción del modelo SNA	28
3.5.5.Quinta Etapa: Análisis y visualización de datos.....	29
3.5.6.Sexta Etapa: Evaluación de usabilidad de la herramienta	29
3.6.Instrumentos de investigación	30
3.7.Tratamiento de datos	30
3.8.Recursos humanos y materiales.....	30
3.8.1.Recursos humanos	30
3.8.2.Equipos y materiales.....	31
3.8.2.1.Hardware	31
3.8.2.2.Software	31

CAPÍTULO IV	32
RESULTADOS Y DISCUSIÓN	32
4.1.Resultados de la extracción de los anuncios de empleo mediante la técnica de <i>Web Scraping</i> y limpieza de datos.....	33
4.2.Resultado de las redes de habilidades profesionales aplicando SNA.....	33
4.3.Resultados de la herramienta web informativa y del reporte de habilidades profesionales requeridas en cada área del conocimiento según los anuncios de empleo	44
4.4.Resultados de la evaluación de usabilidad de la herramienta web	46
4.5.Discusión	50
CAPÍTULO V	51
CONCLUSIONES Y RECOMENDACIONES	51
5.1.Conclusiones.....	52
5.2.Recomendaciones	54
CAPÍTULO VI	55
BIBLIOGRAFÍA	55
CAPÍTULO VII.....	60
ANEXOS	60

ÍNDICE DE TABLAS

Tabla 1: Librerías y APIs para <i>Web Scraping</i>	12
Tabla 2 Software para SNA.....	14
Tabla 3: Descripción de variables	27
Tabla 4: Hardware utilizado para el desarrollo del proyecto.....	31
Tabla 5: Software utilizado para el desarrollo del proyecto	31
Tabla 6: Cantidad de anuncios de empleo y habilidades por área de conocimiento	33
Tabla 7: Cantidad de atributos por elemento de la Red SNA “habilidades”	33
Tabla 8: Indicadores de la red general de habilidades.....	35
Tabla 9: Indicadores para la red de Arquitectura	36
Tabla 10: Indicadores para la red de Ingeniería Eléctrica	37
Tabla 11: Indicadores para la red de Ingeniería en Sistemas	38
Tabla 12: Indicadores para la red de Ingeniería en Software	39
Tabla 13: Indicadores para la red de Ingeniería Ambiental.....	40
Tabla 14: Indicadores para la red de Ingeniería en Telemática.....	41
Tabla 15: Indicadores para la red de Ingeniería Mecánica.....	42
Tabla 16: Indicadores para la red de provincias	43
Tabla 17: Respuestas a las afirmaciones individuales de la escala SUS	47
Tabla 18: Valoración SUS.....	48
Tabla 19: Rango de usabilidad	48

ÍNDICE DE ILUSTRACIONES

Ilustración 1: Fases del Web Scraping	10
Ilustración 2: Malla curricular de Ingeniería de Software de la UTEQ	18
Ilustración 3: Ubicación del Campus “La María”	23
Ilustración 4: Etapas de la metodología de desarrollo	25
Ilustración 5: Proceso para el Web Scraping	26
Ilustración 6: Etapa de clasificación de habilidades profesionales.....	28
Ilustración 7: Red de habilidades profesionales requeridas en distintas carreras de la FCI	34
Ilustración 8: Nube de palabras de habilidades profesionales con mayor demanda	35
Ilustración 9: Red SNA de habilidades requeridas en la carrera de Arquitectura	35
Ilustración 10: Nube de palabras de habilidad demandadas en la carrea de Arquitectura ..	36
Ilustración 11: Red SNA de habilidades requeridas en la carrera de Ingeniería Eléctrica ..	36
Ilustración 7: Nube de palabras de habilidad demandadas en la carrea de Ingeniería Eléctrica	37
Ilustración 13: Red SNA de las habilidades requeridas en la carrera de Ingeniería en Sistemas	37
Ilustración 14: Nube de palabras de habilidad demandadas en la carrea de Ingeniería en Sistemas.....	38
Ilustración 15: Red SNA de habilidades requeridas en la carrera de Ingeniería en Software	38
Ilustración 16: Nube de palabras de habilidad demandadas en la carrea de Ingeniería en Software.....	39
Ilustración 17: Red SNA de habilidades requeridas en la carrera de Ingeniería Ambiental	39
Ilustración 18: Nube de palabras de habilidad demandadas en la carrea de Ingeniería Ambiental	40
Ilustración 19: Red SNA de habilidades requeridas en la carrera de Ingeniería en Telemática	40
Ilustración 20: Nube de palabras de habilidad demandadas en la carrea de Ingeniería en Telemática	41
Ilustración 21: Red SNA de habilidades requeridas en la carrera de Ingeniería Mecánica.	41
Ilustración 22: Nube de palabras de habilidad demandadas en la carrea de Ingeniería Mecánica.....	42
Ilustración 23: Red SNA de las provincias con mayor disponibilidad de ofertas	43

Ilustración 24: Dashboard de la herramienta web	44
Ilustración 25: Gráfico de habilidades y anuncios extraídos	44
Ilustración 26: SNA Habilidades requeridas	45
Ilustración 27: Interfaz para la visualización de la red de habilidades	45
Ilustración 28: Interfaz de usuario	46
Ilustración 29: Rango de usabilidad	48

INDICE DE ANEXOS

Anexo 1: Modelo Físico de la Base de Datos para generar el reporte de habilidades	61
Anexo 2: Documentación del software	61
Anexo 2.1. Casos de Uso.....	61
Anexo 2.2. Caso de uso extendido – Visualizar dashboard.....	62
Anexo 2.3. Caso de uso extendido – Visualizar red de habilidades	62
Anexo 2.4. Caso de uso extendido – Visualizar red de localidades	63
Anexo 2.5. Caso de uso extendido – Generar reporte	64
Anexo 3: Reporte generado desde la aplicación web	65
Anexo 4: Manual de usuario de la aplicación web	66
Anexo 4.1. Dashboard de la aplicación	66
Anexo 4.2. Visualización de las redes de habilidades profesionales por carrera	66
Anexo 4.4. Visualización de redes clasificadas por localización	67
Anexo 4.5. Generación del reporte	68

CÓDIGO DUBLIN

Título:	Visualización de las habilidades profesionales mediante el análisis de anuncios de sitios web de empleos			
Autores:	Fernández Loor, Ariel Oswaldo; Velez Araujo, Dayanna Lizzet			
Palabras claves:	<i>Web Scraping</i>	<i>SNA</i>	<i>habilidades</i>	<i>Currículos</i>
Fecha de publicación:	Diciembre, 2022			
Director del proyecto:	Vicuña Pino, Ariosto Eugenio			
Editorial:	Quevedo: UTEQ, 2022			
Resumen:	<p>Resumen. - Los avances tecnológicos han causado el surgimiento de nuevas ocupaciones, por lo tanto, las universidades deben actualizar sus currículos para que sus estudiantes adquieran las habilidades necesarias para enfrentarse al mundo laboral. Este presente estudio combina la técnica del Web Scraping con el Análisis de Redes Sociales para identificar las habilidades con mayor demanda en distintas áreas de conocimiento. Para el efecto, se construyeron siete redes de habilidades diferentes en base a las habilidades profesionales requeridas para cada área de conocimiento según la información disponible en los anuncios de empleo. Con la interpretación y análisis de las métricas (centralidad, cercanía y densidad) de cada una de las redes se determinaron aquellas habilidades con mayor influencia en el mercado laboral actual. El análisis de las métricas redes de habilidades profesionales determinaron que las habilidades con mayor demanda se encuentran relacionadas con el desarrollo de software y manejo de datos, siendo así SQL, Java y Javascript las habilidades profesionales con mayor número de enlaces según el indicador de centralidad. Demostrando que las habilidades con mayor demanda se encuentran relacionadas con el desarrollo de software y el manejo de datos.</p> <p>Abstract. - Technological advances have caused the emergence of new occupations; therefore, universities must update their curricula so that their students acquire the necessary skills to face the working world. This study combines the Web Scraping technique with Social Network Analysis to identify the skills most in demand in different areas of knowledge. For this purpose, seven different skill networks were constructed based on the professional skills required for each area of knowledge according to the information available in job advertisements. With the interpretation and analysis of the metrics (centrality, closeness and density) of each of the networks, those skills with the greatest influence in the current labor market were determined. The analysis of the metrics of the professional skills networks determined that the skills with the highest demand are related to software development and data management, with SQL, Java and Javascript being the professional skills with the highest number of links according to the centrality indicator. Demonstrating that the skills with the highest demand are related to software development and data management.</p>			
Descripción:	86 hojas: dimensiones, 21 cm x 29,7 cm			
URI:				

INTRODUCCIÓN

Los anuncios laborales que se publican en los portales de empleo pueden proporcionar indicadores sobre la demanda de profesionales en distintas áreas. Los datos de los anuncios de empleo dan a conocer los cambios en la demanda de mano de obra a medida que el mercado laboral evoluciona [1]. Los avances tecnológicos han transformado la manera de cómo las personas perciben las ofertas de empleo disponibles. Esto nos induce a pensar que el mercado laboral, dadas las condiciones de nuevas profesiones y requerimientos, es variable. El mercado laboral se encuentra en constante transformación, lo que causa un seguimiento en cuanto a su evolución [2]. En consecuencia, la evolución del mercado laboral está dando lugar a la creación de nuevos empleos, por lo tanto, identificarlos es importante.

De hecho, la identificación de los requerimientos del mercado laboral tiene su influencia en el contexto universitario dado que los profesionales deben adquirir las habilidades que las empresas requieren en cada uno de sus puestos de trabajo. Analizar los datos sobre la demanda de las habilidades profesionales para clasificar los perfiles en función de las habilidades, así como analizar las áreas de aprendizaje y proporcionar correcciones en los currículos de las universidades. Teniendo en cuenta los factores asociados al impacto que genera el desarrollo de los procesos de actualización curricular, es posible indicar que la renovación de los currículos de distintas áreas del conocimiento, ha permitido revertir las restricciones de las opciones vocacionales y laborales de los estudiantes [3].

Por otro lado, actualizar los currículos busca mejorar los procesos de enseñanza en todas las áreas del conocimiento. Esto implica evaluar los currículos y el perfil de egreso, definiendo estos factores pueden identificarse áreas en las que es necesario reforzar los programas de aprendizaje para disminuir la brecha que existe entre los planes de estudio y las habilidades demandadas en el mercado laboral. En muchos casos, existe una brecha entre lo que los estudiantes aprenden en la universidad y lo que las empresas necesitan de ellos, lo que significa que se requiere un esfuerzo importante para integrar a los recién graduados en un entorno laboral [4].

Esta brecha entre las habilidades profesionales que el mercado laboral requiere y aquellas que los estudiantes poseen al momento de ingresar al campo laboral, ha llevado a las

universidades a iniciar programas para evaluar las habilidades profesionales de sus estudiantes. Aún no se conoce un modelo generalizado para determinar la brecha planteada. Este trabajo se plantea como propósito modelar las habilidades profesionales requeridas en el mercado laboral para un área de conocimiento específica mediante el uso del Social Network Analysis (SNA). El modelo será visualizado como un grafo que proporcionará información sobre las habilidades profesionales que tienen una mayor demanda dentro del ámbito laboral. De esta manera, las universidades tendrán información de la tendencia y evolución de las habilidades que las empresas requieren.

CAPÍTULO I

CONTEXTUALIZACIÓN DE LA INVESTIGACIÓN

1.1. Problema de investigación

1.1.1. Planteamiento del problema

Dentro del contexto universitario, las instituciones de educación superior buscan realizar mejoras en los procesos de enseñanza en todas las áreas del conocimiento. Esto implica analizar la demanda de habilidades en el ámbito laboral, generar conocimiento sobre las habilidades que las empresas buscan en un empleado y proporcionar correcciones en los planes de estudio [3]. Las universidades destinan recursos para obtener información sobre la evolución de los graduados en su vida profesional, haciendo uso de métodos tradicionales como encuestas y/o entrevistas. El uso de estos métodos consume recursos humanos, económicos y de tiempo para las universidades.

Por otra parte, actualmente varias empresas publican sus vacantes en sitios de empleo en los cuales añaden información sobre el perfil (conocimiento y habilidades) que buscan en un empleado [5]. Es cierto que actualmente en Internet se encuentran varios sitios que recogen las ofertas laborales y agrupan las habilidades requeridas para cada una de ellas. Sin embargo, las aplicaciones web que existen no permiten a las universidades visualizar de forma sencilla e interactiva las conexiones entre las habilidades requeridas y cómo se relacionan con distintas áreas del conocimiento de las carreras que ofertan. Además, no permiten conocer la disponibilidad de ofertas laborales en cada una de las provincias del país. Ante la gran cantidad de información es necesario aplicar técnicas de inteligencia artificial que faciliten visualizar las habilidades requeridas para los diferentes puestos de trabajo profesionales.

Aplicar técnicas de Inteligencia Artificial (IA) a la información que se obtiene de los sitios de empleo para analizar los datos de manera eficaz [6]. Existen varias técnicas disponibles, pero se debe adaptar cada técnica para obtener la mejor solución del problema. Existen modelos que relacionan las habilidades laborales con las distintas áreas del conocimiento, utilizando el método de agrupaciones (clúster). Además, las aplicaciones web no facilitan a los usuarios conocer las habilidades con mayor demanda en el mercado laboral, lo que dificulta tener una visión completa sobre la disponibilidad de ofertas laborales y que compare las habilidades que se requieren en el mercado laboral.

Varios investigadores han argumentado que una de las limitaciones de la utilización de datos de anuncios en línea, es la representación incompleta de la demanda de mano de obra [1]. Además, los estudios realizados no aportan con una aplicación web que permita visualizar los resultados que han obtenido. Por lo tanto, las universidades no pueden hacer uso de esta información para conocer las tendencias de las habilidades laborales que son requeridas en el mercado laboral actual y mejorar sus currículos académicos. Detectar la escasez de habilidades a partir de la información disponible en los anuncios de empleo en línea y analizar los datos sobre la demanda de las habilidades laborales permite clasificar los perfiles en función de las habilidades [3].

El presente proyecto de investigación intenta convertir los datos que se extraen de los anuncios de empleo en información útil para los estudiantes y las universidades. Una posible solución es usar la técnica de *Web Scraping* con Análisis de Redes Sociales (SNA, *Social Network Analysis*) para agrupar requerimientos de habilidades y ofertas de trabajo. Estas técnicas estarán contenidas dentro de una aplicación web que sirva como guía para la actualización de sus currículos, mejorando los perfiles de egreso de los estudiantes.

1.1.2. Formulación del problema

¿Cómo el Social Network Analysis y el *Web Scraping* pueden ayudar en la visualización de las relaciones de las habilidades profesionales y las áreas de conocimiento en la web de manera sencilla?

1.1.3. Sistematización del problema

¿Cómo se pueden extraer los datos de anuncios laborales de los sitios de alojamiento de empleo?

¿Qué mecanismo permite la visualización de las habilidades profesionales relacionadas con las distintas áreas del conocimiento?

¿Cómo se puede configurar las opciones necesarias para generar un reporte de las habilidades profesionales requeridas en cada área del conocimiento?

¿De qué manera se puede analizar la aplicación desarrollada para medir la usabilidad de la herramienta?

1.2. Objetivos

1.2.1. Objetivo General

Desarrollar una herramienta web para la visualización de las habilidades profesionales requeridas en las distintas áreas del conocimiento publicadas en los sitios web de ofertas de trabajo.

1.2.2. Objetivos Específicos

- Extraer los anuncios laborales del portal de empleo Indeed haciendo uso de la técnica de *Web Scraping*.
- Utilizar Social Network Analysis (SNA) para la visualización de las habilidades profesionales relacionadas con las distintas áreas del conocimiento.
- Recopilar en un reporte las habilidades profesionales requeridas en cada área del conocimiento según los anuncios de empleo.
- Analizar la aplicación web desarrollada mediante un estudio de usuarios para conocer su nivel de usabilidad.

1.3. Justificación

La educación de las universidades se encuentra en constante cambio. A causa de la evolución del mercado laboral, las universidades deben realizar ajustes curriculares para satisfacer las necesidades de las nuevas profesiones y requerimientos. En este sentido, en varios países de América Latina, han iniciado procesos de actualización en sus currículos para medir los rendimientos de los estudiantes y evaluar la calidad de la oferta de las universidades.

El Consejo de Educación Superior (CES), dispone que las universidades deben seguir un modelo educativo que permita la actualización de la gestión curricular. Así mismo integrar conocimientos que garanticen el desarrollo investigativo y la vinculación con la sociedad. Con lo mencionado las universidades deben retroalimentarse para el diseño, actualización y fortalecimiento de sus currículos.

En consecuencia, las universidades deben orientar sus currículos para adquirir conocimiento, desarrollar habilidades y competencias que aporten a los profesionales la capacidad de

competir en los sectores de su área de conocimiento. La clave para que las universidades realicen los ajustes necesarios en sus currículos se encuentra en conocer las habilidades con mayor demandada en el mercado laboral actual.

Conocer las relaciones que existen entre las distintas áreas del conocimiento profesionales y las habilidades con mayor demanda ayudará a las universidades en la toma de decisiones con respecto al cambio o actualización curricular. La aplicación web brindará la posibilidad de visualizar las habilidades con mayor demanda y las relaciones que existen entre las distintas áreas del conocimiento con las habilidades profesionales que requieren.

Permitiendo a las universidades planificar y evaluar acciones que aseguren la formación del estudiante. Realizando ajustes curriculares para atender los requerimientos de las nuevas ocupaciones que surgen con el avance tecnológico.

CAPÍTULO II

FUNDAMENTACIÓN TEÓRICA DE LA INVESTIGACIÓN

2.1. Marco Conceptual

A continuación, se describen los conceptos principales para proporcionar una mejor comprensión del tema del proyecto de investigación.

2.1.1. Web Scraping

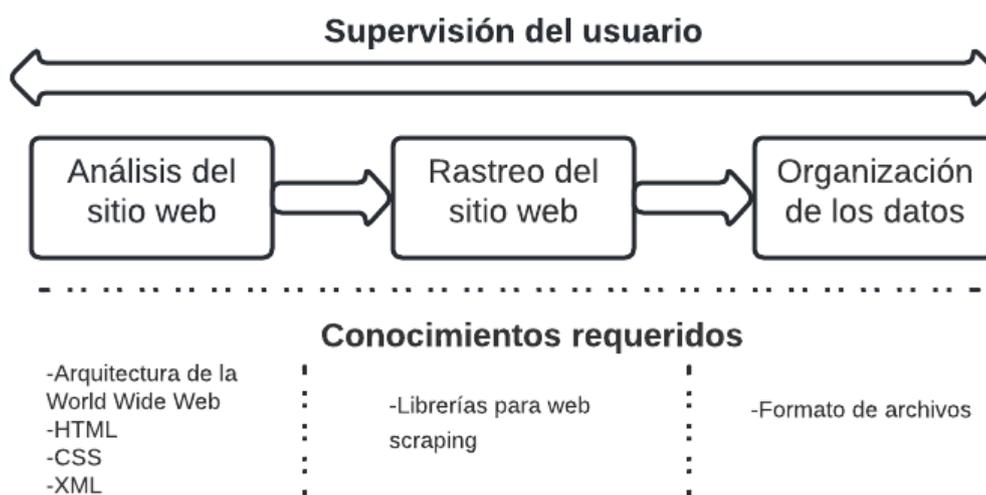
Broucke y Baesens [7] definen al *Web Scraping* como un agente construido para la descarga, análisis y organización de datos de forma automatizada. Por lo tanto, se puede definir al *Web Scraping* como una técnica para la extracción de información disponible en diferentes sitios web. Para Zhao [8], el *Web Scraping* extrae los datos disponibles en Internet haciendo uso del Protocolo de Transferencia de Hipertexto (HTTP) o a través de un navegador web. En definitiva, esta técnica es eficiente para recolectar los datos que se encuentran disponibles en varios sitios web.

El *Web Scraping* permite reunir y procesar grandes cantidades de información. Para Rabiyaoui et.al [9] el objetivo principal de un *web scraper*¹ es recoger los datos disponibles en sitios web para presentarlos en un formato estructurado. Además, esta técnica permite la transformación de los datos no estructurados que se encuentran disponibles en Internet en datos estructurados para que sean almacenados en bases de datos o en hojas de cálculo.

Para Krotov y Tennyson [10] el *Web Scraping* está formado de tres fases: análisis del sitio web, rastreo del sitio web y la organización de los datos extraídos como se muestra en la ilustración 1. La primera fase para el *Web Scraping* corresponde a realizar un análisis de la estructura del sitio web del que se realizará la extracción de los datos. La segunda etapa implica la programación de un script que sea capaz de navegar por el sitio web y realizar de manera automática la extracción de datos. Por último, los datos extraídos deben ser organizados, esta fase implica el preprocesamiento de datos para que permita realizar un análisis.

¹ Software que simula la navegación en el Internet para la recolección de datos.

Ilustración 1: Fases del Web Scraping



FUENTE: KROTOV Y TENNYSON [10]

ELABORADO POR: AUTORES

2.1.2. Legalidad y ética del Web Scraping

Actualmente, no existe una ley que regule las acciones relacionadas con el *Web Scraping*. Por el momento según Dreyer y Stockton [11], el *Web Scraping* se rige por un conjunto de teoría y leyes como la infracción de derechos, incumplimiento de contrato y la Ley de Fraude y Abuso Informático. A continuación, se enlistan algunos detalles de teorías legales que se aplican al Web Scraping.

- Términos de uso
- Material protegido por derechos de autor
- Propósito del *Web Scraping*
- Daños al sitio web

Mientras existen leyes y teorías jurídicas para el Web Scraping, la ética del mismo aún no ha sido abordada. Paul y Elder [12], definen la ética como “*un conjunto de conceptos y principios que nos guían para determinar qué comportamiento ayuda o perjudica a las criaturas sensibles*”. En consecuencia, se debe tener en cuenta que el *Web Scraping* puede causar daños a usuarios que se encuentren vinculados al sitio web.

2.1.3. Lenguajes de programación populares para Web Scraping

Los procesos de *Web Scraping* pueden implementarse utilizando diferentes lenguajes de programación. Los lenguajes de programación con mayor popularidad son Java, Python y PHP [7].

2.1.3.1. Web Scraping con Java

Java es uno de los lenguajes de programación principales utilizados para el *Web Scraping*, debido a que contiene una variedad de librerías para realizar el proceso de recolección de datos. Una de las desventajas de este lenguaje de programación es la sintaxis, debido a que en ocasiones hace que el código sea extenso y difícil de comprender [13].

2.1.3.2. Web Scraping con Python

Python posee una gran cantidad de librerías para el procesamiento de datos como Numpy, Scipy y otras similares, a su vez este lenguaje programación proporciona la facilidad de uso del lenguaje de scripting² [14].

2.1.3.3. Web Scraping con PHP

PHP proporciona diferentes librerías y herramientas que permiten llevar a cabo el proceso de recolección de datos. Además, proporciona APIs para realizar el *Web Scraping*. La limitante de este lenguaje de programación para realizar tareas de *Web Scraping* es la cantidad mínima de datos que se pueden extraer [15].

2.1.4. Librerías y APIs para Web Scraping

Las librerías están basadas en la estructura HTML³, esto facilita la recolección de datos, ofrecen los datos en formatos amigables para posteriormente sean procesados e integrados en distintos como: JSON⁴, XML⁵, CSV⁶, XLS o RSS⁷ [9]. Algunas de las librerías disponibles se describen en la tabla 1.

² Acción de escribir comandos para que sean interpretados por una aplicación.

³ Lenguaje de Marcado de Hipertexto, utilizado para la estructurar y despliegue de las páginas web.

⁴ JavaScript Object Notation, formato para el intercambio de datos.

⁵ Lenguaje de Marcado Extensible, permite almacenar datos de forma legible.

⁶ Formato para almacenar datos en forma de tabla.

⁷ Formato de estructura de datos que facilita el acceso a la información.

Tabla 1: Librerías y APIs para Web Scraping

Nombre de la librería	Descripción
Beautiful Soup	Es una librería de Python, permite la extracción de datos HTML y XML. Beautiful Soup facilita la navegación y extracción de la información disponible en los sitios web [16].
Jsoup	Es una librería de Java, proporciona una API para obtener la URL y extraer datos de sitios web [17].
Jaunt	Es una librería de Java que proporciona un navegador que permite realizar Web Scraping y controlar las peticiones de solicitud/respuesta HTTP. Una de las desventajas de Jaunt, es que no soporta JavaScript [18].
Scrapy	Es un framework que permite construir aplicaciones que realicen Web Scraping, además, Scrapy proporciona un mecanismo que empieza desde la extracción de datos hasta el almacenamiento de los resultados de la extracción [19].
Cheerio	Proporciona una API que permite manipular la estructura de los datos que se encuentran disponibles en los sitios web [20].
Newspaper	Es una librería que permite la extracción de artículos disponibles en los sitios web de noticias [21].
StormCrawler	Es una librería que permite a los usuarios desarrollar un <i>scraper</i> propio según los datos que requieran extraer [22].

FUENTE: KHALIL Y FAKIR [9]

ELABORADO POR: AUTORES

2.1.5. Análisis de Redes Sociales (*Social Networks Analysis, SNA*)

El SNA es un conjunto de herramientas, teorías y procesos que permiten comprender las relaciones entre actores. De acuerdo con Saqr y Alamro [23], el SNA es una agrupación de métodos y herramientas, las cuáles se utilizan para estudiar relaciones, interacciones y comunicaciones. El gráfico SNA representa las conexiones que existen entre los actores

(nodos) como puntos y las relaciones (aristas) como flechas que se originan desde la interacción hacia el objetivo.

Una de las principales características que se puede definir del SNA es su enfoque en la estructura de las relaciones; es decir, el SNA reconoce y mide las relaciones permitiendo conocer fácilmente los actores o unidades que interactúan entre sí [24].

2.1.6. Elementos básicos de una red

Una red está compuesta por tres elementos que la caracteriza, los cuales son: nodos o actores, relaciones y flujo. Los nodos o actores corresponden a las personas o grupos que se encuentran interviniendo, es necesario mencionar que dentro de una red pueden intervenir varios actores. Las relaciones, son las conexiones existentes entre dos o más nodos. Y, por último, el flujo indica la dirección del vínculo [23].

Para Afra et al. [25], mencionan que las aristas de una red evidencian el tipo de relación que existen entre los actores. Además, las relaciones se clasifican dependiendo su direccionalidad y la densidad de la relación.

2.1.7. Métricas en el SNA

La evaluación de la calidad de las relaciones entre los nodos y la comprensión de los patrones de conexión ha generado interés. A continuación, se describen las métricas que permiten realizar el análisis de la red [26].

2.1.7.1. Grado de intermediación (*betweenness*)

El grado de intermediación es la medida en que un nodo se encuentra entre el resto de los nodos de la red. Esta medida tiene en cuenta la forma de conexión de los vecinos del nodo, otorgando un valor alto a aquellos nodos que tienen puentes entre grupos [27].

2.1.7.2. Grado de cercanía (*closeness*)

El grado de cercanía mide la calidad de conexión de un nodo con relación a los demás nodos; en otras palabras, esta medida hace referencia a la cantidad de saltos para llegar a un nodo específico de la red [26].

2.1.7.3. Grado de centralidad (*centrality*)

El grado de centralidad es la cantidad de nodos con los cuales un nodo específico se conectada de manera directa. Esta métrica se divide en grado de entrada y grado de salida [27].

2.1.8. Software para SNA

Las herramientas de análisis de redes sociales se utilizan para representar los nodos y las aristas de una red, y para analizar los datos. En la tabla 2 se presentan herramientas disponibles para realizar el SNA.

Tabla 2 Software para SNA

Nombre de la herramienta	Descripción
Gephi	Es una plataforma que permite visualizar y explorar cualquier tipo de red, además, gráficos dinámicos y jerárquicos. Gephi, es una herramienta que permite explorar y entender los gráficos [28].
Pajek	Es un software utilizado para graficar redes, además, este software puede calcular medidas de centralidad, identificar nodos estructurales, etc. [29].
IGraph	Este software permite graficar y manipular las redes. Contiene implementaciones para problemas de grafos, flujo de la red, además, implementa algoritmos como la búsqueda de la estructura de la comunidad [30].
RStudio	Incluye herramienta que permite el análisis a nivel de los nodos de la red, métodos de distancia estructural y covarianza. Además, permite generar gráficos aleatorios y visualizarlos en 2D/3D [31].

FUENTE: AFRA et al. [25]

ELABORADO POR: AUTORES

2.1.9. Definición de habilidades (*skills*)

El término habilidades es polisémico, es decir, hay una diversidad de interpretaciones conceptuales y es utilizado con varios significados dependiendo la actividad formativa a la que se esté haciendo referencia. Como concepto, tiene múltiples fuentes teóricas; por lo tanto, en el presente trabajo se ha adoptado la definición de habilidades tomando como referencia el término *skills* del inglés que según *Cambridge Dictionary* [32], lo define como “*una habilidad para hacer bien una actividad o trabajo, especialmente porque lo has practicado*”.

A falta de una definición precisa para determinar el nivel de educación y preparación profesional, en el *Information Technology Curricula 2017* [33] definen el término *competence* como las cualidades que un empleado demuestra durante su rendimiento laboral. Mientras que el término *ability* según *Cambridge Dictionary* [34], lo define como “*el poder físico o mental o la habilidad necesaria para hacer algo*”. Por esta razón, el término y definición adoptado en el presente proyecto es *skills*, que como se ha mencionado es la habilidad que un empleado posee para utilizar conocimientos técnicos para cumplir y resolver tareas.

2.1.10. Clasificación de habilidades

2.1.10.1. Habilidades duras

Las habilidades duras según Sopa et al. [35], son aquellos conocimientos que pueden ser documentados de una forma sencilla, además, las habilidades duras pueden ser creadas, escritas y transferidas entre los empleados de una empresa. Es necesario mencionar que las habilidades duras pueden ser descritas de forma general y también ser delimitadas hacia un área del conocimiento específica. Algunos investigadores denominan actividades relacionadas con las habilidades duras a la gestión de proyectos, el dominio de alguna tecnología o herramienta.

2.1.10.2. Habilidades blandas

En el artículo “*Hard skills versus soft skills: Which are more important for indonesian employees innovation capability*” [35], definen a las habilidades blandas como características de la personalidad de una persona. Este tipo de habilidades son consideradas

difíciles de formular y compartir, debido a que forman parte de la personalidad natural de una persona. Las habilidades blandas se forman a partir de acciones, experiencias, ideales, valores, y otros aspectos de la vida de una persona. Por lo tanto, se pueden definir como el conocimiento personal o el conocimiento obtenido por las situaciones a las que se enfrentado cada individuo durante su vida.

2.1.11. E-recruitment

El *e-recruitment*⁸ es el método de reclutamiento de personal haciendo uso de los medios digitales. Este método de selección de personal engloba todos los tipos de reclutamiento que se encuentran disponibles en Internet, en la actualidad existen varios métodos de *e-recruitment* como: correo electrónico, formularios en páginas web, portales de empleo, entre otros. Cabe mencionar que este método permite la introducción de filtros que mejoran la selección de empleos [36].

2.1.12. Portales de empleo

Según un estudio realizado por Adecco e Infoempleo [37], para el año 2020 aproximadamente el 70% de la población de España buscaba ofertas laborales en sitios de alojamiento de empleo. Los portales de empleo en línea son plataformas de contratación disponibles en Internet en las cuales las empresas publican los anuncios de las vacantes de empleo, añadiendo información sobre el perfil que requieren para el empleo. Los portales de empleo en línea que recogen las ofertas laborales se han convertido en uno de los medios de contratación con mayor adecuación a la oferta y demanda de empleo [38].

2.1.13. Ejemplo de portales de empleo en línea

Actualmente, una gran cantidad de empresas anuncian sus vacantes de empleo por medio de anuncios en portales de empleo en línea. Los anuncios publicados se especifica el conjunto de habilidades que debe poseer el candidato [5]. En Internet hay varios portales de alojamiento de empleo, a continuación, se presentan algunos ejemplos.

⁸ Reclutamiento en línea

2.1.13.1. Indeed

Indeed es el sitio de trabajo número 1 en el mundo con más de 250 millones de usuarios únicos. Como empresa, Indeed se esfuerza por dar prioridad a las personas que buscan empleo, brindándoles acceso gratuito para buscar empleos, publicar currículos y empresas de investigación. Los buscadores de empleo vienen a Indeed.com para encontrar sus trabajos perfectos. Los empleadores visitan nuestro sitio web para publicar y patrocinar trabajos [1]. Indeed permite que los empleados proporcionen calificaciones generales de satisfacción, lo que da una visión general de su experiencia laboral con la empresa. Los usuarios también pueden calificar en 5 dimensiones diferentes [2]:

- i) Equilibrio entre trabajo y vida
- ii) Compensación/Beneficios
- iii) Seguridad Laboral/Progreso
- iv) Gestión
- v) Cultura, permitiendo así mayor profundidad.

2.1.13.2. LinkedIn

Se puede decir que LinkedIn es un entorno comunicativo en el que se combinan los contenidos y las audiencias relevantes, permitiendo crear un perfil de comportamiento, que posteriormente entra a formar parte de un público objetivo concreto. La propia evolución de LinkedIn como red profesional en este sentido ha llevado a la necesidad de generar un entorno que permita la captación y fidelización de usuarios, y donde la relevancia del contenido influye en el interés del público [3].

LinkedIn permite a los usuarios incluir información equivalente a su currículum, combinando información personal y profesional. Así, a partir de ella, los profesionales del reclutamiento y la selección son capaces de extraer información representativa sobre las habilidades específicas que muestran los candidatos [4].

2.1.14. Definición de currículo

Dentro del contexto universitario, el currículo encierra el desarrollo de diversos aspectos relacionados con los procesos educativos. El modelo curricular de las universidades

2.2. Marco referencial

El uso de técnicas de *Web Scraping* y SNA para determinar la demanda de habilidades en el mercado laboral se evidencia en los siguientes proyectos de investigación que contienen un contexto similar al presente.

Analizar las funciones y las competencias de los puestos de trabajo es esencial para determinar la brecha que existe entre las habilidades profesionales que el mercado laboral requiere y las que los estudiantes poseen al momento de ingresar al campo laboral. Es por eso que en el artículo titulado “*A Social Network Analysis of Jobs and Skills*” [41], los autores optaron por utilizar SNA para identificar los roles de empleo con competencias similares, además, identificar las relaciones que existen entre las competencias. El objetivo de los autores fue visualizar las relaciones de manera intuitiva y presentar un análisis que proporcione información a los solicitantes de empleo, empleadores y a las agencias de contratación para comprender las tendencias, funciones y competencias laborales necesarias para satisfacer la demanda de mano de obra.

Apatsidis, Georgiou y Mittas [2] presentan un estudio exhaustivo en *Stack Overflow* para examinar las principales características del trabajo remoto que lo discriminan de su contraparte presencial. Mediante la recopilación y el análisis de 8514 anuncios de empleo y el aprovechamiento de metodologías de minería de textos, se señalan los principales elementos que definen cada categoría, desde las tecnologías dominantes hasta los puestos de trabajo y los sectores que más buscan. Los resultados sugieren que el trabajo a distancia presenta diferencias con el trabajo tradicional, y que se asocia principalmente al sector de la ingeniería de software y a las tecnologías de desarrollo de software y análisis de datos más conocidas.

Por otra parte, en el artículo “*Adaptively selecting occupations to detect skill shortages from online job ads*” [1], los autores desarrollan un método basado en datos para generar conjuntos de competencias altamente similares a partir de un conjunto de competencias iniciales utilizando datos de anuncios de empleo en línea. Esto proporciona a los investigadores un método novedoso para seleccionar de forma adaptativa las ocupaciones basándose en datos de habilidades granulares.

En otro artículo, titulado “*Analysis of Skill Demand in Thai Labor Market from Online Jobs Recruitments Websites*” [34], los autores proponen un nuevo método para el análisis de las

competencias demandadas en el mercado laboral haciendo uso de la información disponible en los portales de empleo en línea. El artículo implementa varias técnicas como el *Web Scraping*, extracción de palabras claves y la visualización de la información extraída de los portales de empleo en línea. La presentación de los resultados se hace por medio de una nube de palabras con la descripción de los puestos de trabajo y la cualificación de cada empleo. Con los trabajos relacionados se muestra que la combinación de la técnica de Web Scraping y SNA proporciona información para conocer el desajuste de las competencias; es decir, el tipo de desajuste que existe entre la educación y la demanda de mano de obra.

2.3. Marco Legal

En la actualidad no existe una legislación que regule directamente el Web Scraping, los investigadores que utilizan esta técnica lo hacen bajo un panorama legal incierto, rigiéndose por un conjunto de leyes fundamentales. A continuación, se presentan algunas de las leyes sobre la protección de datos y garantía digital en países de América Latina, España y Estados Unidos.

El Reglamento General de Protección de datos (RGPD) [43], es una normativa de la Unión Europea, por lo tanto, cualquier empresa o aquellas empresas que manejen información personas de cualquier tipo debe acogerse a ella. La regulación se aplica si el controlador de datos (una organización que recolecta datos de residentes de la UE) o el procesador (una organización que trata datos en nombre del controlador de datos, por ejemplo, proveedores de servicios en la nube) o el interesado (persona) tiene su sede en la UE. Además, la regulación también se aplica a las organizaciones con sede fuera de la Unión Europea si recopilan o tratan datos personales de los residentes de la UE.

De la misma manera, la Ley Orgánica de Transparencia y Acceso a la Información Pública (LOTAIP), fue promulgada en el Suplemento del Registro Oficial No. 337 del 18 de mayo del 2004. En el artículo 4 de la LOTAIP [44], se establece que “*la información pública pertenece a los ciudadanos y ciudadanas y que su acceso es gratuito a excepción de los costos de reproducción, así como que el ejercicio de la función pública está sometido al principio de apertura y publicidad de sus actuaciones*”. Además, el artículo 7 de la LOTAIP [44] dispone que “*todas las instituciones, organismos y entidades, personas jurídicas de derecho público o privado que tengan participación del Estado o sean concesionarios de éste, difundirán a través de un portal de información, página web u otros medios a*

disposición del público implementados en la institución, la información que se describe en los literales del artículo enunciado”.

Por otra parte, la Ley Orgánica de Protección de Datos Personales y garantía de los derechos digitales [45] tiene como objetivo proteger la intimidad, privacidad e integridad del individuo de la información de los individuos, en cumplimiento con los artículos escritos en la Constitución Ecuatoriana. Del mismo modo, regula las obligaciones del individuo en todo proceso de transferencia de datos para garantizar la seguridad del intercambio.

Finalmente, en el 2020 en el Estado de California se aprobó la Ley de Privacidad del Consumidor [46], en la cual se responsabiliza a todas las empresas del manejo seguro de los datos personales. Además, esta ley exige que se informe a los usuarios sobre el uso de sus datos y si estos se ven comprometidos de alguna manera.

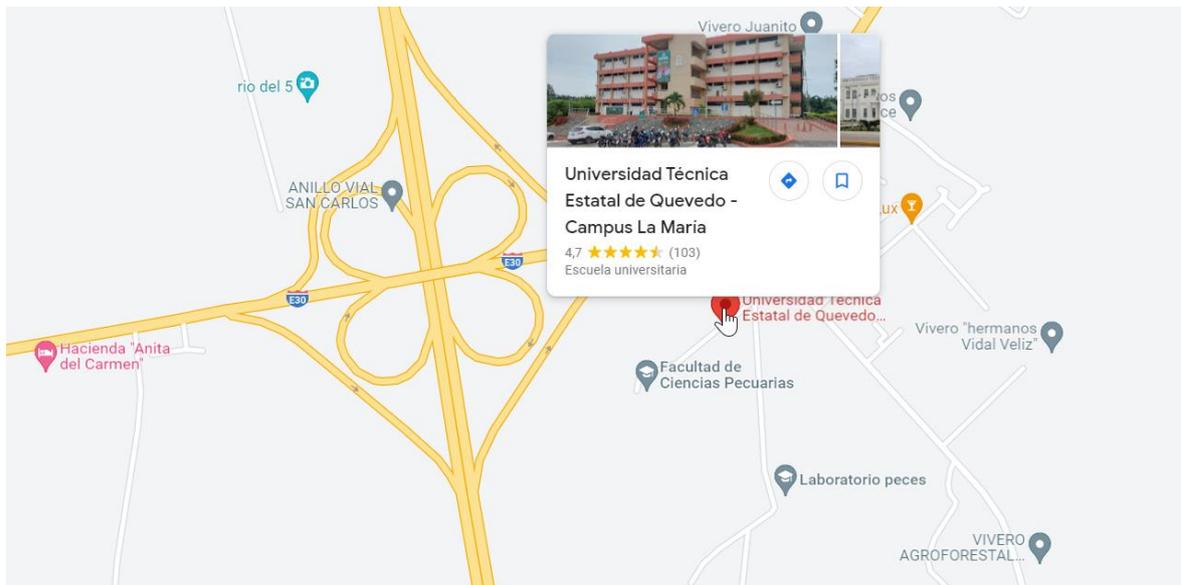
CAPÍTULO III

METODOLOGÍA DE LA INVESTIGACIÓN

3.1. Localización

El proyecto se realizó en la Universidad Técnica Estatal de Quevedo, campus “La María”, ubicado en el km 7 vía Quevedo-El Empalme (Ilustración 3). El proyecto tuvo una duración de 4 meses, desde el mes de junio del 2022 hasta el mes de septiembre del 2022.

Ilustración 3: Ubicación del Campus “La María”



FUENTE: GOOGLE MAPS

3.2. Tipo de investigación

Para el desarrollo del proyecto de investigación se utilizó dos tipos de investigación: documental y descriptiva. La investigación documental permitió recopilar información acerca de la combinación de las técnicas de *Web Scraping* y SNA para determinar la demanda de competencias en el mercado laboral. Se utilizaron artículos científicos de revistas y conferencias internacionales disponibles en bases de datos como IEEE, ACM Y SCOPUS. La investigación descriptiva, permitió caracterizar el proyecto de investigación y determinar el alcance del mismo.

3.3. Método de investigación

3.3.1. Método analítico

El método analítico fue utilizado para la selección del lenguaje de programación para desarrollar el *Web Scraping*, además, este método fue utilizado para conocer las librerías

disponibles para realizar el SNA y cuál de ellas se deben aplicar para obtener una red dinámica. El método analítico fue empleado para denominar las diferentes herramientas y tecnologías que permitieron ejecutar las fases para el desarrollo del *Web Scraping* y los requerimientos para realizar el modelo SNA para conocer las relaciones entre las habilidades de las distintas áreas del conocimiento.

3.3.2. Método deductivo

El método deductivo fue utilizado para analizar las diferentes fuentes de información disponibles que contienen similitud con el proyecto de investigación. Es por esto que este método es de gran utilidad en este proyecto, debido a que permite recopilar la información relacionada existente.

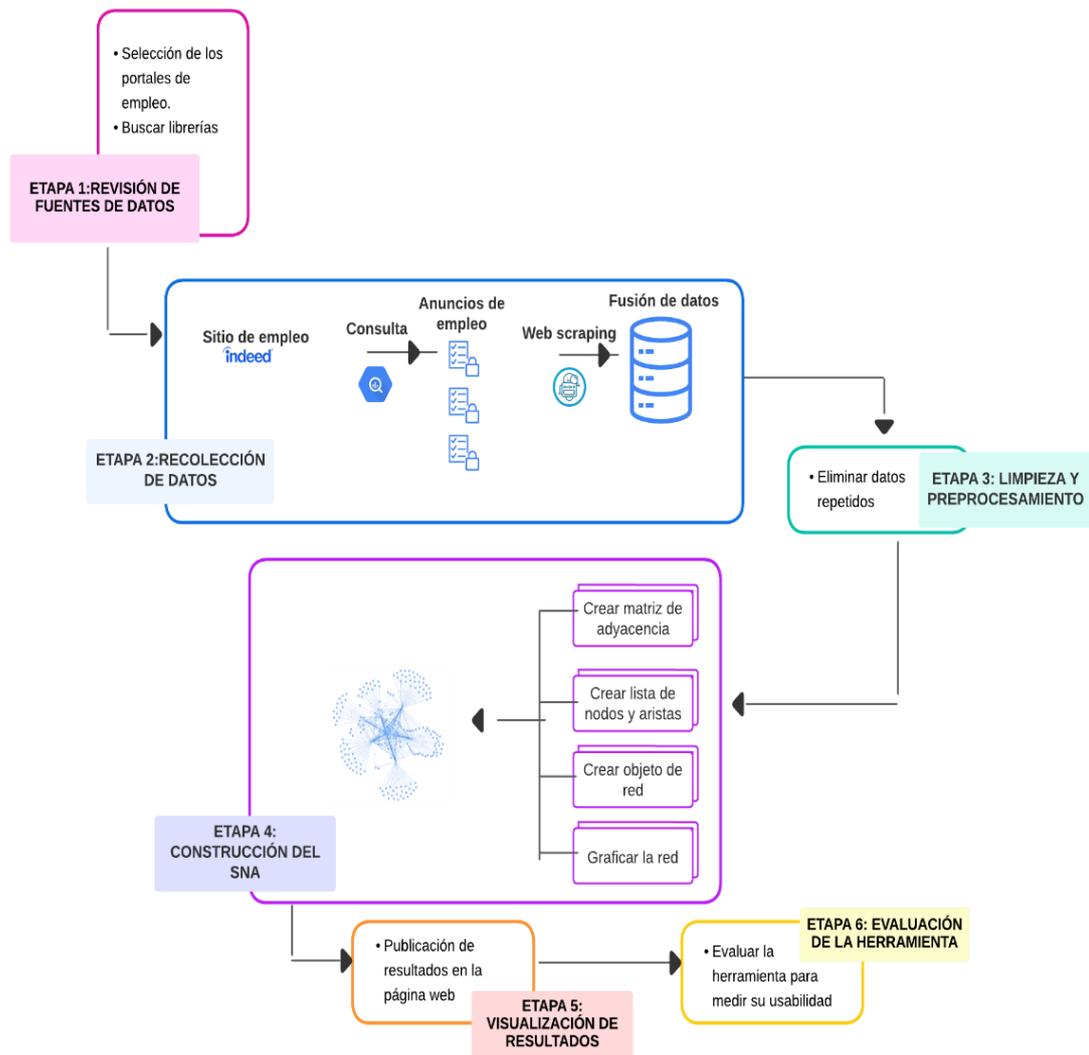
3.4. Fuentes de recopilación de información

Para la fundamentación teórica del proyecto de investigación, se utilizaron fuentes de información tales como: artículos científicos, artículos de revistas, libros y conferencias. El portal de empleo web Indeed Ecuador, fue utilizado como fuente para la extracción de los anuncios de empleo. Esta decisión se tomó debido a la popularidad del sitio, al ser considerado el sitio número 1 para publicar las ofertas de empleo [2], se encuentran una gran cantidad de datos que constituyen una fuente valiosa de información para el descubrimiento de conocimientos.

3.5. Diseño de la investigación

Los datos de las habilidades requeridas se extraerán del portal de empleo (Indeed) que es considerado el sitio con mayor popularidad [2]. Seleccionado el portal de empleo se procede a la recolección de datos usando la técnica de web scraping. La extracción de las habilidades requeridas -para cada puesto de trabajo- se encuentran en la descripción de los anuncios de empleo. Obtenidos estos datos, es necesario realizar la limpieza y preprocesamiento de estos para evitar obtener resultados erróneos. Posteriormente, se realiza el proceso de construcción de la red mediante SNA y se calcula los indicadores de grado de centralidad, grado de cercanía y grado de intermediación de la red. Finalmente, se visualiza la red.

Ilustración 4: Etapas de la metodología de desarrollo



ELABORADO POR: AUTORES

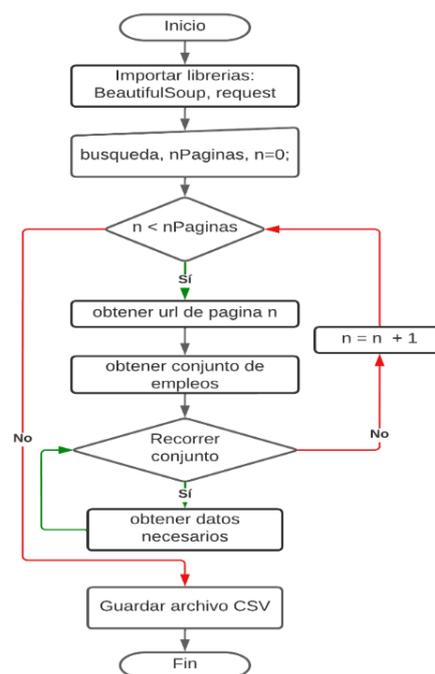
3.5.1. Primera Etapa: Revisión de fuentes de datos

En esta etapa se efectúa la revisión bibliográfica con el afán de recopilar información de trabajos relacionados que se encuentran disponibles en las bases de datos (ACM, IEEE XPLORE, SCOPUS). Se realiza un análisis crítico que conlleve conocer el enfoque y objetivo de los autores que han elaborado trabajos similares al que se propone en este artículo. Además, se revisa las consideraciones para la elección del portal de empleo, las librerías para ejecutar el Web Scraping y las librerías para construir el modelo de SNA. Asimismo, se identifican los criterios o palabras claves para la búsqueda de los anuncios de empleo.

3.5.2. Segunda Etapa: Recopilación de datos

Esta etapa se centra en extraer los anuncios de empleo desde el sitio web seleccionado para la obtención de los anuncios de empleo disponibles en Ecuador. Utilizando como términos de búsqueda los nombres de las carreras de la Facultad de Ciencias de la Ingeniería de la Universidad Técnica Estatal de Quevedo. Para analizar el contenido HTML del portal de empleo se utiliza la biblioteca de Python, BeautifulSoup. Esta librería facilita la extracción de los anuncios de empleo disponibles para cada término de búsqueda. En la ilustración 5 se presenta el algoritmo que se utilizó para la extracción de los anuncios.

Ilustración 5: Proceso para el Web Scraping



ELABORADO POR: AUTORES

De cada anuncio de empleo se van a extraer siete atributos: título del empleo, nombre de la empresa, ciudad, valoración sueldo, fecha de publicación y la descripción del anuncio. Estos atributos se encuentran en cada anuncio de empleo del sitio Indeed.com y se encuentran detallados en la tabla 3. Se debe tener en cuenta que algunos de los datos extraídos presentan un único elemento (por ejemplo, el título del empleo, el nombre la empresa, ciudad), mientras que otros tienen un formato de frases descriptivas (por ejemplo, la descripción del anuncio de empleo). Para extraer los atributos del HTML buscamos las etiquetas que contienen esta información en este caso la etiqueta div o caja. Luego obtenemos de manera cíclica la información de cada etiqueta.

Tabla 3: Descripción de variables

Variable	Descripción
Título de empleo	El nombre del puesto al que corresponde el anuncio de empleo.
Empresa	La empresa en la cual se encuentra la vacante del anuncio publicado.
Ciudad	La ubicación del anuncio de empleo publicado.
Valoración	Calificación del anuncio de empleo.
Sueldo	Sueldo del anuncio de empleo.
Fecha de la publicación	Fecha en la que se creó el anuncio de empleo en Indeed.
Descripción del anuncio	Lista de habilidades necesarias para postular al anuncio de empleo.

**FUENTE: INVESTIGACIÓN
ELABORADO POR: AUTORES**

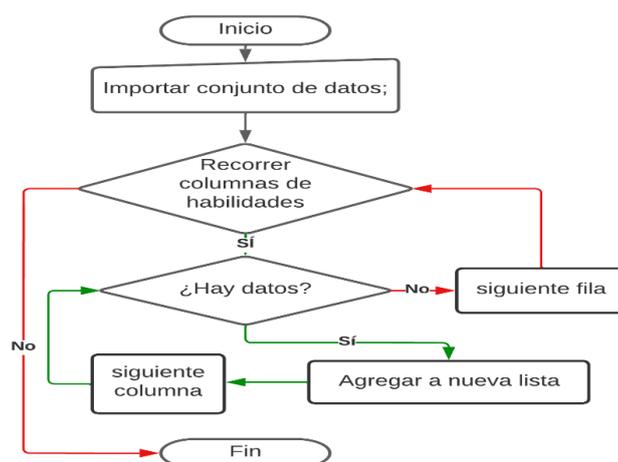
3.5.3. Tercera Etapa: Limpieza y preprocesamiento

Se centra en la revisión del contenido de las variables dentro del conjunto de datos. Como la descripción de los anuncios de empleo contiene un texto en el que explica diferentes aspectos del empleo, por tanto, no es posible utilizar directamente los datos. Se hace necesario extraer las habilidades que se requieren para un empleo en cada anuncio. La limpieza de los datos consistió en aplicar los siguientes criterios:

- Eliminar caracteres especiales
- Anuncios de empleos duplicados
- Anuncios de empleo que no tenían relación con el término de búsqueda.

Para obtener las habilidades de la descripción del anuncio, se implementó el algoritmo de la ilustración 6. El algoritmo comienza importando el conjunto de datos al que se realizó la limpieza para luego recorrer cada columna de habilidad verificando si existe un dato. Si se cumple la condición se agrega a un conjunto auxiliar de habilidades, caso contrario se pasa a la siguiente fila u oferta laboral.

Ilustración 6: Etapa de clasificación de habilidades profesionales



ELABORADO POR: AUTORES

3.5.4. Cuarta Etapa: Construcción del modelo SNA

Para encontrar las relaciones se tomó como términos de búsqueda el nombre de las carreras de la Facultad de Ciencias de la Ingeniería y las habilidades que solicitaban en los anuncios de empleo. Se usará la librería tidyverse para la creación de la lista de nodos y arista, mientras que para crear los gráficos se usó visNetwork y networkD3 que son librerías disponibles en R para la creación y manipulación de redes. Para el SNA, se utilizaron las columnas en donde se encontraba el nombre del área del conocimiento, ciudad, provincia y las habilidades. Se toman las variables del área de conocimiento y habilidades de la carrera para conocer la relación que existe entre cada habilidad y las distintas carreras de la FCI y si estas habilidades son solicitadas en otra carrera. Asimismo, se escogen las variables ciudad y provincia para identificar la disponibilidad de ofertas laborales por carrera.

Para construir una red se debe considerar dos aspectos: los nodos y las conexiones. Para conocer sus relaciones, es necesario construir primero una matriz de adyacencia, que es una matriz cuadrada cuyos nombres de columna y fila son los nodos de la red. Dentro de la matriz de adyacencia, un uno indica que hay una conexión entre los nodos y un cero indica que no hay conexión. La matriz sirve para conocer que habilidades se relacionan con las diferentes carreras de las FCI.

Las librerías utilizan datos numéricos para realizar las relaciones entre sí, por lo tanto, para los nodos se debe crear un campo identificador para reconocer las columnas con datos alfanuméricos. Por otra parte, para las aristas se toma el mismo conjunto de datos

adicionando una columna con el peso de cada arista, cabe mencionar que el peso es el equivalente al número de veces que se repite una habilidad en un área de conocimiento determinada.

Teniendo listos los datos se harán uso de las funciones de los paquetes mencionados para crear gráficos interactivos de las habilidades y sus relaciones con las áreas del conocimiento, ciudades y/o provincias. Además, nos permite agregar parámetros adicionales como el tamaño de nodos, filtros en la red, controles de navegación, modo de presentación. Con la finalidad de brindar una mejor visualización y que la red sea interactiva.

3.5.5. Quinta Etapa: Análisis y visualización de datos

Los diferentes resultados obtenidos tales como la extracción de los datos y el modelo SNA se presentaron en una aplicación web. La aplicación servirá como medio de información para permitir a las universidades y estudiantes contar con información sobre las habilidades laborales que son requeridas en el mercado laboral en un área de conocimiento en específico. El aplicativo muestra información sobre el total de habilidades, las ofertas laborales que se extrajeron, información estadística representada mediante tablas, nube de palabras, gráfico de barras y la red SNA. Además, se adicionó la opción para generar un reporte con las habilidades laborales clasificadas por área de conocimiento.

3.5.6. Sexta Etapa: Evaluación de usabilidad de la herramienta

En la última etapa, se evaluó la herramienta web para conocer el nivel de usabilidad. Para esto se hizo uso del cuestionario SUS (*System Usability Scale*). En esta etapa se realizaron una serie de acciones para determinar la facilidad de uso de la herramienta desarrollada. Para ello se siguieron las siguientes pautas para lograr un mayor aprovechamiento de la herramienta propuesta.

- La herramienta debe ser intuitiva y fácil de utilizar.
- La presentación de la información debe ser rápida.
- La información que presente la herramienta debe estar ordenada.
- El contenido debe ser claro y preciso.

- Las imágenes deben tener colores llamativos y que no provoquen confusión.

3.6. Instrumentos de investigación

Los instrumentos utilizados para el desarrollo del proyecto de investigación fueron los siguientes:

- Reuniones con el director del proyecto de investigación.
- Análisis de los sitios de empleo.
- Análisis de librerías para *Web Scraping* y SNA.
- Cuestionario SUS.

3.7. Tratamiento de datos

Para el tratamiento de los anuncios de empleos obtenidos con el *Web Scraping* al sitio de empleo Indeed Ecuador fueron depurados haciendo uso de los softwares de R y Excel.

Para el caso de la aplicación de SNA en los datos Carrera-Habilidades las habilidades se encontraban separadas por columnas según la oferta laboral correspondiente. Para poder realizar el conteo de habilidades por carrera y anuncio de empleo se unificó cada habilidad a filas para conocer el número total en base al conjunto de datos ingresado.

Por otra parte, en la relación Carrera-Ciudad y Carrera-Provincia sólo se tomó el número total de anuncios de empleo para conocer la cantidad de ciudades y provincias se encontraban en el conjunto de datos ingresado.

3.8. Recursos humanos y materiales

Para el desarrollo de cada una de las etapas del proyecto de investigación fueron necesarias hacer uso de los siguientes recursos humanos y materiales:

3.8.1. Recursos humanos

El proyecto de investigación ha sido realizado por el Sr. Ariel Oswaldo Fernández Loor y la Srta. Dayanna Lizzet Velez Araujo, bajo la dirección del Ing. Ariosto Eugenio Vicuña Pino.

Las personas mencionadas han participado activamente en cada una de las etapas de desarrollo del proyecto.

3.8.2. Equipos y materiales

3.8.2.1. Hardware

Para el desarrollo del proyecto se utilizó el hardware que se presenta en la tabla 4.

Tabla 4: Hardware utilizado para el desarrollo del proyecto

Cantidad	Equipo	Características
1	Laptop HP PAVILION	Procesador: Intel® Core™ i5-1035G1 CPU @ 1.00GHz 1.19 GHz RAM: 8,00 GB Tipo de sistema: Sistema operativo de 64 bits
1	Laptop DELL INSPIRON	Procesador: Intel® Core™ i7-8565U CPU @ 1.80GHz (8CPUs), ~2.0GHz 1.19 GHz RAM: 16,00 GB Tipo de sistema: Sistema operativo de 64 bits

ELABORADO POR: AUTORES

3.8.2.2. Software

En la tabla 5 se detalla el software utilizado para el desarrollo del proyecto de investigación.

Tabla 5: Software utilizado para el desarrollo del proyecto

Software	Función
Visual Studio Code	Desarrollo del algoritmo para realizar Web Scraping
RStudio	Desarrollo de algoritmo de limpieza de datos y SNA en base a los datos de entrada en el archivo .csv
Apache NetBeans IDE 13	Desarrollo de herramienta Web para la visualización de las redes generadas por RStudio.
Social Network Visualizer	Cálculo de los indicadores de centralidad, cercanía e intermediación de las redes.
Microsoft Office	Documentación

ELABORADO POR: AUTORES

CAPÍTULO IV
RESULTADOS Y DISCUSIÓN

4.1. Resultados de la extracción de los anuncios de empleo mediante la técnica de *Web Scraping* y limpieza de datos

La cantidad total de anuncios de empleo extraídos fue de 512, en los que se identificaron un total de 755 habilidades profesionales. Una gran cantidad de los anuncios de empleo extraídos estuvieron relacionados con las carreras de Ingeniería en Sistemas e Ingeniería en Software. Esto se debe a que las empresas dedicadas al desarrollo de software utilizan estos medios para publicar sus ofertas de empleo. Por lo tanto, existe mayor demanda laboral para los egresados dedicados a esta área de conocimiento. La distribución de la cantidad de anuncios de empleos y habilidades extraídos para cada carrera de la FCI se muestra en la tabla 6.

Tabla 6: Cantidad de anuncios de empleo y habilidades por área de conocimiento

Carreras de la Facultad de Ciencias de la Ingeniería	N.º de anuncios de empleo	N.º de habilidades
Ingeniería en Software	110	112
Ingeniería en Sistemas	201	214
Ingeniería Mecánica	99	180
Ingeniería Ambiental	14	57
Ingeniería Eléctrica	53	93
Arquitectura	27	49
Ingeniería Telemática	8	50

FUENTE: INVESTIGACIÓN

ELABORADO POR: AUTORES

4.2. Resultado de las redes de habilidades profesionales aplicando SNA

En la tabla 7 se detalla el resumen de la cantidad de atributos por elemento que construye la red general de habilidades profesionales requeridas en las distintas carreras de la FCI.

Tabla 7: Cantidad de atributos por elemento de la Red SNA “habilidades”

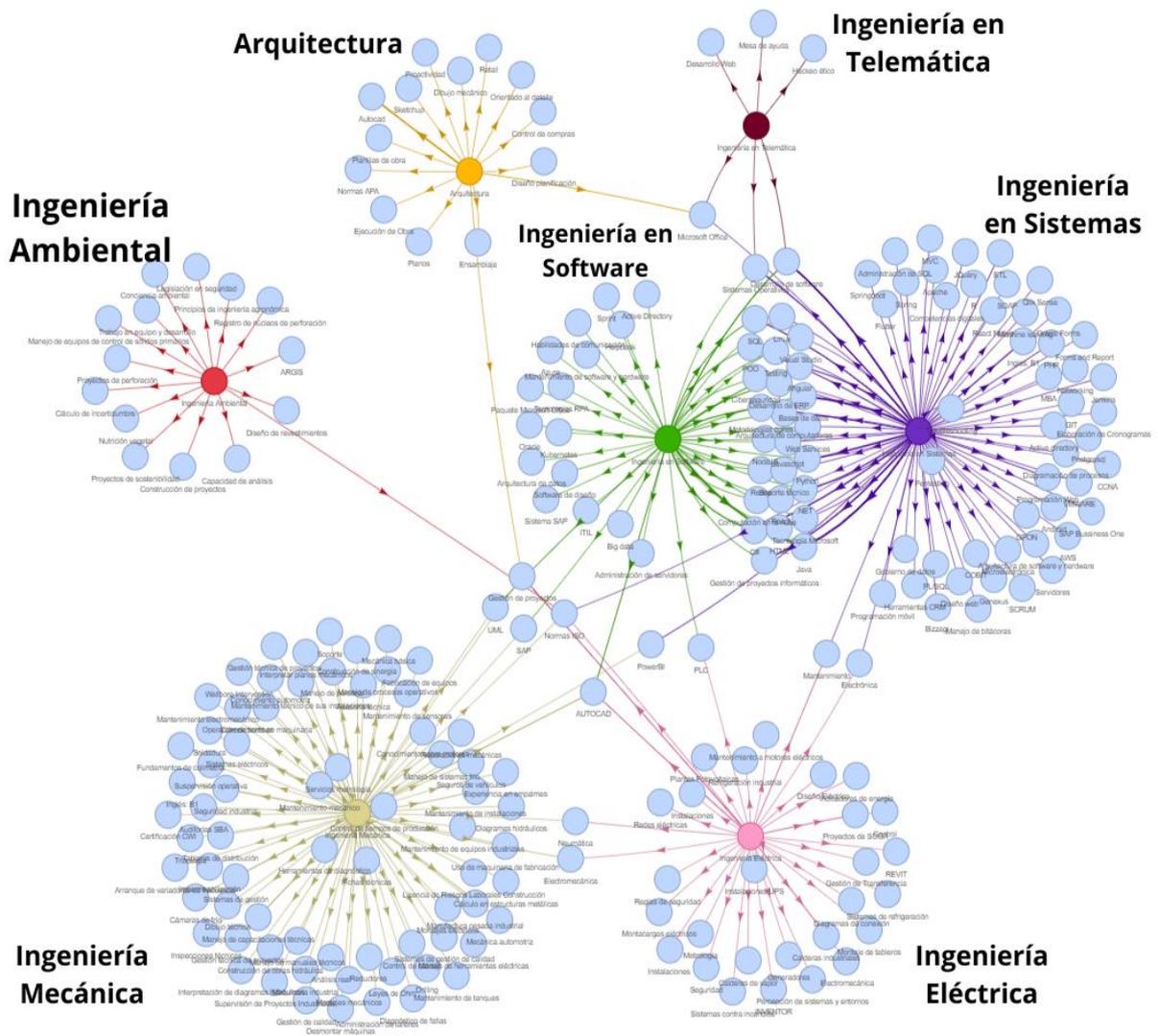
Tipo	Cantidad
Nodos	605
Aristas	768

FUENTE: INVESTIGACIÓN

ELABORADO POR: AUTORES

La relación de las habilidades profesionales con las carreras de la FCI se muestra en la ilustración 7. Del análisis se obtuvo que la mayoría de las habilidades se encontraban relacionadas con el desarrollo de software. La tabla 8 por su parte muestra la información de los nodos y el valor de los indicadores de la red, con lo que se determinó en nodo con mayor influencia en la red. De acuerdo con los resultados de las habilidades profesionales, se graficó una nube de palabra (ilustración 8) con las principales habilidades profesionales.

Ilustración 7: Red de habilidades profesionales requeridas en distintas carreras de la FCI



**FUENTE: INVESTIGACIÓN
ELABORADO POR: AUTORES**

Tabla 9: Indicadores para la red de Arquitectura

Nº	Nodos	Grado de centralidad	Grado de cercanía
8	AutoCAD	16	0.001593
10	Microsoft Office	7	0.001549
9	Revit	4	0.001546
18	Sketchup	3	0.001463
63	Lumion	3	0.001246

Densidad de la red: 0,04= 4%

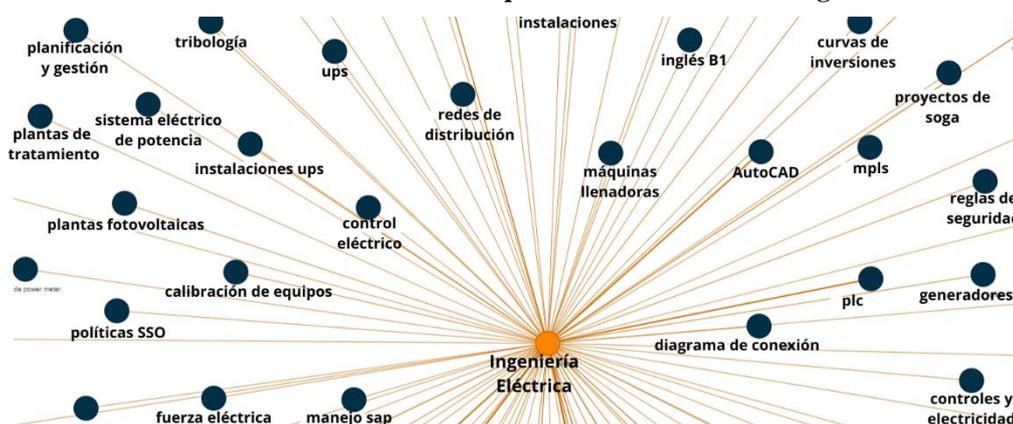


Ilustración 10: Nube de palabras de habilidad demandadas en la carrera de Arquitectura

**FUENTE: INVESTIGACIÓN
ELABORADO POR: AUTORES**

Entre las habilidades con mayor demanda para la carrera de Ingeniería Eléctrica según los indicadores están: Mantenimiento, Diseño eléctrico, AutoCAD, Gestión de proyectos y PLC. La tabla 10 detalla los valores que tienen los nodos obtenidos de la red de Ingeniería Eléctrica (ilustración 11) en cada uno de los indicadores para las cinco principales habilidades para la carrera de Ingeniería Eléctrica según la información de los anuncios de empleo. El análisis reveló que las habilidades profesionales se encuentran relacionadas con el mantenimiento y el diseño de sistemas eléctricos. Se encontró que la habilidad “Mantenimiento” presentó catorce enlaces.

Ilustración 11: Red SNA de habilidades requeridas en la carrera de Ingeniería Eléctrica



**FUENTE: INVESTIGACIÓN
ELABORADO POR: AUTORES**

La nube de palabras para la carrera de Ingeniería Eléctrica (ilustración 12), muestra las principales habilidades que se requieren en esa carrera según la información de los anuncios de empleo.

Tabla 10: Indicadores para la red de Ingeniería Eléctrica

Nº	Nodos	Grado de centralidad	Grado de cercanía
6	Mantenimiento	14	0.012149
3	Diseño eléctrico	12	0.011990
2	Autocad	10	0.011773
4	Gestión de proyectos	10	0.011773
25	PLC	10	0.011773

Densidad de la red: $0,0212766 = 2,12766\%$

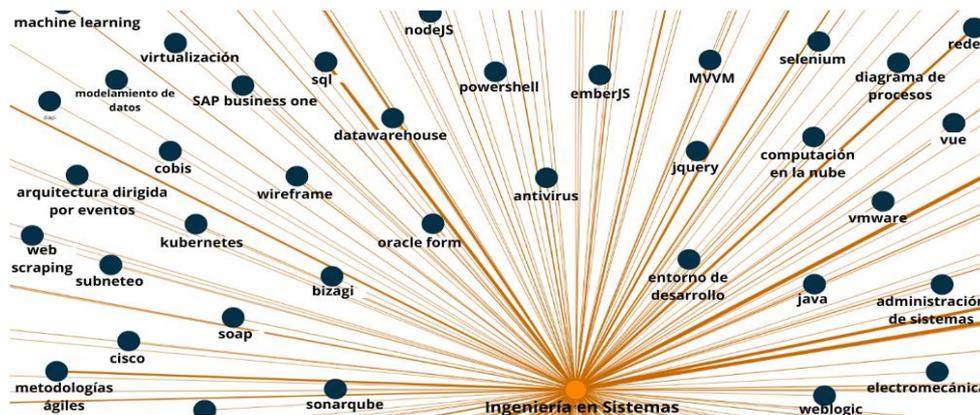


Ilustración 12: Nube de palabras de habilidad demandadas en la carrera de Ingeniería Eléctrica

**FUENTE: INVESTIGACIÓN
ELABORADO POR: AUTORES**

Esta red (ilustración 13) muestra que la habilidad profesional con mayor demanda para la carrera de Ingeniería en Sistemas es “Bases de Datos” presentando en su indicador de grado de cercanía un valor de 48. La información visualizada en la red es analizada a partir de los valores que tiene los indicadores de los nodos con mayor peso, en la tabla 11 se indica las cinco principales habilidades para la carrera de Ingeniería en Sistemas. El análisis de los indicadores mostró que las habilidades con mayor demanda se relacionan directamente con el desarrollo y mantenimiento de sistemas computacionales.

Ilustración 13: Red SNA de las habilidades requeridas en la carrera de Ingeniería en Sistemas



**FUENTE: INVESTIGACIÓN
ELABORADO POR: AUTORES**

En la ilustración 14, se muestra la nube de palabras correspondiente a las habilidades con mayor demanda para la carrera de Ingeniería en Sistemas. Como se puede observar, la mayoría de los datos requieren habilidades relacionadas con bases de datos.

Para el caso de la carrera de Ingeniería Ambiental la nube de palabras (ilustración 18) muestra que los empleos requieren que los empleados tengan conocimientos con herramientas de mapeo e información geográfica como ArcGIS y mapInfo.

Tabla 13: Indicadores para la red de Ingeniería Ambiental

Nº	Nodos	Grado de centralidad	Grado de cercanía
7	ArcGIS	2	0.012048
9	MapInfo	2	0.012048
13	Normas ISO	2	0.012048
20	Proyectos de sostenibilidad	1	0.012048
2	Agricultura sostenible	1	0.009009

Densidad de la red: $0,0344828 = 3,44828\%$

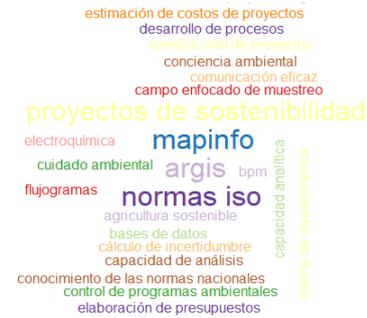


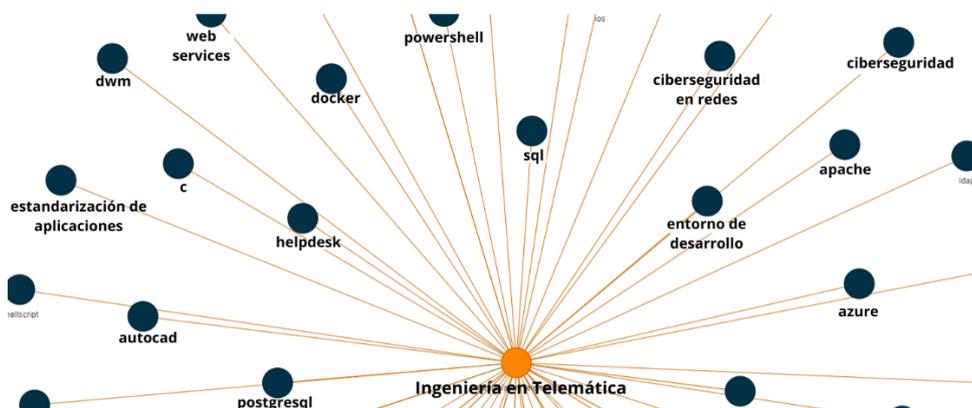
Ilustración 18: Nube de palabras de habilidad demandadas en la carrea de Ingeniería Ambiental

FUENTE: INVESTIGACIÓN

ELABORADO POR: AUTORES

De la misma manera, en la ilustración 19 se observa la red SNA correspondiente a la carrera de Ingeniería en Telemática. La tabla 14 por su parte muestra los valores de centralidad y cercanías para los nodos relacionados con la carrera de Ingeniería en Telemática. Los nodos con mayor influencia en esta red según el valor en el indicador grado de cercanía son: Bases de datos, Java, Normas ISO, Python y Sistemas Operativos, lo que significa que las habilidades se relacionan principalmente con el desarrollo y mantenimiento de sistemas informáticos.

Ilustración 19: Red SNA de habilidades requeridas en la carrera de Ingeniería en Telemática



FUENTE: INVESTIGACIÓN

ELABORADO POR: AUTORES

La nube de palabras para la carrera de Ingeniería en Telemática se muestra en la ilustración 20, el propósito del gráfico es representar las habilidades con mayor influencia en la red.

Tabla 14: Indicadores para la red de Ingeniería en Telemática

Nº	Nodos	Grado de centralidad	Grado de cercanía
2	Bases de datos	3	0.004717
4	Java	3	0.004717
6	Normas ISO	3	0.004717
7	Python	3	0.004717
12	Sistemas operativos	3	0.004717



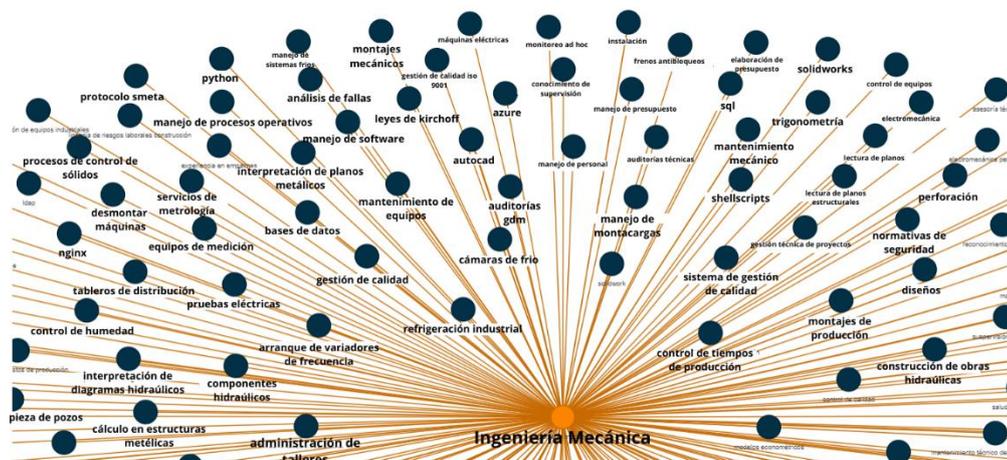
Densidad de la red: $0,0392157 = 3,92157$

Ilustración 20: Nube de palabras de habilidad demandadas en la carrea de Ingeniería en Telemática

**FUENTE: INVESTIGACIÓN
ELABORADO POR: AUTORES**

En la ilustración 21 se puede apreciar la red obtenida para las habilidades profesionales requeridas para la carrera de Ingeniería Mecánica. El indicador de grado de centralidad muestra que el nodo AutoCAD tiene un valor de 9, lo que significa que un gran número de anuncios solicitaban como requerimiento esta habilidad. En la tabla 15, se presentan los valores obtenidos en cada uno de los indicadores. Además, se pudo notar que el nodo seguridad industrial obtuvo un valor de 0.004645 en el indicador de grado de cercanía. La densidad de la red es 0,97561%, este valor depende de la dimensión de la red. Con lo que se puede concluir que la mayoría de los anuncios de empleo requieren habilidades que se encuentran relacionados con el diseño e instalación de equipos mecánicos.

Ilustración 21: Red SNA de habilidades requeridas en la carrera de Ingeniería Mecánica



**FUENTE: INVESTIGACIÓN
ELABORADO POR: AUTORES**

Para el caso de la carrera de Ingeniería Mecánica la nube de palabras (ilustración 22) muestra que la habilidad relacionada con AutoCAD tiene mayor demanda.

Tabla 15: Indicadores para la red de Ingeniería Mecánica

Nº	Nodos	Grado de centralidad	Grado de cercanía
48	AutoCAD	9	0.004788
27	Seguridad industrial Mantenimiento	7	0.004645
5	de instalaciones	6	0.004543
6	Soldadura	5	0.004407
25	Conocimiento automotriz	4	0.004219

Densidad de la red: $0,0097561 = 0,97561\%$



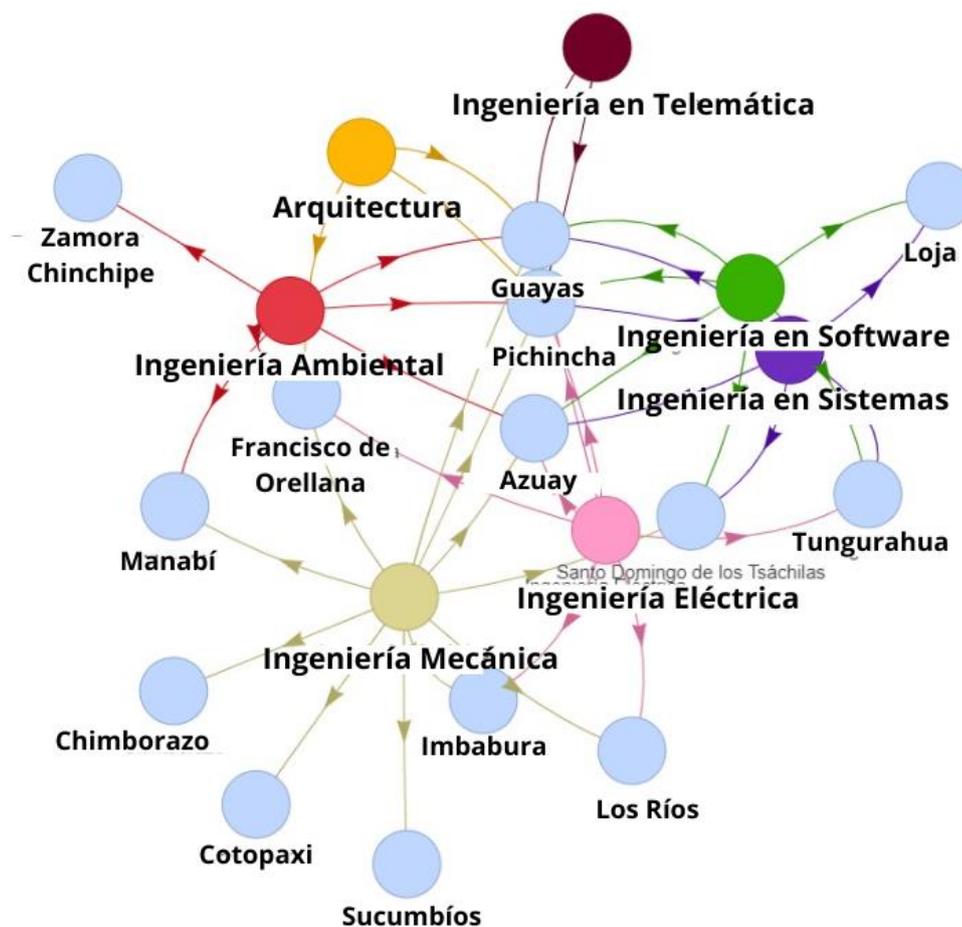
Ilustración 22: Nube de palabras de habilidad demandadas en la carrera de Ingeniería Mecánica

**FUENTE: INVESTIGACIÓN
ELABORADO POR: AUTORES**

La red de provincias (ilustración 23) evidencia la disponibilidad de ofertas laborales en distintas provincias del Ecuador para las carreras de la FCI. En general, se puede notar que la mayor cantidad de ofertas laborales se encuentra en la provincia de Pichincha. La tabla 16 por su parte muestra los valores de enlaces, cercanía y grado de intermediación de los nodos que intervienen en la red. Se observa el identificador del nodo, nombre y los valores obtenidos para cada uno de los indicadores. Estos últimos indican la importancia de los nodos en la red de provincias.

Además, el nodo con mayor influencia para el indicador de grado de centralidad es Pichincha con 223 enlaces, lo que significa que la mayoría de las ofertas de empleo se encuentra en esta provincia. Esto no indica precisamente que solo en esta provincia existe disponibilidad de ofertas laborales, debido a que la columna de grado de centralidad indica que la provincia del Guayas tiene un valor alto (0.122035); es decir, cuenta con una gran relación con el resto de los nodos de la red. Finalmente, la densidad de la red es alta (19,5238%) en comparación con las densidades de las redes de las habilidades, esto se debe a que no existen un gran número de conexiones dentro de la red.

Ilustración 23: Red SNA de las provincias con mayor disponibilidad de ofertas



FUENTE: INVESTIGACIÓN
ELABORADO POR: AUTORES

Tabla 16: Indicadores para la red de provincias

Nº	Nodos	Grado de centralidad	Grado de cercanía	Grado de intermediación
4	Ingeniería en Sistemas	201	0.121567	64.5
5	Ingeniería en Software	110	0.119493	1.5
8	Pichincha	223	0.122231	130
9	Guayas	139	0.122035	23.5

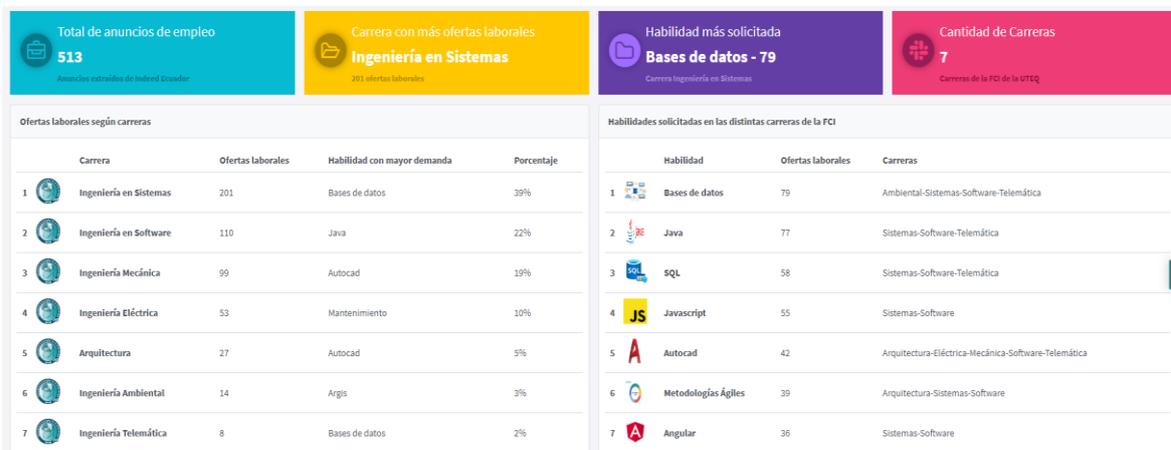
Densidad de la red: 0.195238=19,5238%

FUENTE: INVESTIGACIÓN
ELABORADO POR: AUTORES

4.3. Resultados de la herramienta web informativa y del reporte de habilidades profesionales requeridas en cada área del conocimiento según los anuncios de empleo

Para mostrar los resultados del SNA, se desarrolló una herramienta web en la cual se muestran las distintas redes obtenidas, la cual se puede observar en el anexo 7. Con la información obtenida de los anuncios de empleos extraídos se desarrolló un *dashboard* y el resultado se muestra en la ilustración 24.

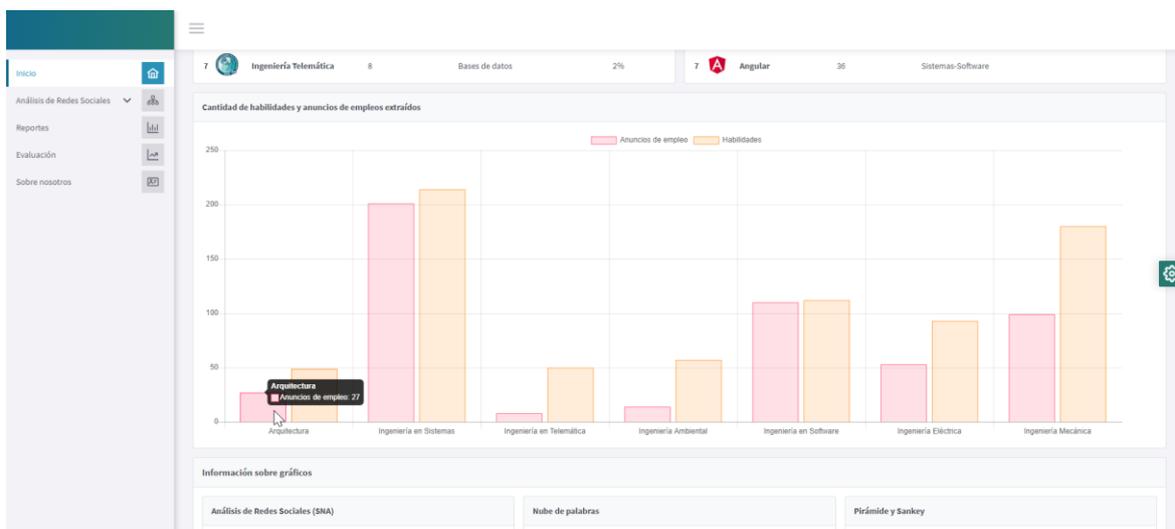
Ilustración 24: Dashboard de la herramienta web



FUENTE: INVESTIGACIÓN
ELABORADO POR: AUTORES

En el *dashborard* también se muestra un gráfico con información sobre el total de anuncios y habilidades extraídas como se presenta en la ilustración 25.

Ilustración 25: Gráfico de habilidades y anuncios extraídos



FUENTE: INVESTIGACIÓN
ELABORADO POR: AUTORES

Para que los usuarios visualicen las redes que se obtuvieron, se optó por clasificar las redes por carreras y por ubicación. En la sección de carreras se muestra las distintas redes de las habilidades necesarias para las áreas de conocimientos analizadas. Por otra parte, en la sección de ubicaciones se muestra la disponibilidad de ofertas laborales en cada ciudad y provincia.

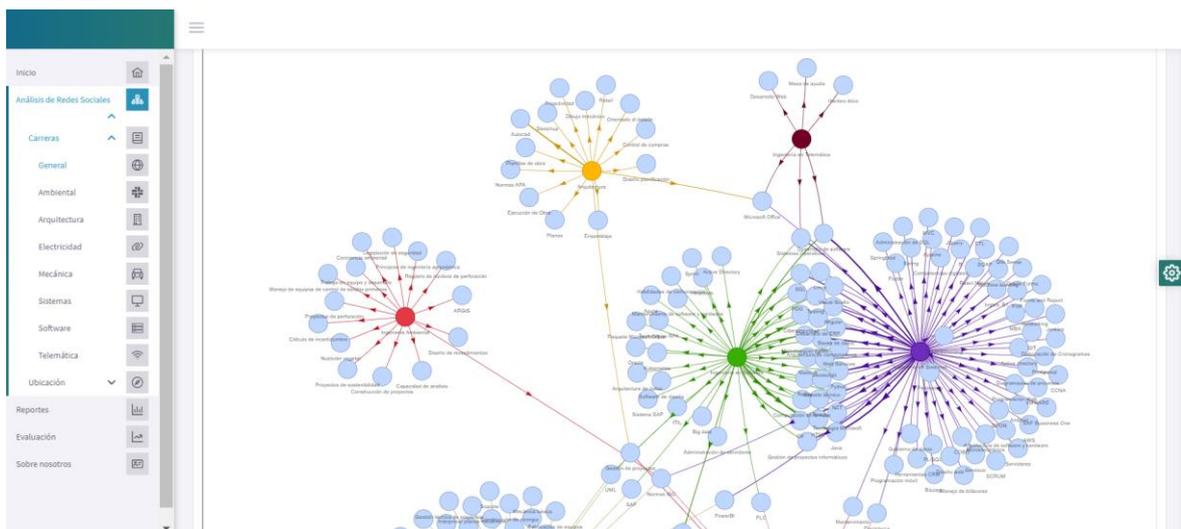
En la ilustración 26 y 27, se puede observar la información relacionada con el SNA general. Además, de presentar la red general de habilidades, se proporciona una tabla con el top de las habilidades con mayor demanda y su respectiva nube de palabras.

Ilustración 26: SNA Habilidades requeridas



**FUENTE: INVESTIGACIÓN
ELABORADO POR: AUTORES**

Ilustración 27: Interfaz para la visualización de la red de habilidades



**FUENTE: INVESTIGACIÓN
ELABORADO POR: AUTORES**

La interfaz de la aplicación web para generar el reporte se muestra en la ilustración 28. El reporte puede ser generado con información sobre las habilidades demandadas para una carrera o puede elegir la información de varias carreras. Como existen carreras que tienen un gran número de habilidades, se proporciona al usuario que elija la cantidad de habilidades que se desea visualizar según el área de conocimiento profesional que seleccione. El reporte se presenta en el anexo 3.

Ilustración 28: Interfaz de usuario



**FUENTE: INVESTIGACIÓN
ELABORADO POR: AUTORES**

4.4. Resultados de la evaluación de usabilidad de la herramienta web

Para conocer la opinión de los usuarios sobre la usabilidad de la herramienta web desarrollada se aplicó el cuestionario SUS, el cual proporcionó una puntuación de referencia sobre la opinión de los usuarios. En la tabla 17 se presenta cada afirmación, media, desviación estándar de las puntuaciones y una puntuación absoluta para ajustar las afirmaciones 2,4,6,8 y 10 para que las respuestas positivas se asocien a un número mayor.

Tabla 17: Respuestas a las afirmaciones individuales de la escala SUS

Pregunta	M	SD	M Ajustado
Me gustaría usar la aplicación frecuentemente	4,36	1,15	4,36
Encontré la aplicación innecesariamente compleja	2,08	0,91	3,92
Pienso que la aplicación fue fácil de usar.	4,28	0,84	4,28
Necesitaría el apoyo de un técnico/profesor para utilizar la aplicación.	1,96	0,89	4,04
Encontré que las diferentes funciones de la aplicación fueron bien integradas (constituyen un todo).	4,28	0,89	4,28
Pienso que había demasiadas inconsistencias en la aplicación.	2,00	1,19	4,00
Imagino que la mayoría de las personas aprenderían a usar rápidamente la aplicación.	3,68	1,49	3,68
Encontré la aplicación muy difícil de usar.	2,08	1,26	3,92
Me sentí muy seguro/cómodo usando la aplicación.	4,04	1,06	4,04
Necesito aprender muchas cosas antes de utilizar la aplicación.	2,44	1,53	3,56

FUENTE: INVESTIGACIÓN

ELABORADO POR: AUTORES

Se encontró que las afirmaciones con mayor puntaje fueron “Me gustaría usar la aplicación frecuentemente” con un promedio de 4,36. La afirmación que tuvo menor fue “Necesitaría el apoyo de un técnico/profesor para utilizar la aplicación.” Con un promedio de 1,96. A continuación, se realizó el cálculo para conocer la valoración de la herramienta web desarrollada. En la tabla 18 se muestra el promedio SUS, la desviación SUS y el coeficiente de variación, lo que permite conocer el nivel de usabilidad de la herramienta web.

Tabla 18: Valoración SUS

Promedio SUS	75,20
Desviación SUS	13,01
Coefficiente de variación	0,17

**FUENTE: INVESTIGACIÓN
ELABORADO POR: AUTORES**

Con los resultados obtenidos, se determinó que según los rangos de usabilidad que se presentan en la tabla 19.

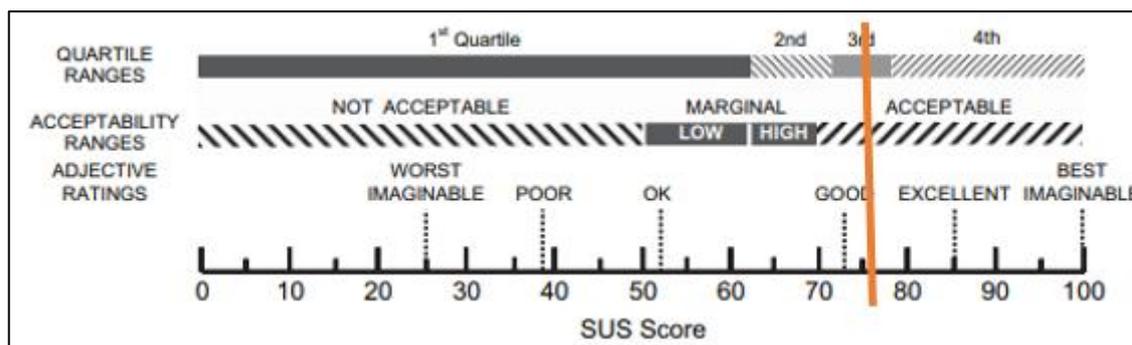
Tabla 19: Rango de usabilidad

Categoría	Rango
Clasificaciones de adjetivos	Entre bueno y excelente
Rangos de aceptabilidad	Aceptables
Rango de cuartil	Cuartil 3

**FUENTE: INVESTIGACIÓN
ELABORADO POR: AUTORES**

En la ilustración 29 se marca el rango de usabilidad determinado para la herramienta web desarrollada.

Ilustración 29: Rango de usabilidad



FUENTE: Bangor et al. [47]

Además, en el cuestionario SUS se añadieron 4 preguntas abiertas. En donde la primera interrogante era “Describa el tipo de usuarios que usted considera que podrían utilizar la aplicación, de manera tan específica como estime pertinente”, las respuestas a esta pregunta fueron variada ya que los usuarios determinaron que la aplicación podía ser utilizada por estudiantes universitarios hasta el personal de recursos humanos de las empresas. Sin

embargo, se encontraron respuestas puntuales en las cuáles determinaban que los usuarios interesados en utilizar la herramienta web estaban relacionados con el ámbito universitario.

Para la pregunta “Si considera que la aplicación puede ser mejorada, indique sus recomendaciones”, los usuarios argumentaron lo siguiente: “Anexar a otras bases de datos que ofertan trabajo”, “Aumentar las carreras analizadas”. De manera similar ocurre para la pregunta “Si considera que la aplicación es útil, describa los usos que le daría.”, en la cual los usuarios manifestaron lo siguiente: “Lo usaría para buscar las ofertas laborales de mi carrera”, “Para saber que habilidades debo adquirir o mejorar.”, y “Buscar trabajo, generar información sobre empleo y desempleo de una ciudad o Provincia”.

Por último, la pregunta “Si tiene algún comentario adicional sobre la aplicación, escríbalo a continuación.”. En esta pregunta se recibieron comentarios positivos para mejorar la herramienta web, como las siguientes: “Agregar información para las otras facultades”, “Me parece muy útil y necesario”. En conclusión, la evaluación de la herramienta web tuvo resultados positivos.

4.5. Discusión

Las librerías utilizadas para el Web Scraping sirvieron para realizar la extracción de los anuncios de empleos que se encontraban disponibles en Indeed Ecuador para formar el conjunto de datos con el que posteriormente se realizaría el SNA. Además, las librerías ofrecieron las opciones para crear gráficos interactivos, logrando presentar las relaciones entre las habilidades con las áreas de conocimiento de manera sencilla.

En el marco referencial se mencionaron algunos trabajos que utilizan la técnica de Web Scraping y SNA para conocer las habilidades y sus relaciones. Por ejemplo en el artículo [41], utilizan el SNA para analizar las distintas habilidades laborales que requieren dentro del mercado laboral de Singapur. Haciendo uso del SNA, identificaron los puestos de trabajo en donde se necesitaban habilidades similares. Sin embargo, no presentan sus resultados en una herramienta web disponible para que varios usuarios observen sus hallazgos.

Es importante resaltar que existen investigaciones que han optado por utilizar el SNA como método para encontrar relaciones entre los requisitos que se solicitan en los empleos, haciendo uso de la extracción de anuncios de empleos disponibles en varios portales de empleo. En contraste con el presente proyecto de investigación, donde se determinaron las habilidades con mayor demanda, además, de presentar los resultados en un aplicativo web.

CAPÍTULO V

CONCLUSIONES Y RECOMENDACIONES

5.1. Conclusiones

En este proyecto de investigación se presenta el análisis de redes sociales para conocer la demanda de habilidades en el mercado laboral según la información que se presenta en los anuncios de empleo. Para esto se crearon las redes en base a las habilidades profesionales requeridas para las distintas carreras de la FCI. En base a la interpretación de cada una de las redes obtenidas se pueden extraer las habilidades que son requeridas en el mercado laboral actual. De tal manera, se llegó a las siguientes conclusiones:

- La técnica de Web Scraping para extraer datos desde páginas web utilizada para la extracción de los anuncios de empleo. Facilitó la recolección de 512 anuncios de empleo, a través de un algoritmo que analizó el contenido del portal Indeed. Además, se determinó para cada carrera la cantidad de anuncios y habilidades correspondientes que se muestran en la tabla 2. Por lo tanto, la técnica aplicada es eficiente para recolectar datos desde páginas web como el caso de los anuncios de empleo disponibles en el portal de Indeed.
- El análisis de redes sociales facilitó que se identifiquen las relaciones existentes entre las distintas carreras de la Facultad de Ciencias de la Ingeniería y las habilidades profesionales que se requieren en cada una de ellas por parte de los empleadores. El análisis de las métricas de centralidad, cercanía y densidad de las redes de habilidades profesionales determinaron que las habilidades con mayor demanda se encuentran relacionadas con el desarrollo de software y manejo de datos, siendo así SQL, Java y Javascript las habilidades profesionales con mayor número de enlaces según el indicador de centralidad. Se concluye, que los algoritmos implementados en el análisis de redes sociales favorecieron en el reconocimiento de las relaciones existentes entre las habilidades profesionales requeridas para las distintas carreras, además, el análisis de las métricas de la red permitió conocer las habilidades profesionales con mayor influencia dentro de las redes de habilidades profesionales.
- El desarrollo de una aplicación web para la visualización de los resultados del SNA, permitió presentar las redes de habilidades, el top 10 y la nube de palabras de las habilidades con mayor demanda de manera sencilla e interactiva, clasificándolas por cada carrera de la FCI. En conclusión, la aplicación web desarrollada cumple con el

propósito de presentar de manera sencilla y clara las relaciones que existen entre las distintas carreras y las habilidades que se requieren.

- De acuerdo al reporte generado por la aplicación web, se observa que el contenido cumple con proporcionar información sobre las habilidades con mayor demanda para las distintas carreras de la FCI. Además, por medio de los filtros disponibles en la aplicación web se puede seleccionar las carreras y la cantidad de habilidades para que se visualicen en el reporte. En definitiva, la generación del reporte agrupa el número de veces que una habilidad es solicitada con cada carrera dando a conocer la demanda de dicha habilidad dentro del campo laboral.
- El cuestionario de usabilidad se realizó con un total de 25 participantes que colaboraron de manera voluntaria para evaluar la aplicación web. Luego de analizar las respuestas, se obtuvo un promedio de usabilidad del 75.2 que, basándonos en la escala de puntuación y valoración de Bangor, la valoración de la aplicación se encuentra en un rango aceptable. Por lo tanto, se concluye que el uso de la aplicación proporciona material de apoyo para conocer las tecnologías y/o herramientas con mayor demanda en el mercado laboral actual.

5.2. Recomendaciones

Una vez concluido el proyecto de investigación, se realizan las siguientes recomendaciones:

- Ampliar el estudio de la investigación con de extracción de datos en otros portales de empleo para la obtención de anuncios disponibles en Ecuador.
- Desarrollar una aplicación móvil que realice la técnica de Web Scraping en tiempo real en base a las distintas carreras, para ser visualizado por los usuarios mediante SNA.
- Realizar un análisis de las habilidades de las distintas carreras de la Facultad Ciencias de la Ingeniería, para ayudar a los solicitantes a saber si cuentan con las competencias necesarias que respondan a la demanda de mercado.
- Realizar evaluación de la herramienta por parte del personal académico de la UTEQ o de alguna otra universidad interesada.

CAPÍTULO VI
BIBLIOGRAFÍA

- [1] N. Dawson, M. A. RizoIU, B. Johnston, and M. A. Williams, "Adaptively selecting occupations to detect skill shortages from online job ads," *Proc. - 2019 IEEE Int. Conf. Big Data, Big Data 2019*, pp. 1637–1643, 2019, doi: 10.1109/BigData47090.2019.9005967.
- [2] I. Apatsidis, K. Georgiou, N. Mittas, and L. Angelis, "A Study of Remote and On-site ICT Labor Market Demand using Job Offers from Stack Overflow," *Euromicro Conf. Softw. Eng. Adv. Appl.*, pp. 252–259, 2021, doi: 10.1109/seaa53835.2021.00039.
- [3] M. Humburg and R. van der Velden, "Skills and the graduate recruitment process: Evidence from two discrete choice experiments," *Econ. Educ. Rev.*, vol. 49, pp. 24–41, 2015, doi: 10.1016/j.econedurev.2015.07.001.
- [4] V. Saunders and K. Zuzel, "Evaluating Employability Skills: Employer and Student Perceptions," *Biosci. Educ.*, vol. 15, no. 1, pp. 1–15, 2016, doi: 10.3108/beej.15.2.
- [5] P. Rogiers, S. Viaene, and J. Leysen, "The digital future of internal staffing: A vision for transformational electronic human resource management," *Intell. Syst. Accounting, Financ. Manag.*, vol. 27, no. 4, pp. 182–196, 2020, doi: 10.1002/isaf.1481.
- [6] P. Garg, R. Rani, and S. Miglani, "Mining Professional's Data from LinkedIn Puneet," *Proc. - 2015 5th Int. Conf. Adv. Comput. Commun. ICACC 2015*, pp. 98–101, 2016, doi: 10.1109/ICACC.2015.35.
- [7] S. vanden Broucke and B. Baesens, *Practical Web Scraping for Data Science: Best Practices and Examples with Python*. New York: Apress, 2018.
- [8] B. Zhao, "Web Scraping," *Encycl. Big Data*, no. December, 2020, doi: 10.1007/978-3-319-32001-4.
- [9] R. Diouf, E. N. Sarr, O. Sall, B. Birregah, M. Bousso, and S. N. Mbaye, "Web Scraping: State-of-the-Art and Areas of Application," *Proc. - 2019 IEEE Int. Conf. Big Data, Big Data 2019*, pp. 6040–6042, 2019, doi: 10.1109/BigData47090.2019.9005594.
- [10] V. Krotov and M. F. Tennyson, "Web Scraping in the R Language : A Tutorial," *J. Midwest Assoc. Inf. Syst.*, vol. 20, no. 1, 2021, doi: 10.17705/3jmw.000066.
- [11] B. Y. A. J. Dreyer, "Internet ' Data Scraping ': Aprimer for Counseling Clients," *A New York Law Journal Special Section*, 2013.
- [12] B. R. Paul and L. Elder, *The Thinker's Guide to Understanding the Foundations of Ethical Reasoning: Based on" Critical Thinking Concepts Et Principles*. 2013.
- [13] K. Dale, *Data Visualization with Python and Javascript*, vol. 53, no. 9. O'Reilly

- Media, Inc., 2016.
- [14] K. Jarmul and R. Lawson, *Python Web Scraping*, vol. 53, no. 9. United Kingdom: Packt Publishing Ltd., 2017.
- [15] J. Ward, *Instant PHP Web Scraping: Get up and running with the basic techniques of web scraping using PHP*. United Kingdom: Packt Publishing Ltd., 2016.
- [16] L. Richardson, “Beautiful Soup Documentation — Beautiful Soup 4.9.0 documentation,” 2020. <https://www.crummy.com/software/BeautifulSoup/bs4/doc/> (accessed Jul. 23, 2022).
- [17] J. Hedley, “jsoup: Java HTML parser, built for HTML editing, cleaning, scraping, and XSS safety,” *MIT License*, 2009. <https://jsoup.org/> (accessed Jul. 23, 2022).
- [18] T. Cervenka, “Jaunt - Java Web Scraping & JSON Querying,” *Apache Licence*, 2018. <https://jaunt-api.com/> (accessed Jul. 23, 2022).
- [19] R. Arifanto, Y. D. W. Asnar, and M. M. I. Liem, “Domain Specific Language for Web Scraper Development,” *Proc. 2018 5th Int. Conf. Data Softw. Eng. ICoDSE 2018*, pp. 1–6, 2018, doi: 10.1109/ICODSE.2018.8705842.
- [20] V. Kafidoff, E. Klem, J. Davis, N. Singh, G. Chandra, and C. Severance, “Cheerio.” <https://cheerio.js.org/> (accessed Jul. 23, 2022).
- [21] L. Ou-Yang, “Newspaper3k: Article scraping & curation,” 2013. <https://newspaper.readthedocs.io/en/latest/> (accessed Jul. 23, 2022).
- [22] L. DigitalPebble, “StormCrawler,” 2014. <http://stormcrawler.net/> (accessed Jul. 23, 2022).
- [23] M. Saqr and A. Alamro, “The role of social network analysis as a learning analytics tool in online problem based learning,” *BMC Med. Educ.*, vol. 19, no. 1, pp. 1–11, 2019, doi: 10.1186/s12909-019-1599-6.
- [24] O. Serrat, “Social Network Analysis,” *Blackwell Guid. to Res. Methods Biling. Multiling.*, no. February, pp. 263–274, 2009, doi: 10.1002/9781444301120.ch15.
- [25] S. Afra, T. Ozyer, J. Rokne, and R. Alhaji, “NetDriller version 2: A powerful social network analysis tool,” *IEEE Int. Conf. Data Min. Work. ICDMW*, vol. 2018-Novem, pp. 1475–1480, 2019, doi: 10.1109/ICDMW.2018.00211.
- [26] F. Grando, D. Noble, and L. C. Lamb, “An analysis of centrality measures for complex and social networks,” *2016 IEEE Glob. Commun. Conf. GLOBECOM 2016 - Proc.*, 2016, doi: 10.1109/GLOCOM.2016.7841580.
- [27] A. Farooq, G. J. Joyia, M. Uzair, and U. Akram, “Detection of influential nodes using social networks analysis based on network metrics,” *2018 Int. Conf. Comput. Math.*

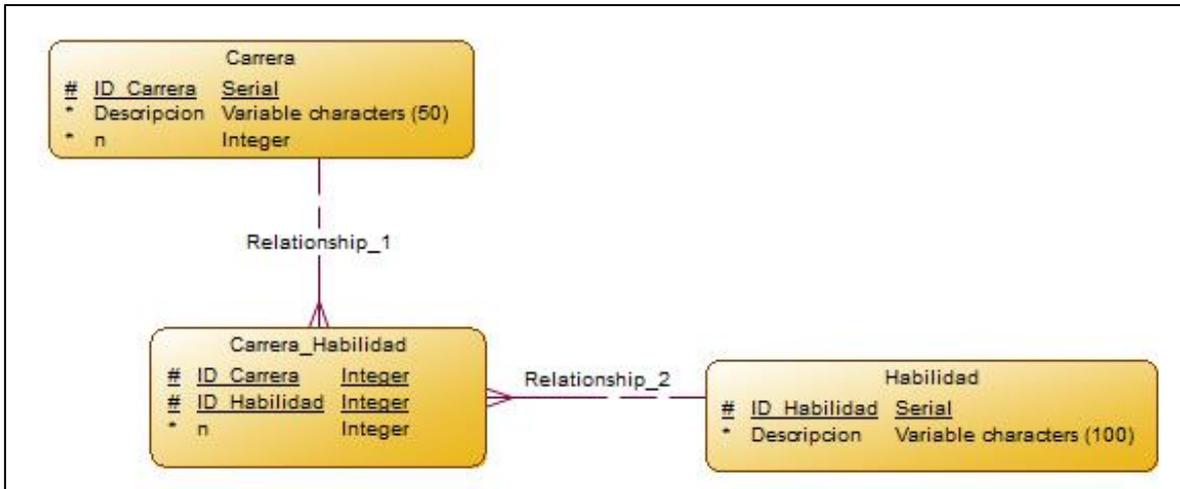
- Eng. Technol. Inven. Innov. Integr. Socioecon. Dev. iCoMET 2018 - Proc.*, vol. 2018-Janua, pp. 1–6, 2018, doi: 10.1109/ICOMET.2018.8346372.
- [28] M. Bastian, S. Heymann, and M. Jacomy, “Gephi: An open source software for exploring and manipulating networks. BT - International AAAI Conference on Weblogs and Social,” *Int. AAAI Conf. Weblogs Soc. Media*, pp. 361–362, 2009, [Online]. Available: <https://gephi.org/>.
- [29] P. Rani and J. Shokeen, “A survey of tools for social network analysis,” *Int. J. Web Eng. Technol.*, vol. 16, no. 3, pp. 189–216, 2021, doi: 10.1504/IJWET.2021.119879.
- [30] B. Lu, H. Sun, P. Harris, M. Xu, and M. Charlton, “Shp2graph: Tools to convert a spatial network into an igraph graph in R,” *ISPRS Int. J. Geo-Information*, vol. 7, no. 8, 2018, doi: 10.3390/ijgi7080293.
- [31] “Social Network Analysis in R | R-bloggers.” <https://www.r-bloggers.com/2021/04/social-network-analysis-in-r/> (accessed Jul. 31, 2022).
- [32] C. University, “Meaning of SKILL in the Cambridge English Dictionary,” *Cambridge University Press* 2022. <https://dictionary.cambridge.org/es/diccionario/ingles/skill> (accessed Aug. 23, 2022).
- [33] M. Sabin *et al.*, *Information Technology Curricula 2017*. Association for Computing Machinery New York NY United States, 2017.
- [34] Cambridge University, “Meaning of ABILITY in the Cambridge English Dictionary,” 2022. <https://dictionary.cambridge.org/es/diccionario/ingles-espanol/ability?q=abilities> (accessed Sep. 03, 2022).
- [35] A. Sopa *et al.*, “Hard skills versus soft skills: Which are more important for Indonesian employees innovation capability,” *Int. J. Control Autom.*, vol. 13, no. 2, pp. 156–175, 2020.
- [36] I. P. Monteiro and M. B. Correia, “Transforming a company’s staffing process: Implementing e-recruitment,” *J. Spat. Organ. Dyn.*, vol. 7, no. 2, pp. 144–157, 2019.
- [37] Adecco and Infoempleo, “Informe Infoempleo & Adecco 2020,” España, 2021. [Online]. Available: <https://cdnazure.infoempleo.com/infoempleo/documentacion/Informe-infoempleo-adecco-2020.pdf>.
- [38] P. G. Lovaglio, M. Cesarini, F. Mercurio, and M. Mezzanzanica, “Skills in demand for ICT and statistical occupations: Evidence from web-based job vacancies,” *Stat. Anal. Data Min.*, vol. 11, no. 2, pp. 78–91, 2018, doi: 10.1002/sam.11372.
- [39] M. F. Benavidez Torres, “La Importancia De La Gestión Curricular Universitaria En

- Programas a Distancia, Estudio Institución De Educación Superior Suramericana,” *In Crescendo*, vol. 10, no. 1, p. 13, 2019, doi: 10.21895/incres.2019.v10n1.02.
- [40] M. S. Vanessa, V. R. Garcés, and S. C. José, “Ecuador, leyes de educación y currículo nacional. Comparación,” no. 1, pp. 401–405, 2019, [Online]. Available: <http://revistamapa.com>.
- [41] D. Ming *et al.*, “A Social Network Analysis of Jobs and Skills,” in *Institutional Knowledge at Singapore Management University A social network analysis of jobs and skills*, 2020, pp. 5747–5749, doi: 10.1109/bigdata50022.2020.937813010.1109/bigdata50022.2020.9378130.
- [42] A. Phaphuangwittayakul, S. Saranwong, S. N. Panyakaew, P. Inkeaw, and J. Chaijaruwanich, “Analysis of Skill Demand in Thai Labor Market from Online Jobs Recruitments Websites,” *Proceeding 2018 15th Int. Jt. Conf. Comput. Sci. Softw. Eng. JCSSE 2018*, pp. 1–5, 2018, doi: 10.1109/JCSSE.2018.8457393.
- [43] European Parliament and of the Council, “Regulation on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive (General Data Protection Regulation),” *Off. J. Eur. Communities*, vol. OJ L 119/1, pp. 1–88, 2016, [Online]. Available: <http://data.europa.eu/eli/reg/2016/679/oj>.
- [44] Asamblea Nacional del Ecuador, “Ley Orgánica de Transparencia y Acceso a la Información Pública,” no. 337, pp. 1–13, 2004, [Online]. Available: <http://biblioteca.defensoria.gob.ec/handle/37000/3372>.
- [45] Asamblea nacional, “Ley Orgánica de protección de datos personales,” *Ley Orgánica Protección Datos Pers.*, no. 459, pp. 1–40, 2021, [Online]. Available: [file:///C:/Users/Usuario/Documents/3ER SEMESTRE DERECHO/Derecho Informático/Reglamentos y Leyes/Ley Orgánica de Protección de Datos Personales.pdf](file:///C:/Users/Usuario/Documents/3ER%20SEMESTRE%20DERECHO/Derecho%20Informático/Reglamentos%20y%20Leyes/Ley%20Orgánica%20de%20Protección%20de%20Datos%20Personales.pdf).
- [46] “California Consumer Privacy Act (CCPA),” *Proyecto de Ley de la Asamblea*, 2018. https://leginfo.legislature.ca.gov/faces/codes_displayText.xhtml?division=3.&part=4.&lawCode=CIV&title=1.81.5 (accessed Sep. 25, 2022).
- [47] A. Bangor, P. T. Kortum, and J. T. Miller, “An empirical evaluation of the system usability scale,” *Int. J. Hum. Comput. Interact.*, vol. 24, no. 6, pp. 574–594, 2008, doi: 10.1080/10447310802205776.

CAPÍTULO VII

ANEXOS

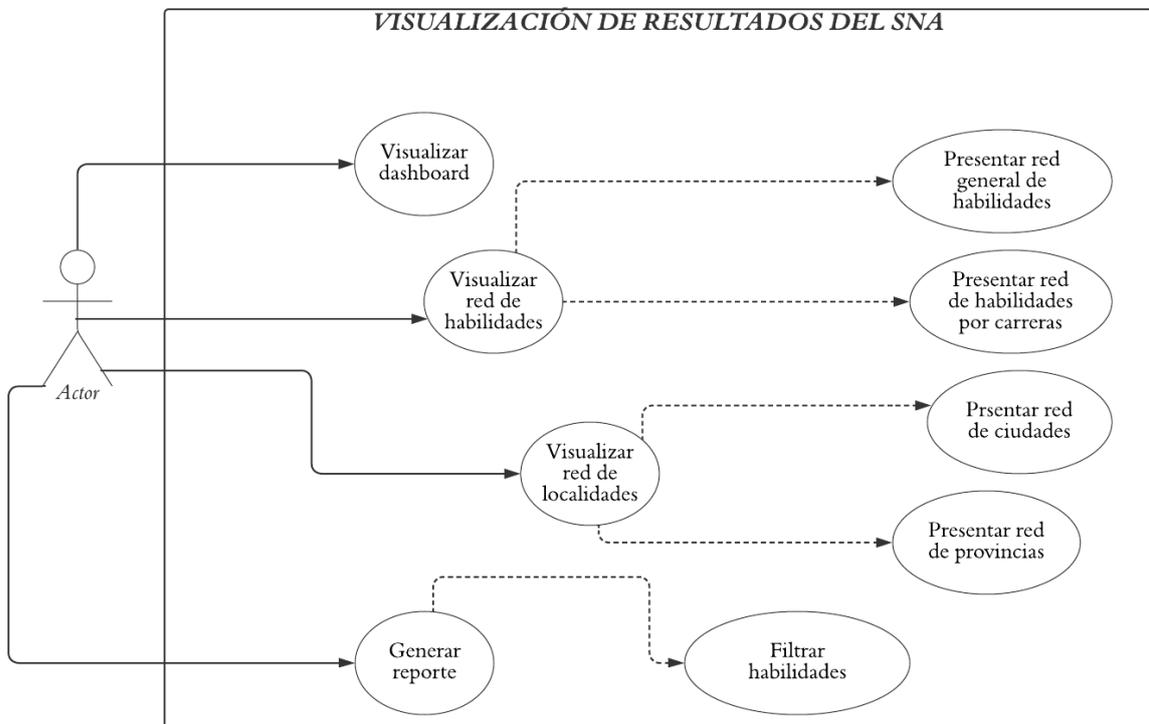
Anexo 1: Modelo Físico de la Base de Datos para generar el reporte de habilidades



Anexo 2: Documentación del software

En la siguiente sección se presentan los diagramas de casos de uso y la explicación detallada de cada uno.

Anexo 2.1. Casos de Uso



Anexo 2.2. Caso de uso extendido – Visualizar dashboard

Caso de uso	Visualizar dashboard
Actor	Investigador
Propósito	Mostrar datos estadísticos de la cantidad de anuncios de empleos y habilidades profesionales extraídas.
Resumen	El investigador visualiza los datos obtenidos mediante el Web Scraping.
Tipo	Primario
Flujo normal de eventos	
Acción del actor	Respuesta del sistema
1. El caso de uso inicia cuando el usuario ingresa al dashboard de la aplicación.	2. El sistema muestra la información estadística de los datos obtenidos.
Flujo alterno de eventos	

Anexo 2.3. Caso de uso extendido – Visualizar red de habilidades

Caso de uso	Visualizar red de habilidades
Actor	Investigador
Propósito	Mostrar las redes obtenidas de las habilidades profesionales.
Resumen	El investigador visualiza los resultados obtenidos en el SNA.
Tipo	Primario
Flujo normal de eventos	
Acción del actor	Respuesta del sistema
1. El caso de uso inicia cuando el usuario desea visualizar las redes de las habilidades profesionales.	2. El sistema muestra la información de las redes.
Flujo alterno de eventos	

Sección: Visualizar red general de habilidades profesionales	
Acción del actor	Respuesta del sistema
	1. El sistema muestra la información del resultado del SNA general de las habilidades.
Sección: Visualizar red de habilidades por carreras	
Acción del actor	Respuesta del sistema
1. El caso de uso inicia cuando el usuario selecciona la carrera que desea visualizar las redes.	2. El sistema muestra la información relacionada con el SNA de la carrera seleccionada.

Anexo 2.4. Caso de uso extendido – Visualizar red de localidades

Caso de uso	Visualizar red de localidades
Actor	Investigador
Propósito	Mostrar las redes obtenidas de la disponibilidad de ofertas laborales.
Resumen	El investigador visualiza los resultados obtenidos en el SNA.
Tipo	Primario
Flujo normal de eventos	
Acción del actor	Respuesta del sistema
1. El caso de uso inicia cuando el usuario desea visualizar las redes de la disponibilidad de mano de obra en distintas partes del Ecuador.	3. El sistema muestra la información de las redes.
Flujo alterno de eventos	
Sección: Visualizar red de ciudades	
Acción del actor	Respuesta del sistema
	1. El sistema muestra la información del resultado del SNA general de las ofertas laborales en distintas ciudades.
Sección: Visualizar red de provincias	

Acción del actor	Respuesta del sistema
	2. El sistema muestra la información del resultado del SNA general de las ofertas laborales por provincias.

Anexo 2.5. Caso de uso extendido – Generar reporte

Caso de uso	Generar reporte
Actor	Investigador
Propósito	Generar un reporte con las habilidades requeridas en cada una de las carreras según los anuncios de empleo.
Resumen	El investigador genera un reporte con las habilidades profesionales.
Tipo	Primario
Flujo normal de eventos	
Acción del actor	Respuesta del sistema
1. El caso de uso inicia cuando el usuario selecciona la opción para generar un reporte. 3.El usuario selecciona la carrera y la cantidad de datos que desea para generar el reporte.	2. El sistema muestra la interfaz para filtrar las carreras. 4.El sistema genera el reporte con los datos que ingresó el usuario.
Flujo alterno de eventos	

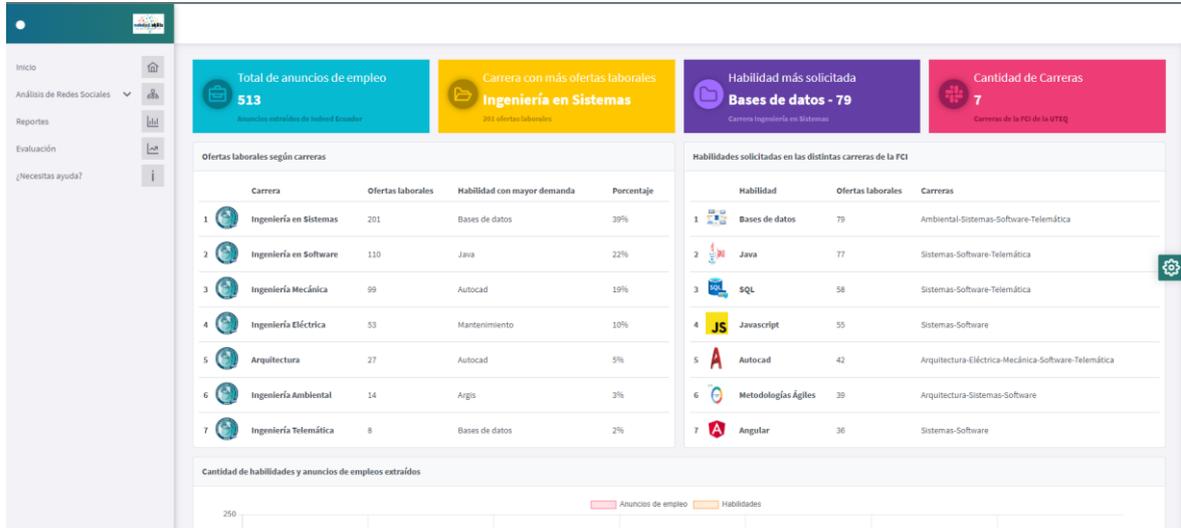
Anexo 3: Reporte generado desde la aplicación web

Habilidades solicitadas en empleos según indeed.com.ec	
Carrera(s): Arquitectura, Ingeniería Ambiental, Ingeniería Eléctrica, Ingeniería en Sistemas	
Habilidad	Total
Arquitectura	27
Autocad	16
Microsoft office	7
Revit	4
Sketchup	4
Lumion	3
Ingeniería Ambiental	14
Argis	2
Mapinfo	2
Normas iso	2
Proyectos de sostenibilidad	2
Agricultura sostenible	1
Ingeniería Eléctrica	53
Mantenimiento	14
Diseño eléctrico	12
Autocad	10
Gestión de proyectos	10
Plc	10
Ingeniería en Sistemas	201
Bases de datos	48
Java	44
Sql	35
Javascript	34
Redes	27
domingo 09 octubre 2022	
Página 1 de 1	



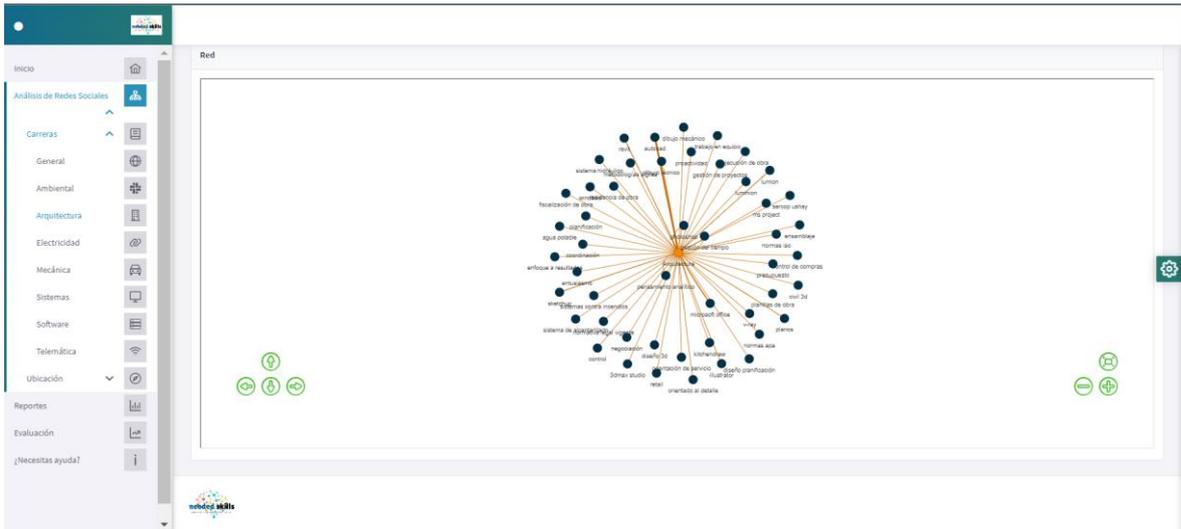
Anexo 4: Manual de usuario de la aplicación web

Anexo 4.1. Dashboard de la aplicación

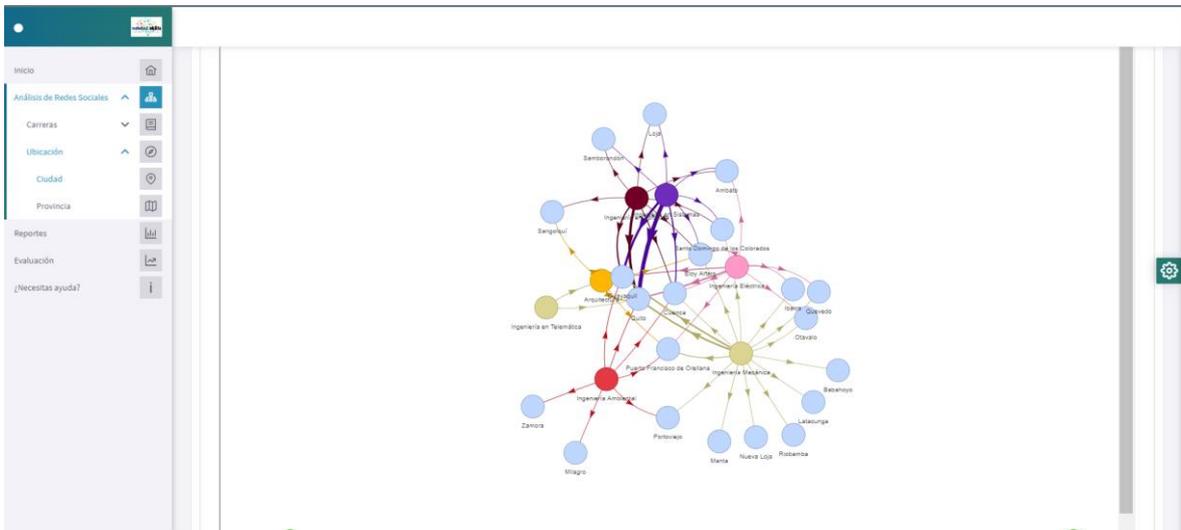


Anexo 4.2. Visualización de las redes de habilidades profesionales por carrera





Anexo 4.4. Visualización de redes clasificadas por localización



Anexo 4.5. Generación del reporte

Reporte del SNA de habilidades profesionales

En esta sección se genera el reporte con la información de las habilidades profesionales según área de conocimiento.

Carrera: Total a mostrar:

+

Búsqueda

Carrera	Mostrar	Acciones
No se encontraron registros.		

robot skills