

Universidad de Lima
Facultad de Ingeniería y Arquitectura
Carrera de Ingeniería de Sistemas



**PLATAFORMA DE RECOMENDACION DE
HABILIDADES TECNOLOGICAS SEGUN
PUESTO DE TRABAJO PARA
PROFESIONALES DE TI, EN FUNCION DE
LA DEMANDA EN LAS BOLSAS DE
TRABAJO DIGITALES**

Trabajo de suficiencia profesional para optar el Título Profesional de Ingeniero de
Sistemas

Diandra Esli Sanchez Bermudez
Código 20101873

Asesor
Pedro Humberto Saravia Torres

Lima – Perú

Febrero, 2020

**TECHNICAL SKILLS RECOMMENDATION
PLATFORM ACCORDING TO WORK
POSITION FOR IT PROFESSIONALS,
DEPENDING ON DIGITAL WORK BAGS
USING SCRAPING AND NLP TECHNIQUES**

SCIENTIA ET PRAXIS
MCMLXII

TABLA DE CONTENIDO

RESUMEN.....	x
ABSTRACT.....	xi
CAPITULO I: INTRODUCCION	1
CAPITULO II: CONCEPTOS Y FUNDAMENTOS.....	8
2.1 Web Scraping	8
2.2 Procesamiento del Lenguaje Natural.....	10
2.2.1. Coincidencia aproximada de cadenas	14
2.3 Trabajadores del conocimiento.....	18
2.4 Modelo de Negocio Freemium.....	19
2.5 Metodologías Ágiles.....	20
2.6 Design Thinking	21
CAPITULO III : FUNDAMENTACIÓN DEL PROYECTO	24
3.1 Fundamentación de la deseabilidad del proyecto	24
3.1.1. Descubrimiento	25
3.2 Fundamentación de la factibilidad del proyecto.....	40
3.2.1. Canvas para “usuarios de la herramienta”	40
3.2.2 Canvas para “empresas”	43
3.2.3 Propuesta de modelo de las tecnologías a usar para cada proceso del desarrollo	45
3.3 Fundamentación de la viabilidad del proyecto	54
3.3.1 Descripción del proyecto	54
3.3.2 Ingresos.....	55
3.3.3 Flujo de caja.....	66
CAPITULO IV: DEFINICIÓN DEL PROYECTO	69
4.1 Diseño de propuesta de valor.....	69
4.2 Alcance	74
4.3 Objetivos generales y específicos	75
4.3.1. Objetivo general.....	75
4.3.2 Objetivos específicos	75

4.4	Modelo de negocio propuesto.....	76
4.4.1	Modelo de negocio	76
4.4.2	Roles de equipo de proyecto.....	78
CAPITULO V : DESARROLLO DEL PROTOTIPO		84
5.1	Empatizar.....	84
5.1.1	Propósito.....	85
5.1.2	Método.....	85
5.1.3	Participantes.....	86
5.1.4	Ubicación.....	86
5.2	Definición	87
5.3	Ideación.....	90
5.4	Prototipado.....	91
5.5	Testeo.....	95
5.5.1.	User Test Versión 1	95
5.5.2.	User test versión 2.....	103
5.5.3.	Casos de Uso 1.....	112
5.5.4.	Caso de Uso 2	114
CONCLUSIONES.....		118
RECOMENDACIONES.....		119
GLOSARIO DE TÉRMINOS.....		120
REFERENCIAS.....		121
BIBLIOGRAFÍA.....		126
ANEXOS.....		127

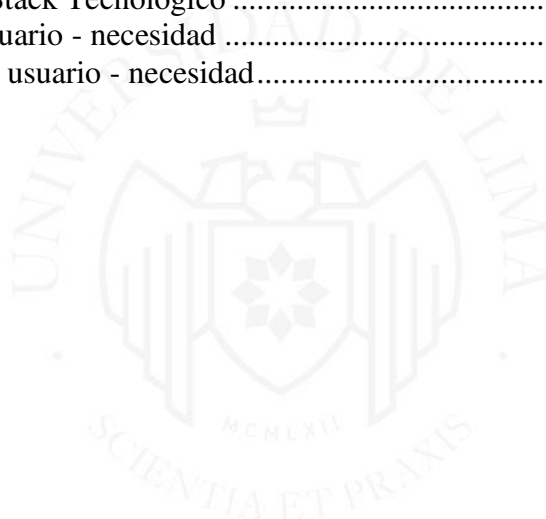
ÍNDICE DE FIGURAS

Figura 1.1	Porcentaje de personas que desean adquirir habilidades digitales para asegurar su empleabilidad.....	3
Figura 1.2	Porcentaje de personas que piensan que necesitan adquirir diferentes habilidades digitales de las que actualmente poseen (%).....	4
Figura 1.3	Desafíos para diseñar política de habilidades inclusivas para la era digital.....	5
Figura 1.4	Uso de datos de puestos de trabajo Burning Glass para inferir la especialización y el nivel de habilidad.....	6
Figura 2.1	La ecuación básica de scraping.....	9
Figura 2.2	Generaciones vectores de incrustación de palabras para frases de habilidad.....	13
Figura 2.3	Fórmula distancia de Jaro Winkler.....	16
Figura 2.4	Representación vectorial en matriz de habilidades.....	17
Figura 2.5	Representación de relación de skills.....	18
Figura 2.6	Definición de una satisfactoria innovación: Intersección de 3 fuerzas:.....	21
Figura 2.7	Etapas del Design Thinking.....	22
Figura 2.8	Marco del modelo diamante para el proceso del design thinking.....	23
Figura 3.1	Esquema para fundamentación de la deseabilidad del cliente.....	24
Figura 3.2	Encuesta laboral 2018.....	26
Figura 3.3	Proporción de empleo en alta demanda por nivel de habilidad.....	27
Figura 3.4	Indicadores clave de la Competitividad del Talento Global en Perú 2017.....	28
Figura 3.5	Incremento en el % de escasez de talento global en los últimos 10 años.....	30
Figura 3.6	Estrategias de empleadores están usando para superar la escasez de talentos.....	31
Figura 3.7	Estructura de la entrevista realizada.....	32
Figura 3.8	Medios por los que actualizan habilidades (%).....	33
Figura 3.9	User Persona de usuario premium.....	34
Figura 3.10	User Persona de usuario regular.....	35
Figura 3.11	Job to be done de definición sobre tareas necesitan satisfacer los usuarios.....	37
Figura 3.12	Canvas para el usuario de HAB.....	40
Figura 3.13	Proceso ETL.....	45
Figura 3.14	Keywords para extracción de puestos de trabajo.....	46
Figura 3.15	Descripción de trabajo extraídos por puestos de trabajo.....	46
Figura 3.16	Extracción y almacenamiento de puestos de trabajo.....	47
Figura 3.17	Dataset de habilidades técnicas.....	48
Figura 3.18	Pre procesamiento de descripciones de trabajo.....	49
Figura 3.19	Resultado comparación de algoritmos de similitud de cadenas.....	51
Figura 3.20	Procesamiento de datos en MS TL.....	53
Figura 3.21	Proceso de Web Scraping y NLP.....	54
Figura 3.22	Dificultad para cubrir vacantes según el tamaño de la empresa.....	56
Figura 3.23	Demanda potencial de empresas.....	57
Figura 3.24	Demanda potencial de personas.....	57
Figura 3.25	Precios de consultoría Open Talent.....	59
Figura 4.1	Definición de la propuesta de valor.....	70
Figura 4.2	Priorización de necesidades del perfil de cliente.....	71
Figura 4.3	Lean Canvas actualizado del modelo de negocio.....	73
Figura 4.4	Definición de MVP de HAB.....	74
Figura 4.5	Comparación de tarifas de Cloud Computing.....	80
Figura 4.6	Diagrama de Infraestructura de aplicación.....	81

Figura 4.7 Diagrama de Despliegue UML.....	81
Figura 4.8 Diagrama de Componentes UML.....	82
Figura 4.9 Diagrama de Flujo UML	83
Figura 5.1 Objetivos por etapas del proceso design thinking	84
Figura 5.2 Mapa de empatía.....	87
Figura 5.3 User Journey Map.....	89
Figura 5.4 Brainstorming y clusterización de ideas en herramienta Miró	90
Figura 5.5 Página Job Tech de referencia	91
Figura 5.6 Versión 0 a mano alzada.....	92
Figura 5.7 Versión 1: Pantalla de inicio, login y registro en herramienta “Marvel”	93
Figura 5.8 Versión 1: Pantalla habilidades y cursos en herramienta “Marvel”	94
Figura 5.9 User Test en la herramienta Shape	96
Figura 5.10 Pregunta 1 y 2 del user test versión 1 realizada en la herramienta Shape	97
Figura 5.11 Pregunta 3 y 4 del user test versión 1 realizado en la herramienta Shape	98
Figura 5.12 Pregunta 5 del test de usuario versión 1 realizada en la herramienta Shape	99
Figura 5.13 Respuestas de los 8 usuarios a la pregunta abierta sobre la percepción que tenían del propósito de la página respecto a las pantallas realizada en el user test versión 1	100
Figura 5.14 Resultados de respuestas de usuarios realizada en el user test versión 1. (%) ..	101
Figura 5.15 Resultados de User Test versión 1.....	102
Figura 5.16 Respuestas de usuarios realizada en user test e nuevas funcionalidades.....	102
Figura 5.17 Respuestas sobre percepción del producto de acuerdo a las pantallas mostradas, realizadas en el user test versión 2	103
Figura 5.18 Dashboard de respuestas de usuarios realizada en el test de usuario versión 2 (%)	104
Figura 5.19 Resultados del user test versión 2.....	105
Figura 5.20 Página de Bienvenida	106
Figura 5.21 Registro de perfil	107
Figura 5.22 Registro de habilidades.....	108
Figura 5.23 Página de inicio de sesión.....	109
Figura 5.24 Página principal de habilidades y cursos	110
Figura 5.25 Página de listado de habilidades relacionadas.....	111
Figura 5.26 Registro Perfil.....	112
Figura 5.27 Registro de habilidades técnicas.....	112
Figura 5.28 Datos guardados del perfil en base de datos.....	113
Figura 5.29 Habilidades y pesos relacionados para el skill “Appium”.....	113
Figura 5.30 Dashboard de habilidades técnicas y cursos.....	114
Figura 5.31 Registro Perfil.....	114
Figura 5.32 Registro de Habilidades.....	115
Figura 5.33 Datos guardados del perfil del usuario que se registró en DB	116
Figura 5.34 Skill y pesos relacionados para el skill “BDD”	116
Figura 5.35 Dashboard de habilidades y cursos.....	117

ÍNDICE DE TABLAS

Tabla 2.1	Tabla de descripción de habilidades de empleos.....	14
Tabla 2.2	Explicación de las variables de fórmula de Jaro Winkler.....	16
Tabla 3.1	Tabla de matriz de relación de habilidades con descripciones del trabajo	52
Tabla 3.2	Tabla de descripción de nivel de puestos de trabajo.....	59
Tabla 3.3	Descripción de nivel de puestos de trabajo.....	60
Tabla 3.4	Detalle de la proyección de sueldo/cantidad.....	61
Tabla 3.5	Activos Fijos	62
Tabla 3.6	Detalle del gasto único.....	62
Tabla 3.7	Detalle de gastos mensuales	63
Tabla 3.8	Demanda anual de puestos de tecnología por tipo de empresa.....	64
Tabla 3.9	Ingresos proyectados por año y tipo de empresa	65
Tabla 3.10	Flujo de caja.....	66
Tabla 3.11	Cálculo del VAN y el TIR	67
Tabla 3.12	Análisis del VAN disminuyendo 5% los ingresos por parte de las empresas	68
Tabla 4.1	Roles y responsabilidades	78
Tabla 4.2	Descripción Stack Tecnológico	79
Tabla 5.1	Definición usuario - necesidad	87
Tabla 5.2	Definición de usuario - necesidad.....	88



ÍNDICE DE ANEXOS

Anexo 1 Apelación HIQ LABS V. LINKEDIN por caso Scraping	128
Anexo 2 Guía de Entrevista para profesionales	130
Anexo 3 Factores concurrentes de categorización empresarial: Número de trabajadores y el nivel de ventas anuales.....	132
Anexo 4 Valores de equivalencia de UIT por año	134
Anexo 5 Cuestionario a empresas.....	135
Anexo 6 Número promedio de personas que se demandarán para puestos de trabajo del sector tecnología entre técnicos y profesionales. MINISTERIO DE TRABAJO Y PROMOCIÓN DEL EMPLEO (2018)	137
Anexo 7 Análisis de Stack Tecnológico	139



RESUMEN

El panorama global cambia rápidamente, la velocidad de los avances tecnológicos, la automatización y la transformación digital hacen que los modelos de negocios y sus objetivos creen puestos de trabajo multidisciplinarios y con conocimientos diversos, así mismo las empresas compiten para contratar a los mejores profesionales, es por ello que es fundamental contar con información unificada y centralizada, que brinde una visión de lo que el mercado laboral busca.

Este proyecto propone el desarrollo de una plataforma, que ayude a las personas a visualizar las habilidades técnicas que el mercado laboral requiere con mayor demanda para un puesto de trabajo y brindarles una visibilidad general basada en el análisis de las bolsas de trabajo más importantes del país, así como la sugerencia de organizaciones donde estudiarlas. Dicha plataforma estará basada en el uso de tecnologías emergentes como el procesamiento de lenguaje natural y scraping, para extraer y analizar las habilidades técnicas demandadas en los puestos de trabajo de las principales bolsas laborales digitales.

Por lo tanto, se puede concluir que la información generada brindará una orientación profesional que aún deben adquirir para alcanzar sus objetivos, así como las organizaciones donde poder estudiarlas.

Esta propuesta podría adicionalmente brindar de crecimiento como el procesamiento de la información para evaluar tendencias de contratación que se pueden utilizar para informar y guiar las decisiones, así como creación de contenido educativo de acorde a las necesidades reales de las personas.

Palabras Clave: Web scraping., clasificación ocupacional, procesamiento de lenguaje natural, habilidades técnicas, distancia de Levenshtein

ABSTRACT

The global landscape changes rapidly, the speed of technological advances, automation and digital transformation make business models and their objectives create multidisciplinary jobs with diverse knowledge, as well as companies compete to hire the best professionals, that is why it is essential to have unified and centralized information that provides insight into what the labor market is looking for.

This project proposes the development of a platform, which helps people visualize the technical skills that the labor market requires with greater demand for a job and provide them with a general visibility based on the analysis of the most important job exchanges in the country, as well as the suggestion of organizations where to study them. This platform will be based on the use of emerging technologies such as natural language processing and scraping, to extract and analyze the technical skills demanded in the jobs of the main digital job exchanges.

This proposal could additionally provide growth such as information processing to assess hiring trends that can be used to inform and guide decisions, as well as creating educational content according to the real needs of people.

Keywords: Web Scraping., occupational Classification, natural language processing, technical Skill, levenshtein distance

CAPITULO I: INTRODUCCION

La búsqueda de empleo siempre es un desafío y más aún en el rubro de tecnología, la escasez de habilidades es más crítica y se crea un cuello de botella en el crecimiento de los profesionales de este ámbito. Por ello se necesita que los mismos optimicen la búsqueda de habilidades técnicas que el mercado laboral necesita y desarrollarlas.

Actualmente uno de los más grandes cambios, es el avance de la tecnología y la manera en la que ha impactado en la forma de trabajo de las personas. Muchas tecnologías emergentes impactan de manera transversal a todas las industrias, y genera la oportunidad de incrementar las capacidades y calzar en las necesidades disruptivas de los puestos de trabajo que las empresas impactadas necesitan en mayor o menor grado por dichas tecnologías. Según Ellis (2019), el 81% por ciento de encuestados de empresas indican que enfrenta una escasez de habilidades críticas que impiden a su organización se mantenga al día con el ritmo de estos cambios.

Con la introducción de la tecnología, las personas tienen que actualizarse constantemente para asegurar que sus habilidades tanto técnicas como blandas son las necesarias para los puestos de trabajos deseados. Los millennials tienden a ser menos impulsados por las posesiones materiales y los viejos marcadores de estatus, como la propiedad; ellos valoran las experiencias sobre las "cosas", en sus carreras, están más interesados en proyectos innovadores y en aprender nuevas habilidades, más que en salarios y seguridad laboral. (Haro, 2017).

Es por ello que es importante mantener actualizado el marco de habilidades técnicas, para las trayectorias de progresión profesional deseada, así como las ocupaciones adhoc, de acuerdo a los requerimientos de nuevos puestos laborales para las nuevas tecnologías emergentes que aparecen, que no se posee y que son necesarias adquirirlas para el futuro. Por otro lado, las empresas tienen la oportunidad de brindar capacitaciones y/o cursos de estas habilidades técnicas, basada en la información detallada que actualmente existe en el

mercado laboral como en las descripciones de los puestos de trabajo y que necesitan ser impartidas o potenciadas.

De acuerdo con el informe a nivel global reveló que un 14% de los profesionales busca oportunidades laborales para cambiar su vida y que este porcentaje está creciendo 2 puntos por año aproximadamente (Randstand Workmonitor ,2016). Otro estudio indica que las aplicaciones o conocimientos tecnológicos son las herramientas predilectas para las personas que plantean un cambio drástico en su ámbito profesional. (Manpower Group Solutions ,2018).

Si bien es cierto, actualmente existen diversas plataformas digitales para la búsqueda de trabajo como LinkedIn o Infojobs que muestran información sobre los puestos de trabajo, existen estudios que demuestran algunas necesidades no cubiertas por estas herramientas, como la centralización de datos y visión del perfil laboral basada en habilidades con las que cuentan las personas. Otro estudio, muestra que el 42% de los candidatos a nivel global discrepan sobre el propósito de las publicaciones que hacen las empresas, ya que los mismos se sienten dentro de un proceso que no va más allá de la búsqueda de trabajo, y no los ayuda en una búsqueda de identidad y que vaya de acuerdo a sus propósitos. (Manpower Group Solutions ,2018)

Zumaeta (2018) comentó que los profesionales que están en la búsqueda de alguna oferta laboral deben estar informados sobre el comportamiento del mercado laboral para que de esa manera puedan encontrar mejores opciones de capacitación y empleabilidad. También resalta que las carreras con más demanda de empleos, serán: Ingeniería, Servicios, Comercio, Finanzas, y Tecnología, dada la transformación digital que actualmente se vive a nivel mundial. Por ende, el cambio tecnológico está perturbando al mercado laboral, la demanda de personas con habilidades técnicas está creciendo, y en muchos casos superando a la oferta. Dado que hay una escasez, es crucial desarrollar una estrategia para seguir atendiendo la demanda masiva de habilidades técnicas actualizadas.

Según La Carpio(2019) , especialista legal del Consejo Nacional de Ciencia, Tecnología e Innovación tecnológica (Concytec), un total de 41 empresas peruanas bajo el esquema de incentivos tributarios, han destinado 117 millones de soles para realizar proyectos de investigación, desarrollo tecnológico e innovación que se orientan principalmente a los sectores de tecnologías de la información, de servicios, aprobados por la institución Concytec, lo que nos da una visibilidad de la demanda de personas en el rubro digital y tecnológico.

Así mismo Dávila (2019) viceministro Mype e Industria, comenta que en esta época es trascendental elaborar estrategias para el desarrollo de capital humano especializado (talento) en innovación, absorción tecnológica y digitalización, a partir de un diagnóstico de los perfiles competentes. Como se muestra en el estudio global un promedio de 80% de personas en el mundo desean adquirir habilidades técnicas para garantizar su información en el futuro y el 58% de personas piensan que requieren más habilidades técnicas que difieren de lo que actualmente poseen. (Randstad Workmonitor ,2018).

Figura 1.1

Porcentaje de personas que desean adquirir habilidades digitales para asegurar su empleabilidad

80% want to acquire more digital skills to guarantee their employability in the future.

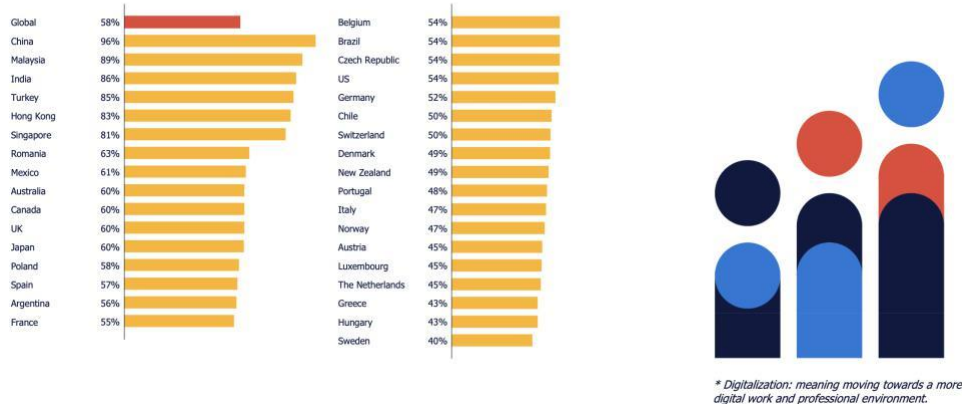


Nota: Son un total de 400 entrevistas por país, de “Global report randstad workmonitor Q4”, por Randstad Digitalization in Work, 2018

Figura 1.2

Porcentaje de personas que piensan que necesitan adquirir diferentes habilidades digitales de las que actualmente poseen (%)

58% think digitalization* requires a different skill set than what they currently possess.



Nota: Son un total de 400 entrevistas por país, de “Global report randstad workmonitor Q4”, por Randstad Digitalization in Work, 2018

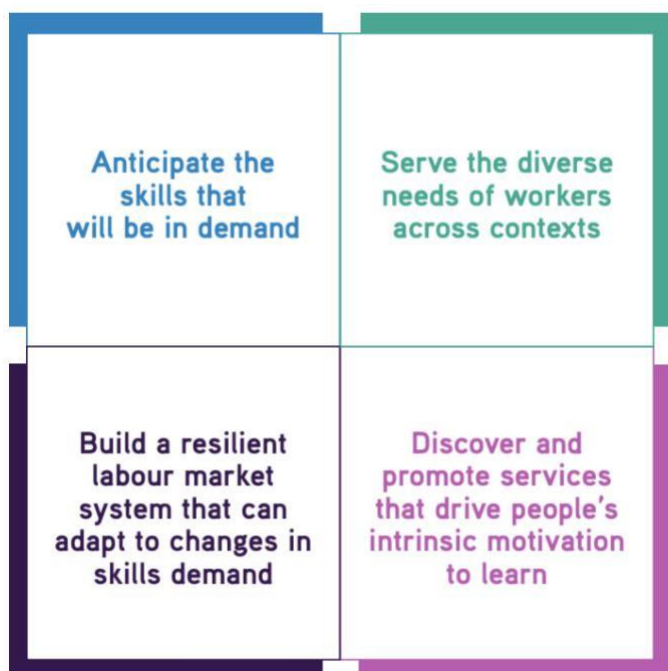
En la figura 1.1 y 1.2 se reconoce la necesidad de la existencia de herramientas que ayuden a poder visualizar qué habilidades técnicas son o serán valiosas para el interés de cada persona. Por lo tanto, es una prioridad, fomentar mercados laborales adaptables en los que todos puedan participar de la mano de la centralización y agrupamiento de datos.

Mecanismos tradicionales como las encuestas, reportes genéricos o planos, no nos dan una visión rápida y confiable del mercado laboral actual, por eso existe la necesidad de tener nuevas formas de acceder a la oferta laboral para así anticipar o ir a la par de los cambios en la demanda y la oferta de habilidades de manera más conveniente.

Así mismo se identificó cuatro desafíos clave que se deberían priorizar para fomentar un ecosistema laboral que sea adecuado para el futuro y poder responder a los cambios en la demanda de habilidades. (Digital Frontrunners,2018).

Figura 1.3

Desafíos para diseñar política de habilidades inclusivas para la era digital



Nota: De “*Report Designing inclusive skills policy for the digital age*”, por Orlik, Jack; Casasbuenas, Juan; Helkkula, Karoliina , 2018 (<https://www.bl.uk/collection-items/digital-frontrunners-designing-inclusive-skills-policy-for-the-digital-age>)

En la figura 1.3, uno de los cuatro desafíos, es construir un resistente sistema de mercado laboral que puede adaptarse a los cambios en demanda de habilidades técnicas y la promoción de servicios que impulsan a las personas a aprender.

Djumalievá, Lima y Sleeman (2018), proponen una metodología para clasificar las ocupaciones en función de las habilidades requeridas que son proporcionadas en los anuncios de trabajo en línea. Dicha clasificación ocupacional consta en agrupar trabajos en ocupaciones basadas en el requerimiento de habilidades similares.

La metodología propuesta para desarrollar esta clasificación ocupacional utiliza métodos de procesamiento del lenguaje natural, como la agrupación de palabras distribuidas similares, a los anuncios de empleo en línea del Reino Unido. La clasificación ocupacional se alinea directamente con las necesidades del empleador y agrupa los trabajos en ocupaciones

basadas en requisitos de habilidades similares. Dicha metodología basada en habilidades prioriza la especialización de habilidades sobre el nivel de habilidades.

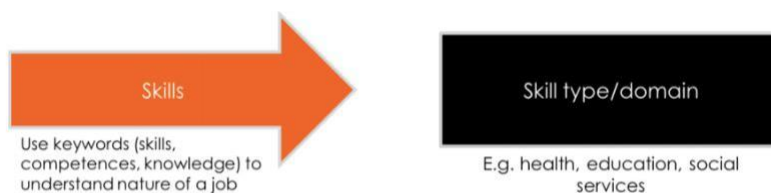
Se realiza un análisis utilizando anuncios de trabajo en línea proporcionados por Burning Glass Technologies, portal laboral, donde los requisitos de habilidades, experiencia y calificación se extraen de las ofertas de trabajo y se estandarizan.

Además de los anuncios de empleo, también se utiliza recursos disponibles públicamente: el Índice de la Oficina de Estadísticas Nacionales(ONS) del 2010 que proporciona una lista de referencia de títulos de trabajo conocidos y un conjunto de términos utilizados para describir las ocupaciones y el Diccionario Europeo de Habilidades y Competencias (DISCO) que divide las habilidades en 9 categorías no específicas de dominio y 25 categorías específicas de dominio, para finalmente para asignar ofertas de trabajo a categorías de habilidades relevantes.

Como se visualiza en la figura 1.4 hacen uso de las habilidades mencionadas en un anuncio de trabajo para comprender la naturaleza del trabajo (especialización de habilidades).

Figura 1.4

Uso de datos de puestos de trabajo Burning Glass para inferir la especialización y el nivel de habilidad



Nota. De “*Classifying Occupations According to Their Skill Requirements in Job Advertisements*”, por Economic Statistics Centre of Excellence (ESCoE) , 2018 (<https://ideas.repec.org/p/nsr/escoed/escoe-dp-2018-04.html>)

Para la metodología de clasificación, se aplica técnicas de aprendizaje automático semi-supervisadas para un conjunto de datos de 37 millones de anuncios de trabajo en línea recopilados en el Reino Unido. La clasificación propuesta tiene el potencial de permitir la medición de la progresión profesional de un individuo dentro del mismo dominio de habilidades, así como recomendar trabajos a los individuos en función de sus habilidades

Según Carnevale, Jayasundera, y Repnikov (2014) algunas estimaciones indican que actualmente el 70% de las ofertas de trabajo se publican en línea. Una de las ventajas que se mencionan de utilizar anuncios de trabajo en línea para analizar las demandas de habilidades, son características como los campos de texto libre en los anuncios de trabajo, que permiten a los empleadores expresar directamente sus necesidades, así como incluir descripciones específicas de habilidades, calificaciones y credenciales (p. 2).

Por otro lado, como se menciona hay algunas desventajas como problemas de representatividad y calidad de los datos, sin embargo, las ventajas de los anuncios de trabajo en línea lo convierten en una fuente útil de información sobre la demanda del mercado laboral (p. 17).

Los estudios también demuestran cómo los requisitos de habilidades en los datos en línea pueden ayudar a refinar las estadísticas económicas, así como Deming y Kahn (2017) establecieron un vínculo positivo entre los requisitos para las habilidades sociales y cognitivas mencionadas en los anuncios y las diferencias salariales (p. 4)

La propuesta, es centralizar a los datos de requerimientos de los puestos de trabajo de manera masiva y mediante la recolección automática de estos, mostrar las habilidades técnicas requeridas por la mayoría de empresas y en una comparación con las habilidades de cada persona , mostrarle dónde puede adquirir las que aún no tiene , generando perspectivas de crecimiento en cada ocupación.

CAPITULO II: CONCEPTOS Y FUNDAMENTOS

Se detalla el marco teórico utilizado en el proceso de innovación del proyecto.

2.1 Web Scraping

Mitchell (2015) define web Scraping como el proceso de extraer y crear representación estructurada de datos de un sitio web mediante técnicas automatizadas. Es la práctica de recopilar datos por cualquier medio que no sea un programa que interactúa con una API (o, como obviamente, a través de un humano usando un navegador web). Esto se logra consultando a un servidor web, luego solicitar datos (generalmente en forma de HTML), y luego analizar estos datos para extraer la información necesaria.

Al principio los datos parecen poco estructurados, la data está dispersa. Es necesario identificar cuál es el patrón que siguen los datos, para poder crear y ejecutar una lógica de extracción y procesamiento de los mismos. (Mitchell ,2015).

El web scraping es una solución intermedia entre técnicas como la recolección manual de datos y el acceso automatizado a los mismos basado en un protocolo predeterminado como una interfaz de programación de aplicaciones (API), en su mayoría se usa cuando la cantidad de datos que se desea extraer tiene un volumen muy grande y los datos son masivos. Mooney (2015) indica que el web scraping es reconocido por ser una técnica eficiente para recolectar big data.

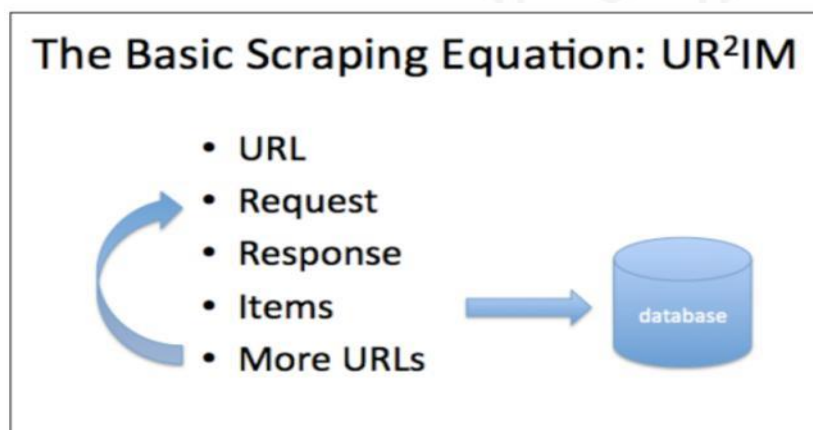
El procedimiento para hacer web scraping no es único, depende del objetivo que se quiera lograr, por ejemplo, Zhao (2018) comenta que se puede dividir en dos pasos secuenciales; donde el primero consta de adquirir recursos web y luego extraer la información deseada de los datos adquiridos. Por otro lado, López (2018) sugiere revisar el código fuente de las páginas web de donde se desea extraer los datos, para encontrar la forma o patrones en que están organizadas y sobre esa base definir rutas e identificadores que permitan descargar las páginas y extraer los datos requeridos para procesarlos.

Específicamente, se comienza formando una solicitud HTTP para adquirir recursos del sitio web específico. Esta solicitud puede formatearse en una URL que contenga una consulta GET o una consulta POST. Una vez que el sitio web objetivo recibe y procesa con éxito la solicitud, el recurso solicitado se recupera del sitio web. El recurso puede estar en múltiples formatos, como páginas web que son construidas a partir de HTML, fuentes de datos en XML o en formato JSON o datos multimedia como imágenes, audio, o archivos de video, etc. Después de descargar los datos web, el proceso de extracción continúa analizando, formatea y organiza los datos la estructura que uno haya establecido.

Una de las librerías más usadas para hacer web scraping, analizar y extraer información del código HTML sin procesar, es BeautifulSoup, esta proporciona funciones basadas en python para navegar, buscar y modificar con ayuda de un grupo de herramientas para descomponer los archivos HTML y extraer la información deseada con ayuda de otras librerías de manera más fácil y rápida.

Figura 2.1

La ecuación básica de scraping



De "Basic Crawling", por Dimitrios Kouzis-Loukas, *Learning Scraping* (pp. 31-32), 2016, Packt Publishing.

Por otro lado, la técnica de web scraping ha sido cuestionada por su validez legal en torno a obtener datos de otras fuentes web que no sean las propias. Sin embargo, el pasado 2017, hiQ interpuso una demanda a LinkedIn para alegar que el scraping que realizaba era legal, el fallo se dio a su favor donde, se le dijo a LinkedIn que era permitida la adquisición de cualquier data que no necesite iniciar sesión en el servicio.

En la apelación de LinkedIn sobre hiQ, que se dió el día 9 de septiembre del 2019 en la corte de apelaciones de Estados Unidos (ver anexo 1). El tribunal de distrito, *Bernhardt v. Los Angeles* (2019) señala “cada parte del tribunal afirma que su posición beneficiaría al interés público al maximizar el flujo libre de información en Internet”. (p. 35). HiQ señala que el raspado de datos es un método común de recopilación de información, utilizado por motores de búsqueda, investigadores académicos y muchos otros.

LinkedIn es una de las redes sociales en la que el web scraping se aplica de manera más frecuente para extraer la información de una gran cantidad de perfiles para guardar datos de candidatos según los criterios que se consideren relevantes como experiencia, estudios, habilidades, etc. (Gestión, 2018). La corte falló a favor de HiQ considerando un precedente bajo el sustento de que se podía “raspar” cualquier información que apareciese en una página web pública sin protección de contraseña.

2.2 Procesamiento del Lenguaje Natural

Hernández (2018) define el procesamiento de lenguaje natural como una disciplina de la Computación que concierne al tratamiento de información desestructurada. Los datos pasan por etapas de procesamiento, una de ellas es utilizar palabras embebidas, una técnica de procesamiento de lenguaje natural, que captura similitudes semánticas de términos basados en su distribución en grandes campos de texto. Si bien es cierto, existen diferentes enfoques de inserción de palabras para asignar palabras a su representación distribuida, la salida resultante es típicamente un vector numérico con longitud 300, donde las dimensiones representan implícitos conceptos semánticos como lo mencionan Chen, Corrado, Dean, Mikolov y Sutskever (2013).

Una de las tareas principales de procesamiento de lenguaje natural es la clasificación de textos, el cual es el proceso de asignar etiquetas o categorías al texto de acuerdo con su contenido. Extraer un texto puede ser complicado por la data no estructurada, la clasificación de texto con aprendizaje automático aprende a hacer clasificaciones basadas en observaciones pasadas, estos clasificadores son más fáciles de mantener y siempre se puede etiquetar nuevos ejemplos.

Entre los algunos algoritmos de clasificación de texto más populares se tiene:

- **Naive Bayes:** Es una familia de algoritmos estadísticos que podemos utilizar al hacer la clasificación de texto, se basa en el Teorema de Bayes. Una de sus principales ventajas es que puede obtener resultados realmente buenos cuando los datos disponibles no son muchos. Naive Bayes trabaja de manera que cualquier vector que represente un texto tendrá que contener información sobre las probabilidades de aparición de las palabras del texto dentro de los textos de una categoría determinada para que el algoritmo pueda calcular la probabilidad de que ese texto pertenezca a la categoría.
- **Clasificador Lineal:** Aquí se utiliza la regresión logística que mide la relación entre la variable dependiente categórica y una o más variables independientes mediante la estimación de probabilidades utilizando una función logística / sigmoidea.
- **Máquina de vector soporte:** es un algoritmo supervisado de aprendizaje automático que se puede utilizar para desafíos de clasificación o regresión.
- **Modelo Bag Words:** Implementando un modelo Random Forest, que son un tipo de modelos de conjunto. Forman parte de la familia modelo basada en árboles. El modelo de bolsa de palabras es fácil de entender e implementar y ha tenido un gran éxito en problemas como el modelado de idiomas y la clasificación de documentos. Brownlee (1917) comenta que el enfoque es muy simple y flexible, y puede usarse de muchas maneras para extraer características de los documentos.

Básicamente el modelo depende de una "bolsa de palabras", que es una representación de texto que describe la aparición de palabras dentro de un documento que implica, un vocabulario de palabras conocidas, una medida de la presencia de palabras conocidas.

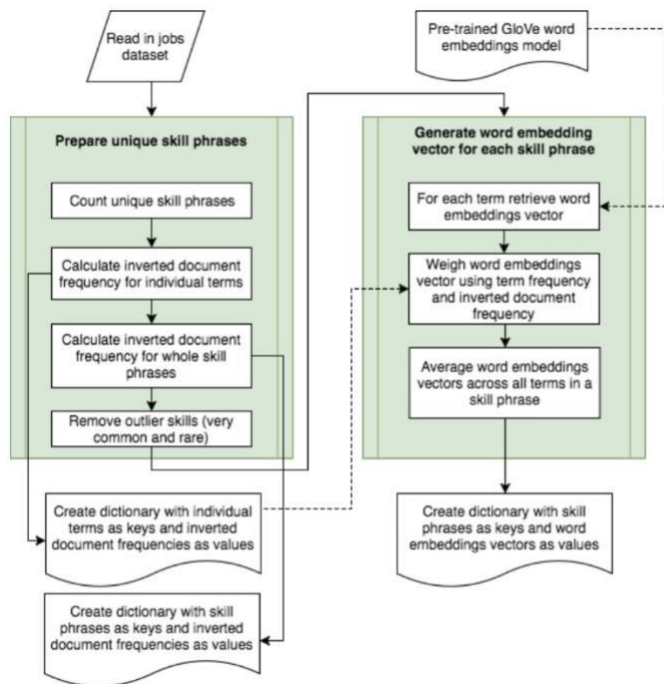
Es importante recalcar que el orden o la estructura de las palabras en el documento se descarta. El modelo solo tiene que ver con si las palabras conocidas aparecen en el documento, no en qué parte del documento. La bolsa de palabras puede ser tan simple o compleja como se desee, la complejidad se presenta al diseñar el vocabulario de palabras conocidas (tokens) como el calificar la presencia de palabras conocidas.

Las palabras embebidas son más flexibles que otras técnicas como la de bolsa de palabras (bag-of-words), estas representan documentos como conjuntos múltiples de palabras ignorando el orden de las palabras. El uso de palabras embebidas permite comparar similitudes de documentos (es decir, cosas como anuncios de trabajo, descripciones de habilidades) que contienen términos, que son semánticamente similares, pero no exactamente lo mismo.

Hay modelos de palabras embebidas pre-entrenados disponibles públicamente. Para este caso se utiliza el modelo llamado GloVe, que contiene un vocabulario de 2.2 millones de palabras, el procedimiento se visualiza en la figura 2.2.

Figura 2.2

Generaciones vectores de incrustación de palabras para frases de habilidad



De “Classifying Occupations According to Their Skill Requirements in Job Advertisements”, por Jyldyz Djumalieva & Antonio Lima & Cath Sleeman, 2018 (<https://www.escoe.ac.uk/wp-content/uploads/2018/03/ESCoE-DP-2018-04.pdf>)

Deming y Kahn (2017) estudian la variación en la demanda de habilidades para profesionales en todas las empresas y mercados laborales, donde se clasifica una amplia gama de palabras clave encontradas en anuncios de trabajo en 10 habilidades generales previamente segmentadas como se observa en la tabla 2.1 donde codifican un anuncio que tienen un requisito particular de habilidad laboral, para ver si tiene al menos una de las palabras o frases clave enumeradas y para estudiar la heterogeneidad en las demandas de habilidades de los empleadores. Se utilizaron palabras y frases clave del texto real de las vacantes laborales para finalmente mostrar la prevalencia de anuncios de habilidades laborales de alto nivel.

Tabla 2.1*Tabla de descripción de habilidades de empleos.*

Job Skills	Key words and phrases
Cognitive	Problem Solving, Research, Analytical, Critical Thinking, Math, Statistics
Social	Communication, Teamwork, Collaboration, Negotiation, Presentation
Character	Organized, Detail-oriented, Multi-tasking, Time Management, Meeting Deadlines, Energetic
Writing	Writing
Customer Service	Customer, Sales, Client, Patient
Project Management	Project Management
People Management	Supervisory, Leadership, Management (not project), Mentoring, Staff
Financial	Budgeting, Accounting, Finance, Cost
Computer (general)	Computer, Spreadsheets, Common Software (e.g. Microsoft Excel, Powerpoint)
Software (specific)	Programming language or specialized software (e.g. Java, SQL, Python, etc.)

Notes: Authors categorization of open text fields in Burning Glass data.

Nota. De “Skill Requirements across Firms and Labor Markets: Evidence from Job Postings for Professionals”, por David Deming, Lisa B. Kahn, 2018, *Journal of Labor Economists*, (<https://doi.org/10.1086/694106>)

2.2.1 Coincidencia aproximada de cadenas

Técnica también conocida como fuzzy string searching, la cual encuentra cadenas que coincidan aproximadamente con un patrón. El problema de la coincidencia aproximada de cadenas generalmente se divide en dos etapas, las cuáles son: encontrar coincidencias aproximadas de subcadenas dentro de una cadena dada y encontrar cadenas de diccionario que coincidan aproximadamente con el patrón.

Hay diversos algoritmos de similitud de cadenas, uno de los más comunes es el de edición basada en la distancia: los algoritmos que se incluyen en esta categoría intentan calcular el número de operaciones necesarias para transformar una cadena en otra, donde mientras más sea el número de operaciones, menos es la similitud entre las dos cadenas.

Baeza-Yates y Navarro (1999), plantean que dado un largo texto de longitud n , y un patrón de longitud m , se recupera todos los segmentos u ocurrencias del texto cuya distancia de edición al patrón es como máximo k . La distancia de edición $ed()$ entre dos cadenas está definida como el número mínimo de inserciones de caracteres, eliminaciones y reemplazos necesarios para igualarlos (p. 1).

Li (2018) comenta que la coincidencia de cadenas similares es un tipo de búsqueda que encontrará coincidencias incluso cuando los usuarios escriban mal las palabras o ingresen solo palabras parciales para la búsqueda. También se conoce como coincidencia aproximada de cadenas.

2.2.1.1 Distancia de Levenshtein

Esta distancia se calcula al encontrar el número de ediciones que transformarían una cadena en otra. Las transformaciones permitidas son de inserción: agregar un nuevo carácter; eliminación: eliminar un carácter y sustitución: reemplazar un carácter por otro.

Al realizar estas tres operaciones, el algoritmo intenta modificar la primera cadena para que coincida con la segunda. Así al final se obtiene una distancia de edición.

Junior (2018) explica que la idea principal del algoritmo de Levenshtein de programación dinámica es llenar las entradas de una matriz m , cuyas dos dimensiones son iguales a las longitudes de las dos cadenas cuya distancia de edición se está calculando.

2.2.1.2 Algoritmo Jaro-Winkler

En informática y estadística, la distancia Jaro-Winkler es una métrica de cadenas para medir la distancia de edición entre dos secuencias. La distancia Jaro-Winkler utiliza una escala de prefijo que otorga calificaciones más favorables a las cadenas que coinciden desde el principio para una longitud de prefijo establecida. La distancia de Jaro entre dos secuencias s_1 y s_2 es definida por:

Figura 2.3

Fórmula distancia de Jaro Winkler

$$d_j = \frac{1}{3} \left(\frac{m}{|s_1|} + \frac{m}{|s_2|} + \frac{m-t}{m} \right)$$

Jaro distance formula

Nota. De “String Similarity Algorithms Compared”, por Coralie Collignon, 2018 (<https://medium.com/@appaloosastore/string-similarity-algorithms-compared-3f7b4d12f0ff>)

Tabla 2.2

Explicación de las variables de fórmula de Jaro Winkler

Variable	Significado
dj	Distancia de Jaro
m	Número de caracteres coincidentes (caracteres que aparecen en s1 y en s2)
t	Mitad del número de transposiciones (compare el carácter i-ésimo de s1 y el carácter i-ésimo de s2 dividido por 2)
s1	Longitud de la primera cadena
s2	Longitud de la segunda cadena

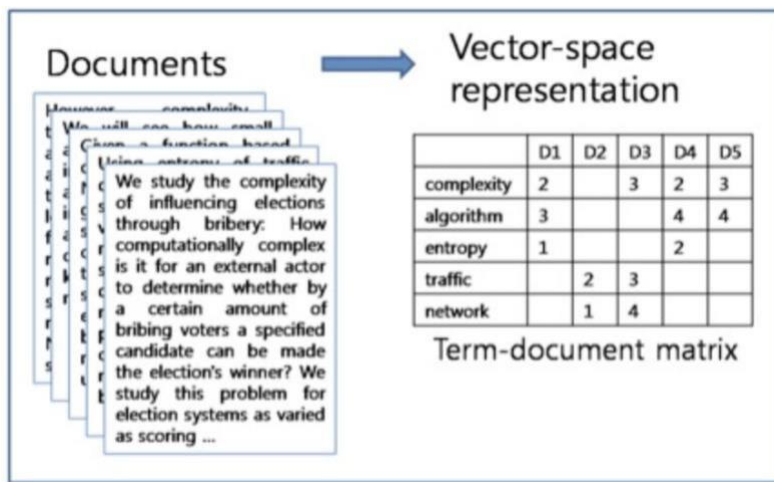
Nota. Adaptado de “String Similarity Algorithms Compared”, por Coralie Collignon, 2018 (<https://medium.com/@appaloosastore/string-similarity-algorithms-compared-3f7b4d12f0ff>)

Ilyankou (2014), explica que el puntaje de similitud entre dos cadenas que da el algoritmo es un valor de 0 a 1, donde 0 corresponde a ninguna similitud y 1 a una coincidencia completa.

Adicionalmente Hughes (2015), explica que hay un área de investigación en la ciencia de procesamiento de lenguaje natural, llamada “semántica distributiva”, que ayuda en el proceso de determinar la relación de 2 palabras y que intenta determinar el significado de una palabra examinando los contextos en los que aparece.

Hughes (2015) también comenta el proceso por el que DICE (plataforma web de trabajo y rutas de habilidades) pasa para poder mostrar información relevante de la relación de habilidades a partir de la extracción de habilidades tecnológicas de las ofertas de trabajo. En su proceso se construye una "Matriz de términos de documentos" a partir de una lista de habilidades, también llamada "Representación del espacio vectorial", esta matriz presenta una fila por habilidad, junto con una columna para cada documento (o publicación de trabajo en este caso), como se muestra en la imagen 2.4.

Figura 2.4
Representación vectorial en matriz de habilidades

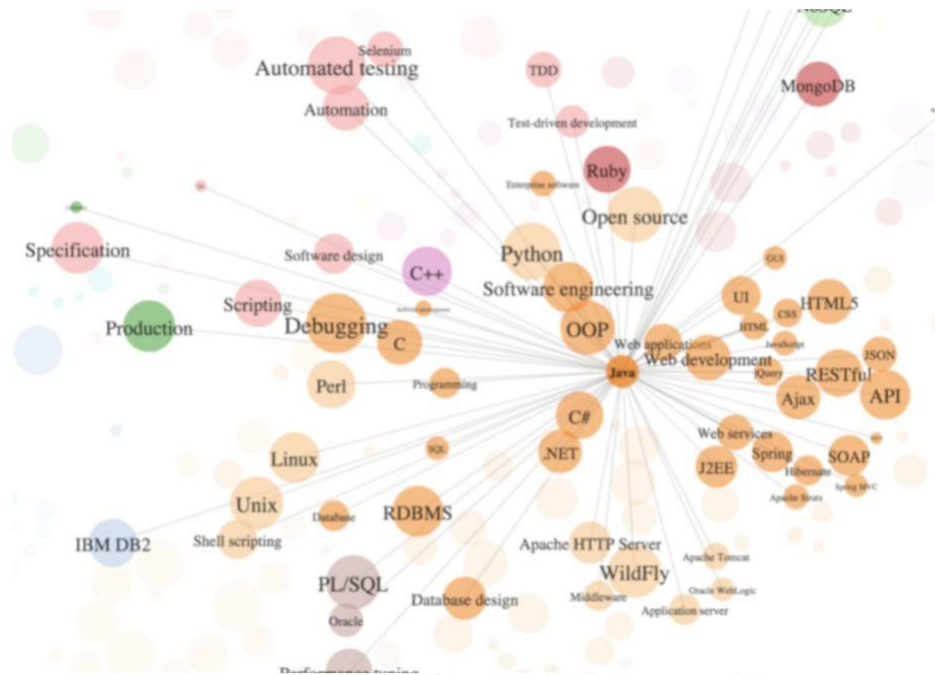


Nota. De "How We Data-Mine Related Tech Skills", por Simon Hughes, 2015

En este formato, cada habilidad se representa como un vector (una lista de números) que codifica los documentos en los que se produce, así mismo propone un enfoque utilizando análisis semántico latente con una suposición de que las habilidades están relacionadas con habilidades que ocurren en el mismo contexto, donde los contextos son puestos de trabajo.

Una vez que se tienen los vectores para cada habilidad, se usa la métrica de similitud de coseno para calcular qué tan similares son dos habilidades de sus vectores, y así calcular las habilidades más similares o relacionadas para una habilidad dada. Como finalmente se plasma en el ejemplo de la figura 2.5

Figura 2.5
Representación de relación de skills



Nota. De “How We Data-Mine Related Tech Skills” , por Simon Hughes, 2015, (<https://insights.dice.com/2015/03/16/how-we-data-mine-related-tech-skills>)

2.3 Trabajadores del conocimiento

Según Falco (2003), un trabajador del conocimiento es aquel que posee un saber que le sirve para hacer cosas y lo utiliza para trabajar. Así mismo Drucker (1999) el trabajador del conocimiento es y continuará siendo el activo más valioso de cualquier organización, el cual introdujo el concepto de trabajador del conocimiento en los años 70 como personas que no hacen tareas solo rutinarias y, por lo tanto, su aportación depende de su capacidad pensar y aprender nuevos conocimientos.

A diferencia del trabajador manual, el trabajador del conocimiento es una persona que se identifica con su área de especialización y no con su empleador. La organización significa para un espacio donde aplicar sus conocimientos. Su compromiso no se consigue a través del salario sino a través de la oferta de oportunidades de desarrollo.

Una de las principales características de ser un trabajador del conocimiento implica auto gestionarse, es decir reconocer sus puntos fuertes y débiles o los que les falta; considerar aprender conocimientos nuevos; asumir la responsabilidad por su desarrollo y crecimiento. (Falco, 2003)

Por lo que es importante incorporar una estrategia de ascenso para mejorar continuamente el conocimiento y las habilidades de cada persona en la fuerza laboral y garantizar oportunidades para su desarrollo profesional y aumentar los ingresos. (Departamento de Salud, Educación y Bienestar del Instituto Nacional de Educación, “Resources in Education” ,2010).

2.4 Modelo de Negocio Freemium

Pahwa (2019), comenta que menos del 20% de los usuarios de LinkedIn pagan por los servicios premium proporcionados por la empresa. Sin embargo, la compañía obtuvo aproximadamente \$ 1 mil millones en ingresos en el primer trimestre de 2017.

Freemium es un modelo de negocio basado en Internet donde los servicios básicos se proporcionan de forma gratuita, pero se cobran cargos por funciones premium adicionales, es un acrónimo de gratuito y premium y la estrategia es bastante popular. Incluso los líderes del mercado como LinkedIn, Tinder, Youtube, Candy Crush, etc. utilizan un modelo freemium para aumentar su base de usuarios y generar más ingresos mediante la implementación de estrategias de micropagos.

El modelo de negocio Freemium fue creado por la industria del software en la década de los ochenta como una estrategia de tiempo limitado o de características limitadas, el modelo de negocios tiene como objetivo atraer a tantos clientes potenciales como sea posible para probar el producto de forma gratuita y convertirse en miembros premium después de pagar una cierta tarifa de suscripción.

Dicho modelo trabaja con el efecto de red, es decir distribuir el servicio de forma gratuita es la mejor manera de obtener más clientes y, aunque la mayoría de estos clientes no alcanzan el nivel premium, actúan como un imán para atraer a más clientes premium potenciales. La oferta de freemium se compone de una propuesta de valor básica regalada de forma gratuita y una prima oferta de pago con características elaboradas (Duarte y Gailly ,2016)

2.5 Metodologías Ágiles

Rosello (2019) comenta que las metodologías ágiles son aquellas que permiten adecuar la forma de trabajo, consiguiendo flexibilidad e inmediatez en la respuesta para amoldar el proyecto y su desarrollo a las circunstancias específicas del entorno. Las cuales se originaron como reacción a las metodologías existentes con el propósito de disminuir la burocracia que implica el desarrollo de las metodologías tradicionales que tienen muchas actividades que hacen que se retrase la etapa de desarrollo. (Herrera y Quiñonez-Xu ,2019).

Existen diferentes tipos de metodologías ágiles de desarrollo más exitosas como:

- Scrum: En scrum un proyecto se ejecuta en iteraciones (llamados sprints). Cada iteración tiene que proporcionar un entregable completo, un incremento de producto que sea susceptible de ser entregado con el mínimo esfuerzo cuando el cliente lo solicite.

- XP (eXtreme Programming): Promueve que todo el código sea escrito en parejas trabajando en el mismo ordenador. La programación en parejas incrementa la calidad del código sin impactar en la fecha de entrega.

- Kanban: Su objetivo es gestionar de una manera más general como se van completando tareas, entre sus principales reglas están, visualizar el trabajo y las fases del flujo de trabajo y medir el tiempo de completar una tarea.

2.6 Design Thinking

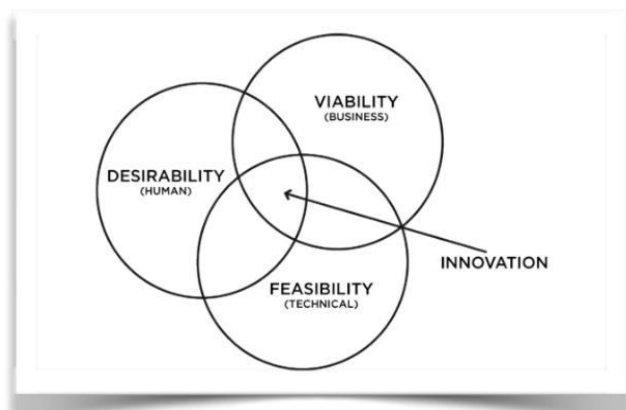
La nomenclatura de "design thinking" aparece por primera vez de manera prominente en un libro de ese título escrito por Peter Rowe (1987). En su uso actual, como proceso de pensamiento, la nomenclatura se atribuye a la consultora de innovación IDEO (Kelley y Littman, 2005).

El ex presidente de la Design Management Institute, una asociación líder de los profesionales del diseño que trabajan en negocios ha ofrecido una definición más detallada del pensamiento de diseño: "un proceso de innovación, innovador que enfatiza la observación, colaboración, aprendizaje rápido, visualización de ideas, rápido. Concepto de prototipado y análisis empresarial concurrente". (Lockwood, 2009).

Adicionalmente, Brown (2017) plasma que el pensamiento de diseño se centra en la definición de innovación, que viene a ser la unión de las 3 fuerzas, deseabilidad, factibilidad y viabilidad.

Figura 2.6

Definición de una satisfactoria innovación: Intersección de 3 fuerzas:



Nota. De "De Edison a Tim Brown: algunas ideas sobre Design Thinking" , por Pablo Mondragón Valero, 2018 (<https://blog.antropologia2-0.com/es/edison-tim-brown-design-thinking/>)

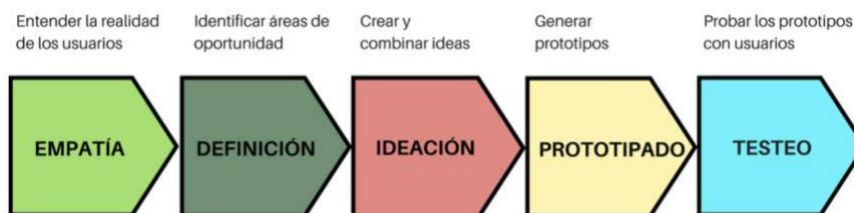
Design Thinking es una metodología que impregna todo el espectro de actividades de innovación con un espíritu de diseño centrado en el hombre. Su esencia es conocer lo que las personas quieren y necesitan en sus vidas.

El enfoque, Design Thinking no sólo se centran en la creación de productos y servicios, se basa en la capacidad para: ser intuitivo, reconocer patrones, construir ideas con significado emocional, funcional, y expresarse en los medios de comunicación que no sean palabras o símbolos. En por ello que desde años atrás Buchanan (1992) afirma:

De hecho, los signos, las cosas, las acciones y los pensamientos no solo están interconectados, sino que también el interpenetrarse y fusionarse en el pensamiento de diseño contemporáneo tiene sorprendentes consecuencias para la innovación. Estas áreas sugieren el linaje del pasado del diseño y presente, así como señalar hacia dónde se dirige el diseño en un futuro (p. 9).

A continuación, se muestra las 5 etapas o pasos fundamentales del Design Thinking (Castillo V, 2014), como se puede observar en la figura 2.7

Figura 2.7
Etapas del Design Thinking



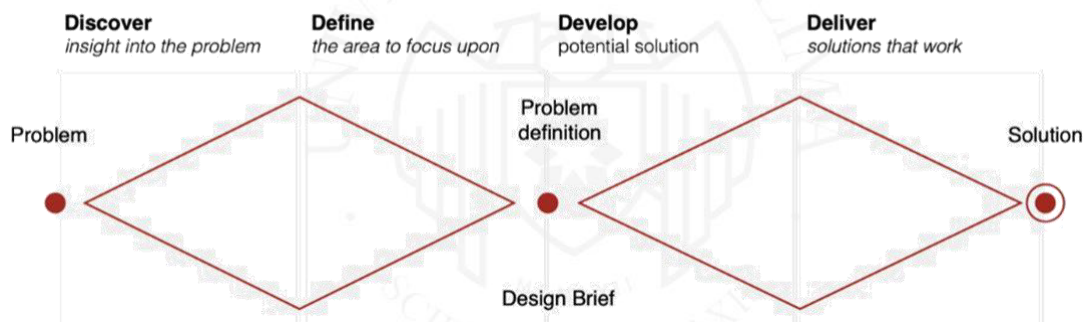
Nota. De “De Edison a Tim Brown: algunas ideas sobre Design Thinking” , por Pablo Mondragón Valero, 2018 (<https://blog.antropologia2-0.com/es/edison-tim-brown-design-thinking/>)

Mondragón (2018) también explica que el design thinking se basa en la comprensión de cómo viven los consumidores y en una experimentación para lograr crear valor y soluciones transformadoras.

Por otro lado, el doble diamante representa un mapa visual que ayuda a seguir todas estas etapas mencionadas en el proceso de diseño y se divide en cuatro fases: descubrir, definir, desarrollar y entregar. El modelo enfatiza la combinación de pensamiento divergente y convergente para diferentes etapas del proceso creativo. El gráfico de la figura 2.8 lo representa, donde se plasma que para el proceso creativo es necesario un pensamiento convergente para dos propósitos: primero, confirmar la definición del problema y segundo definir la solución y un pensamiento divergente para desarrollar y entregar el producto.

Figura 2.8

Marco del modelo diamante para el proceso del design thinking



Nota. De “What is the framework for innovation? Design Council's evolved Double Diamond” , por Design Council, 2018 (<https://www.designcouncil.org.uk>)

CAPITULO III : FUNDAMENTACIÓN DEL PROYECTO

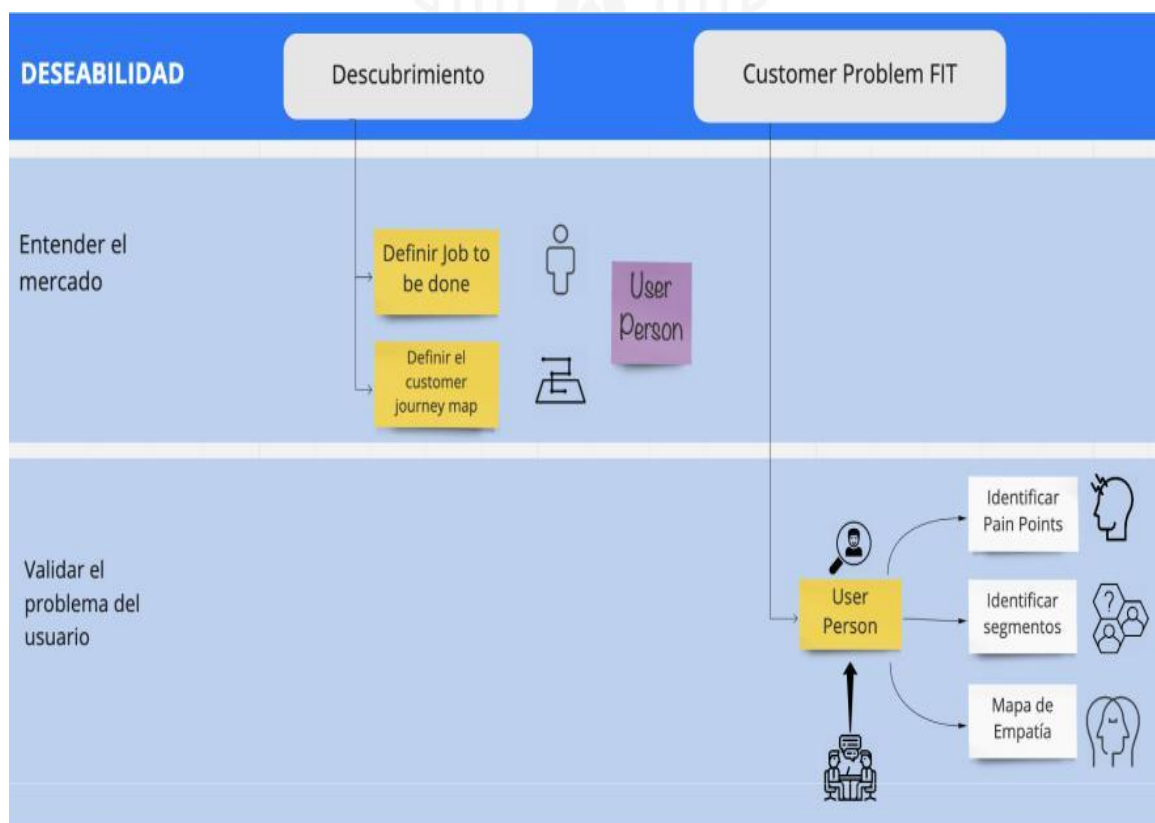
3.1 Fundamentación de la deseabilidad del proyecto

El desarrollo de este proyecto está centrado en la importancia de evaluar la información recopilada de las ofertas laborales de las bolsas de trabajo del sector de tecnología, para construir una clara comprensión de las necesidades que se requieren en la actualidad y que encajan con nuestro perfil.

Como explica Jennersjö (2018) se debe saber que el futuro laboral tiene que ver con la compatibilidad de competencias en lugar de buscar profesiones. Las profesiones vendrán y se irán, pero las competencias se mantendrán. Por lo que debe enfocarse más en qué competencias tiene una persona, cuáles se vinculan entre sí y qué podría hacer una persona con estas competencias si su profesión como tal desaparece.

Figura 3.1

Esquema para fundamentación de la deseabilidad del cliente



3.1.1 Descubrimiento

Para mejorar las habilidades de un trabajador necesita acceso a información relevante y de alta calidad del mercado laboral (Delivering Digital Skills, Nesta, 2018). Es así, que innovar con técnicas para garantizar que esta información satisface la necesidad oportuna, para mapear habilidades relevantes y proporciona información útil sobre las necesidades del entorno laboral.

Actualmente existen diversas plataformas para conseguir empleo de manera digital y conocer los requerimientos de las empresas, pero muchas de ellas sólo prestan su servicio bajo un modelo tradicional donde los solicitantes buscan títulos y sectores de trabajo específicos, pero desconocen las competencias y habilidades que son requeridos en el mercado laboral.

Aproximadamente el 43% de empleadores en el Perú, tienen dificultades para encontrar trabajadores con las competencias necesarias, y esto se debe a que los postulantes suelen carecer de factores como experiencia, conocimientos técnicos y habilidades blandas. (Gestión, 2019)

Así mismo se advierte que la escasez de talento reduce la competitividad de las empresas en todos los sectores productivos, por no contar con profesionales adaptados a un contexto más global y cada vez más cambiante, por lo que la falta de profesionales locales obliga a gastar más en contratación de personal extranjero.

Adicionalmente el desempleo entre los jóvenes es cuatro veces el de los adultos, a su vez sostiene que una de las principales causas de este subempleo es la falta de capacitación, pues las empresas requieren personal calificado, entrenado y especializado que genere productividad. (Germán Lora, 2017)

Por su lado el investigador peruano, Castro (2017) señaló:

No es solo que las universidades o institutos no estén dando las habilidades profesionales adecuadas, hay una falta de aprendizaje valioso que el peruano promedio desconoce, como cuáles son las cosas valoradas para aprender (p. 1).

Una de las causas raíz de esta problemática, es que el mercado laboral ha sufrido cambios impulsado por dos factores: el tema generacional y el avance rápido de la tecnología. Es por ello, que estos factores han hecho que el mercado laboral se torne mucho más complejo e innovador; y por ende haya más competencia y las empresas busquen talentos diferentes, lo que obliga a una continua capacitación para generar perfiles diferenciales que el mercado laboral necesita.

Lizárraga (2018) explica:

Ya no basta con personas puramente administrativa o responsables, ahora se busca gente con soluciones y servicios que lo distinguen de la competencia, las personas deben actualizar sus competencias continuamente de maneras más creativas y eficiente y que vaya de la mano con el rubro tecnológico (p. 2).

Figura 3.2

Encuesta laboral 2018



Nota: Se encuestaron a 18000 personas. De “La revolución de habilidades”, por ManpowerGroup, 2018

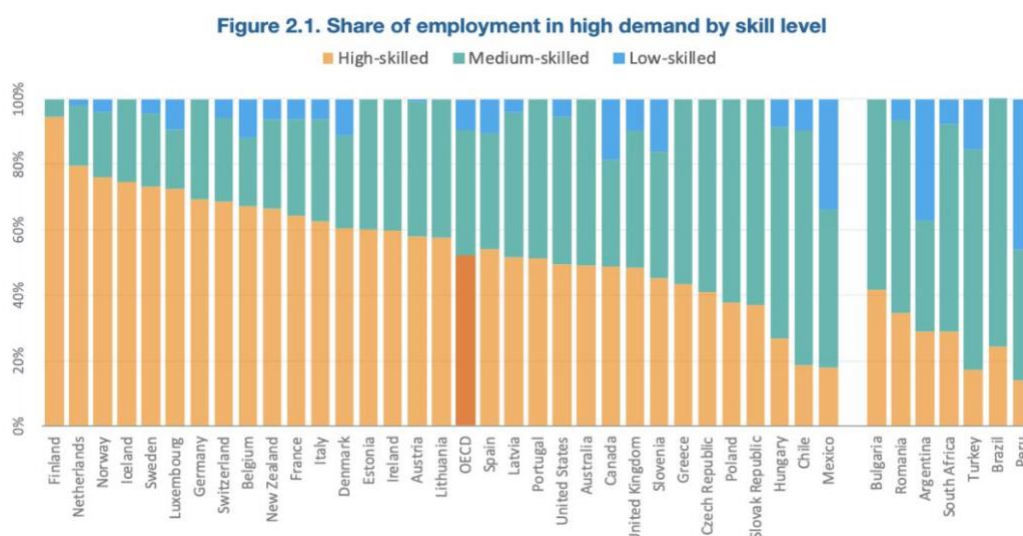
La figura 3.2, muestra la visibilidad de la demanda que tiene el sector tecnología y la importancia de conocer qué habilidades se requieren desarrollar, según Dehaze (2017) la digitalización ha desencadenado una profunda automatización del trabajo, con el impacto de pasar de funciones rutinarias y repetitivas a "trabajadores del conocimiento". Además, comenta que las tecnologías de la información y de la comunicación son un claro ejemplo de demanda superando ampliamente la oferta, se necesita un promedio de cuatro meses para cubrir un puesto de tecnología.

La encuesta citada por la Organización Internacional del Trabajo (OIT), sugiere que el problema no es tanto la falta de aperturas - las tasas de vacantes a menudo pueden ser contra intuitivamente altas - sino la inadecuación de los candidatos para los puestos de trabajo disponibles. Alrededor del 40% de los empleadores dicen no encontrar jóvenes con las habilidades adecuadas para su negocio.

La Organización para la Cooperación y el Desarrollo Económicos (OCDE) propone información sobre su base de datos "Skill for Jobs", el caso de México, Perú o Chile, donde la demanda de profesionales altamente calificados es significativamente menor, aproximadamente 2 de cada 10 trabajos tienen escasez de empleados "altamente calificados" como se puede visualizar en la figura 3.3. (OCDE, 2018)

Figura 3.3

Proporción de empleo en alta demanda por nivel de habilidad



Nota: De "Skills for job", por OECD , 2018

Según el Índice de Competitividad Global de Talento específicamente Perú ocupa el puesto 90 en el pilar "Habilidades Vocacionales y Técnicas" y el puesto 82 en "Habilidades de Conocimiento Global" de los 118 países evaluados a nivel global; cómo podemos observar en la imagen 3.4 , donde se plasma la necesidad de desarrollar estrategias de utilidad para hacer crecer el talento e impulse la competitividad y ayudar a que las personas sepan en que potenciar sus habilidades digitales, actualizarlas y saber combinarlas de forma adecuada con las capacidades que ya manejan. (ADECCO, 2017)

Figura 3.4

Indicadores clave de la Competitividad del Talento Global (GTCI) en Perú 2017

Índice de Competitividad del Talento Global (GTCI) 2017

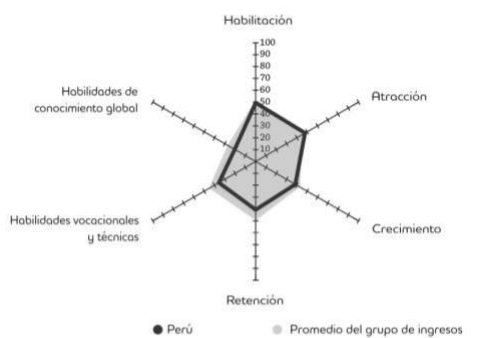
PERÚ

Indicadores clave

Rango (de un total de 18) **83**
 Grupo de ingresos **Ingreso medio-alto**
 Grupo regional.. **América Central, América Latina y el Caribe**
 Población (en millones)..... **31.38**

PBI per cápita (PPA US\$) **12,402.42**
 PBI (en miles de millones de US\$) **192.08**
 Puntuación GTCI **38.76**
 Puntuación GTCI (promedio según grupo de ingresos)..... **42.66**

GTCI 2017 Perfil del país por pilar



Pilar	Puntuación	Ranking
1 HABILITACIÓN	49.40	74
1.1 Panorama normativo	43.26	79
1.2 Panorama de mercado	54.31	49
1.3 Panorama de negocio y entorno laboral	50.65	95

Pilar	Puntuación	Ranking
2 ATRACCIÓN	47.94	55
2.1 Apertura Externa	44.19	38
2.2 Apertura Interna	51.69	89
3 CRECIMIENTO	38.51	76
3.1 Educación Formal	13.57	99
3.2 Aprendizaje permanente	57.25	37
3.3 Acceso a oportunidades de crecimiento	44.70	76
4 RETENCIÓN	40.63	89
4.1 Sostenibilidad	35.13	89
4.2 Estilo de vida	46.14	89
5 HABILIDADES VOCACIONALES Y TÉCNICAS	35.72	90
5.1 Habilidades de Nivel medio	28.87	78
5.2 Empleabilidad	42.56	109
6 HABILIDADES DE CONOCIMIENTO GLOBAL	20.35	82
6.1 Habilidades de Alto nivel	21.74	79
6.2 Impacto del Talento	18.96	81

Nota. De "Índice de competitividad del talento Global", por ADECCO , 2017

Alicia Sánchez, directora de RRHH de Altran (2018), comenta que: "el conocimiento está al alcance de todos y cada día más empaquetado; se transforma y cambia de manera continua y rápida", pero las habilidades y competencias tienen que cuidarse y alimentarse para que las personas las tengan siempre actualizando.

Así mismo, el 60 % de las organizaciones en Perú tendrá articulada su estrategia de transformación digital para el 2020, lo que resultaría en un incremento de la demanda de profesionales de las TI. (IDC, 2018). Así mismo, los profesionales relacionados a las Tecnologías de la Información, son de los más requeridos en los sectores económicos analizados. (Encuesta de Demanda Ocupacional, Ministerio de Trabajo y Promoción del Empleo del Perú, 2018)

Por otro lado TicJob.co muestra información que al 2018, en Estados Unidos, los salarios de algunas ocupaciones de TI superan los 100.000 dólares anuales; en Chile promedia los 46.500 dólares anuales; y en Perú llega a los 45.190 dólares al año, siendo un

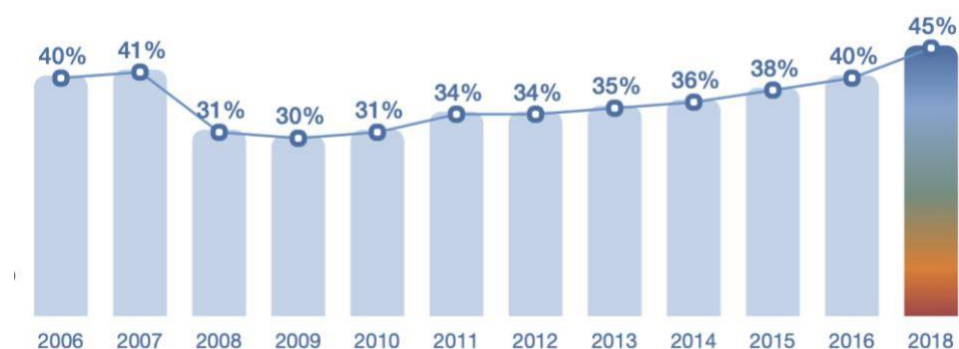
sector que abre gran interés a los profesionales de este sector y es necesaria la actualización de habilidades para mantenerse vigente en el mercado.

El fenómeno de escasez de profesionales se ha incrementado como nunca antes se ha visto en las últimas décadas, de tal manera que los empleadores no logran encontrar a las personas que necesitan con la adecuada combinación de habilidades técnicas y fortalezas humanas, de acuerdo con la reciente Encuesta de Escasez de Talento Perú 2018. (ManpowerGroup, 2018)

Así, el estudio expone que el 54% de la gran empresa (+250 trabajadores) tiene la mayor dificultad para cubrir vacantes. Incluso un 5% más que el año pasado.

Figura 3.5

Incremento en el % de escasez de talento global en los últimos 10 años



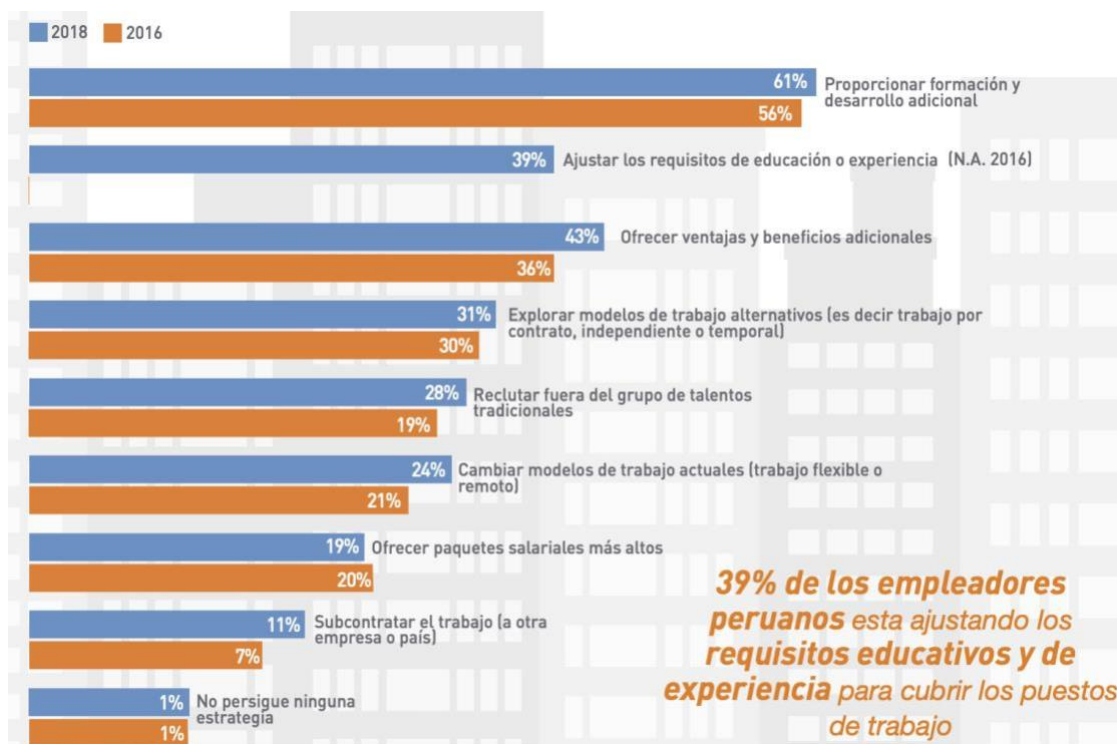
Nota. De “Encuesta de expectativa de empleo”, por ManpowerGroup , 2018

En la actualidad, la mayoría de los 10 puestos más demandados, dentro de los cuales está tecnologías de información como se señala en la encuesta de expectativa de empleo (Manpower, 2018); tanto en Perú como a nivel global, se requiere de una segunda capacitación y no siempre de un título universitario completo. Tecnologías de información, de especialización intermedia, representan el 40% de todos los empleos en la OCDE y cuya demanda es cada vez mayor. (OCDE, 2018)

Casi un tercio de los empleadores afirma que la falta de experiencia de los candidatos constituye la razón principal por la que no pueden cubrir los puestos vacantes, así mismo, el 22% de los empleadores peruanos piensa que los solicitantes carecen de las habilidades duras necesarias. El enfoque en el servicio al cliente en básicamente todas las industrias implica tener la combinación de habilidades adecuadas, por ejemplo, en las funciones de TI requieren de capacidades técnicas, además de las referentes a gestión de personal.

Figura 3.6

Estrategias de empleadores están usando para superar la escasez de talentos



Nota. “De Encuesta de expectativa de empleo”, por ManpowerGroup , 2018

Como se observa en la imagen 3.6, es importante contar con plataformas que ayuden a que los empleadores pueden mapear y mejorar, como nunca antes, su fuerza de trabajo existente y potencial. En la encuesta de expectativa de empleo aseguran los empleadores deben enfatizar el concepto de que el aprendizaje continuo constituye un acuerdo mutuo: esencial para que las personas mejoren sus decisiones de carrera, y fundamental para que las compañías desarrollen el talento que necesitan. Para el éxito en la era digital, resulta clave identificar el potencial futuro de acuerdo a las habilidades que tienen y así impulsar una cultura de capacidad de aprendizaje y detectar programas de entrenamiento acelerado. (Manpower Group , 2018).

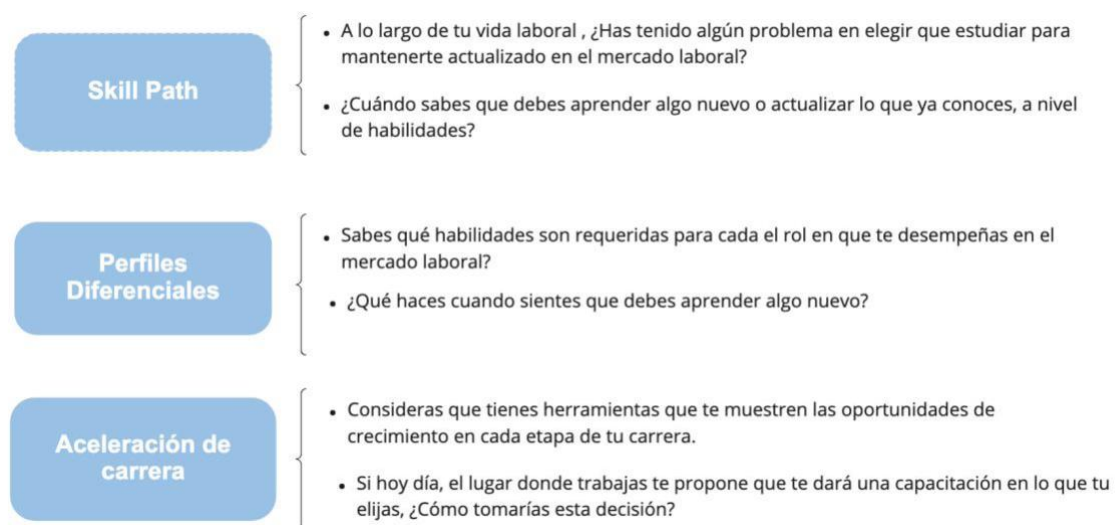
Por lo revisado anteriormente plasman el cambio del mercado de profesionales de TI, esto debido al rápido surgimiento de tecnologías y la necesidad de gente preparada para desarrollar estas. Debe haber estrategias de utilización de esta información de cara a estos profesionales para darles una visión de cómo el mercado se está moviendo y permitirles la opción de aprenderlas, para ello se realizaron entrevistas para ver la perspectiva que estos tienen y así descubrir las oportunidades que pueden haber entre ellos como profesionales y

como se apalanca con las necesidades del mercado en su desarrollo de habilidades, se considera que para un nivel inicial una propuesta de valor sea sugerir qué habilidades técnicas demanda el mercado en su mayoría, a través del análisis de las publicaciones de los puestos de trabajo en las bolsas laborales, y que proporcione información para adquirirlas acompañando a los usuarios en las recomendaciones para su continuo aprendizaje. Actualmente no existe una plataforma en el mercado peruano de dicha utilidad.

3.1.1.1 Entrevistas

Se realizó una entrevista a una muestra de 20 personas entre hombres y mujeres en el rango de 21 a 30 años que trabajan en rubro de tecnologías de la información, el formato de la entrevista se encuentra en anexo 2. Las principales preguntas de la encuesta cubrían algunos aspectos sobre el análisis del entorno considerados anteriormente

Figura 3.7 Estructura de la entrevista realizada



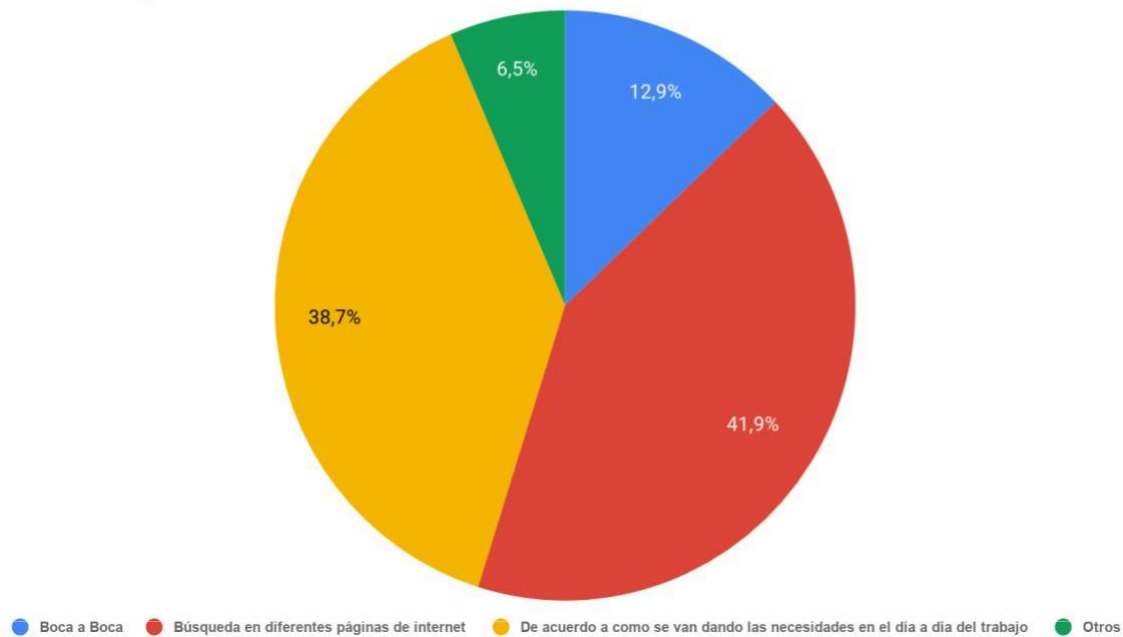
Para la generación de perfiles diferenciales, y conocer la manera de cómo se enteran actualmente de las habilidades técnicas que están siendo solicitadas por la mayoría de empresas, para lograr mantenerse actualizados o encontrar un mejor trabajo. En el gráfico 3.6 se observa que las diferentes respuestas obtenidas indican que, dentro del 41,9% sobre la respuesta de “búsquedas en diferentes páginas de internet, son personas que realizan búsquedas

en diferentes páginas de empleos en el exterior. Para buscar referencias de los requerimientos o tendencias que probablemente actualmente no están desarrolladas en Perú. Otro dato interesante es que en el sector tecnología dependen de mucho de las referencias que puedan tener de grandes empresas tecnológicas reconocidas como GitHub, Dropbox, o comunidades con una frecuencia regular debido a que las herramientas tecnológicas cambian constantemente. Como conclusión de las entrevistas se puede afirmar que las actuales herramientas digitales en torno al empleo tecnológico no permiten a las personas tener más control sobre el desarrollo de sus habilidades técnicas.

Figura 3.8

Medios por los que actualizan habilidades (%)

Porcentajes acerca del medio por el que las personas se actualizan de las habilidades que el mercado requiere



Nota: Total de personas entrevistadas (20) entre hombres y mujeres de 21 - 30 años

3.1.1.2 User Persona

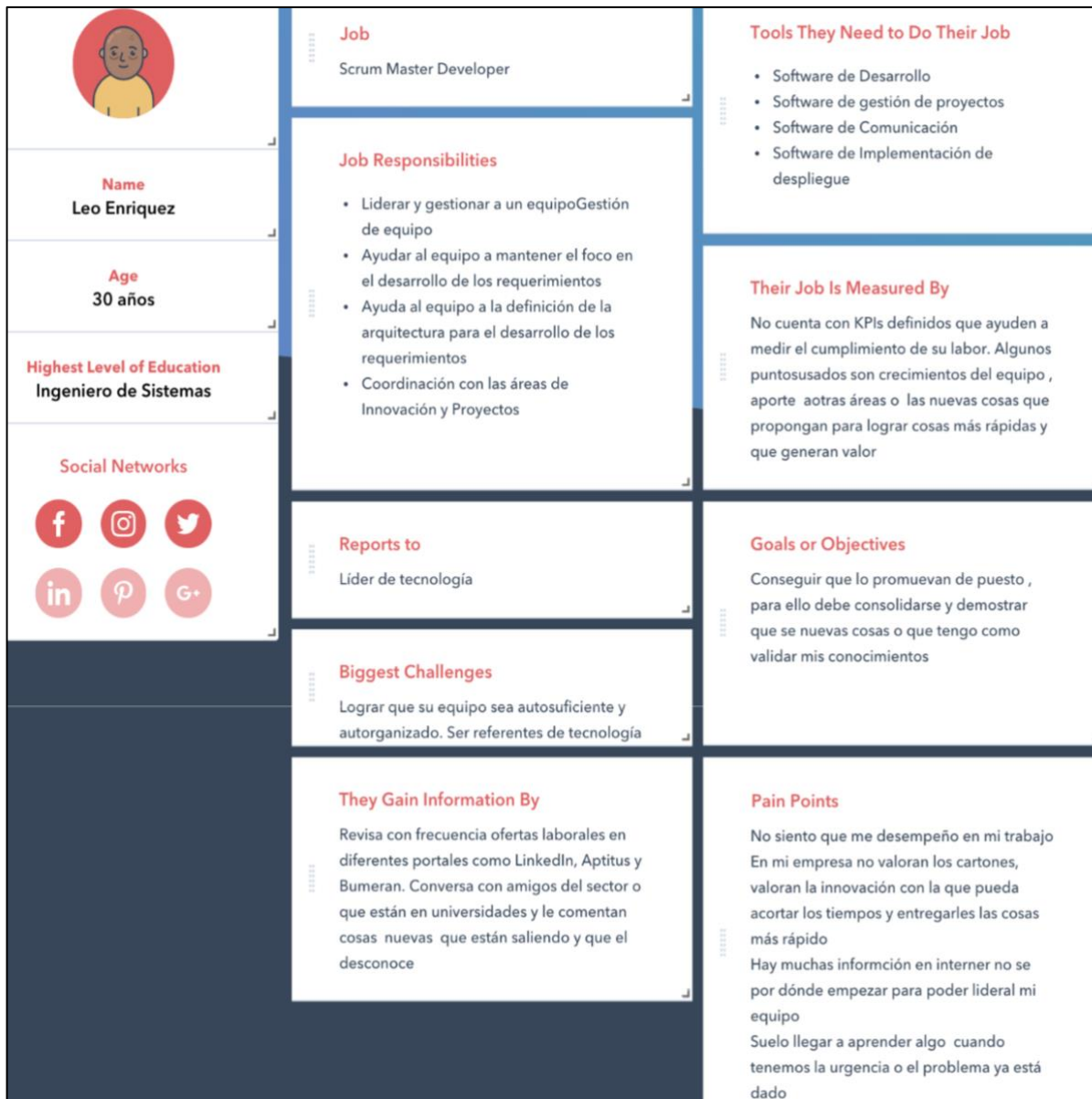
En base a estas entrevistas se sacaron perfiles destacados que representan los patrones que se encontraron en base a nuestra entrevista para conocer más sobre potenciales usuarios y sus motivaciones

Figura 3.9
User Persona de usuario premium



Figura 3.10

User Persona de usuario regular



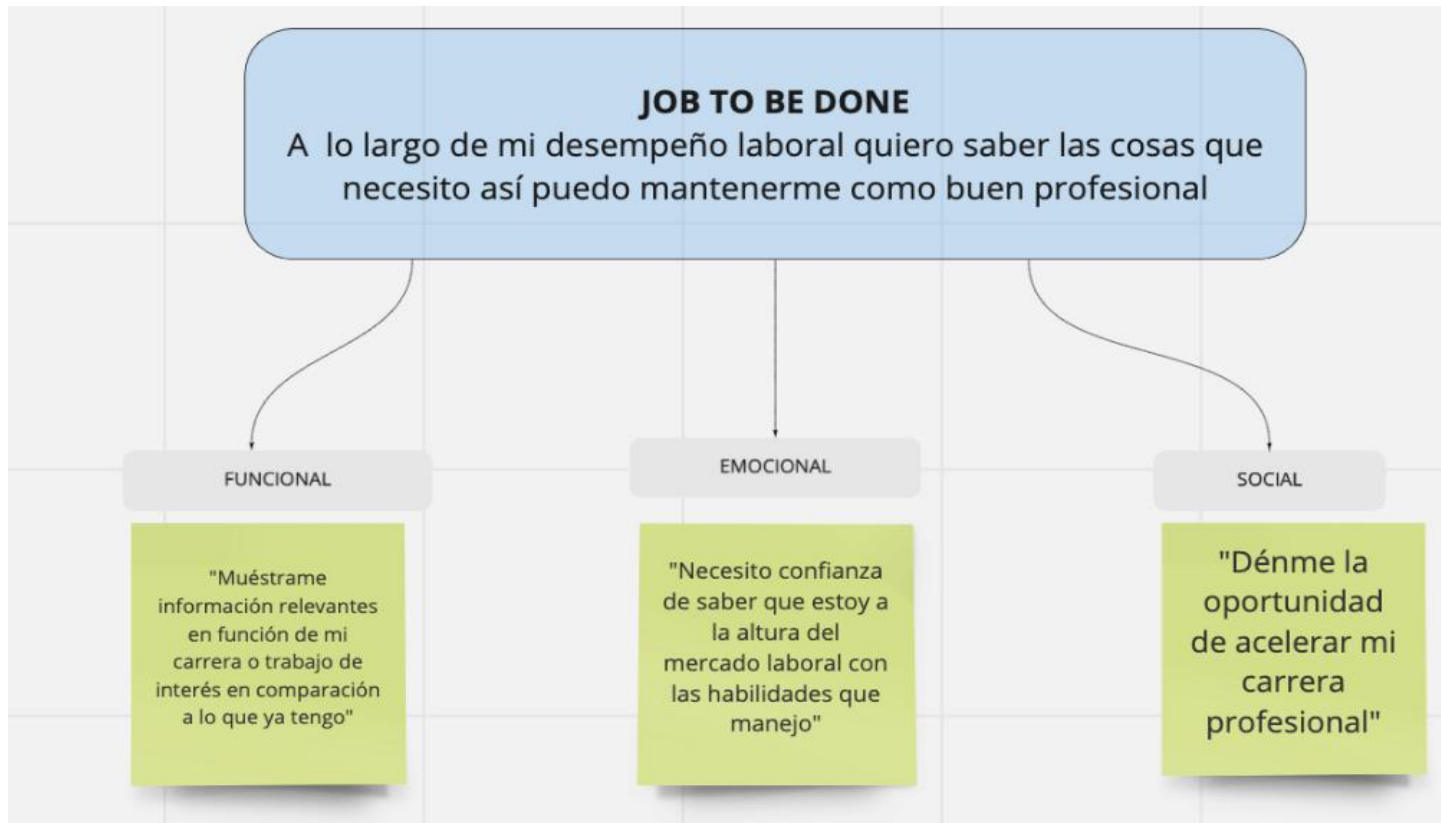
Por lo que se puede concluir, que los usuarios “Premium” son personas que buscan regularmente siempre actualizarse y aprender de manera más proactiva que reactiva, mientras que los usuarios regulares, son las personas son aquellas que solo buscan un ascenso cuando lo creen necesario o tienen la necesidad de aprender algo.

A partir de los perfiles de usuario, se desarrolló el mapa de empatía que se ha especificado en el proceso de design thinking , para tener más claro el perfil del cliente ideal y a partir de ello, podemos construir el “Job to be done” final para comprender todas las necesidades que debe satisfacer nuestro producto final propuesto en torno a nuestro cliente y sus objetivos definidos. (Ver figura 3.9)



Figura 3.11

Job to be done de definición sobre tareas necesitan satisfacer los usuarios.



3.1.1.3 Análisis del mercado potencial de la demanda de profesionales tecnológicos para las empresas

El gigante tecnológico Cisco e IDC, consultora especialista en temas de TI, el 60 % de las organizaciones tendrá articulada su estrategia de transformación digital para el 2020 en América Latina, lo que resultaría en un incremento aún mayor de la demanda de profesionales de tecnología (Oficina de Estadísticas Laborales de los Estados Unidos, 2019). Para poder lograr una estrategia de precios se ha segmentado el mercado peruano, de nuestros potenciales clientes

- Medianas Empresas: Con número de empleados alto (más de 100 empleados) y número de ventas superior a 1700 (UIT). Dentro de las cuales nos enfocamos a 2 perfiles de empresas. (Ver Anexo 3)
- Empresas especializadas en desarrollo de software: Son las empresas del rubro de tecnología, que tienen una alta demanda de perfiles técnicos. Ejemplo: Consultoras como Globant, Everis, Accenture. Se tomó como referencia los datos brindados por la empresa Globant y su proceso de reclutamiento en tecnología en Lima. Ellos tienen un promedio de demanda entre 20 perfiles al año de Perú. Estas empresas serían los clientes potenciales por su volumen de demanda.
- Empresas no especializadas en desarrollo de software: Son las empresas que no pertenecen al rubro de tecnología. Se tomó como referencia una universidad y su proceso de reclutamiento en puestos de tecnología, la cual tiene una demanda de perfiles tecnológicos en promedio de 12 perfiles al año.
- Pequeña Empresas: Con número de empleados de hasta 100 trabajadores y con número de ventas máximo de 1700 (UIT). Dentro de las cuales nos enfocamos a 2 perfiles de empresas.

- Empresas especializadas en desarrollo de software: Son las empresas del rubro de tecnología, en este perfil encajan los startups que según la asociación Peruana de Capital Semilla y Emprendedor (Pecap) que en el primer trimestre del 2019 la inversión en estas aumentó 295%, por lo que estas también serían clientes potenciales por crecimiento del sector. Por ejemplo, la startup “Kambista” y su rotación de personal en el sector tecnología en el último año es de 5 personas anuales en promedio pero esto aumentará de acuerdo al crecimiento de la empre



3.2 Fundamentación de la factibilidad del proyecto

3.2.1 Canvas para “usuarios de la herramienta”

Figura 3.12

Canvas para el usuario de HAB

Aliados Clave Usuarios Empresas que proveen cursos de capacitación en el ámbito tecnológico.	Actividades Clave Desarrollo de lista de habilidades para los diferentes puestos buscados Recomendación de cursos y/o capacitaciones, aprovechando la lista de habilidades construido a partir de la base de datos de publicaciones Análisis de información de correlación de habilidades Estudio de la información de cada perfil registrado por puesto laboral	Propuesta de Valor Ayudar a las personas del rubro tecnológico a conocer y potenciar su progresión laboral facilitándoles la información de las habilidades más relevantes que sus intereses necesitan ante las necesidades cambiantes de la fuerza laboral y mostrarle dónde pueden adquirir dichas habilidades	Relación con los Clientes Estrategia boca a boca Efectos de la red Rankings de usuarios Comunidad de Soporte	Segmento de Clientes Usuarios Premium: Hombres y mujeres en el rango de edad 21-26 que pertenecen al sector tecnológico con un seniority junior/semisenior Usuarios Regular Hombres y mujeres en el rango de edad 27-30 que pertenecen al sector tecnológico con un seniority senior
	Recursos Clave Página Web Infraestructura Adquisición de los Usuarios de la página		Canales Página Web Soporte Redes Sociales (Facebook)	
Estructura de Costos Desarrollo de la Plataforma Hosting de solución Publicidad y Marketing			Vías de Ingreso A corto plazo: no se tendrá ganancias A largo plazo : pago por suscripción de perfiles Pro, por características con funcionalidades extras.	

- Aliados Clave: Los usuarios que harán crecer de manera orgánica a la plataforma y las empresas que proveen cursos de capacitación en el ámbito tecnológico, las cuales serán mostradas en la plataforma
- Actividades Clave:
 - Desarrollo de la lista de habilidades demandadas en mayor cantidad por las empresas para los diferentes puestos buscados
 - Recomendación de cursos, aprovechando la lista de las habilidades construido a partir de la base de datos de publicaciones
 - Análisis de información de correlación de habilidades
- Propuesta de Valor:
 - Ayudar a las personas del rubro tecnológico a conocer y potenciar su progresión laboral facilitándoles la información de las habilidades más relevantes que sus intereses necesitan ante las necesidades cambiantes de la fuerza laboral y mostrarle dónde pueden adquirir dichas habilidades
- Relación con el Cliente:
 - Estrategia boca a boca, por el crecimiento orgánico de la plataforma durante un año.
 - Efectos de la red, a través de las campañas de marketing en redes sociales como Facebook y Youtube
 - Comunidad de Soporte ante cualquier duda o comentario.
- Segmento de Clientes:
 - Usuarios Premium: Hombres y mujeres en el rango de edad 21-26 que pertenecen al sector tecnológico con un seniority junior/semisenior.

- Usuarios Regular: Hombres y mujeres en el rango de edad 27-30 que pertenecen al sector tecnológico con un seniority senior.
- Estructura de Costos: Los costos en los que se incurren son de la mano de obra para la plataforma, el hosting de la solución, publicidad y marketing en redes sociales
- Estructura de Ingresos: Desde el usuario no se tendrá una ganancia a corto plazo, los primeros años de operación no se tendrá ningún ingreso por esta parte debido a que será el periodo en el cual debemos crear una base de usuarios y tráfico en la página, y fidelizar clientes así se mostrarán las sugerencias de cursos a llevar de manera gratuita. A partir del 3er año se podrá tener la creación de perfiles pro, donde se les cobrará por tener funcionalidades extras en la web.



3.2.2 Canvas para “empresas”

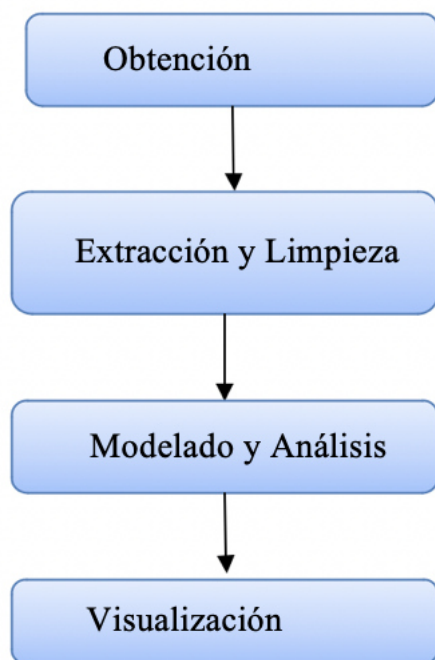
Aliados Clave Recursos humanos de las empresas en su proceso de reclutamiento de perfiles tecnológicos Proveedor de Marketing Personal de Ventas	Actividades Clave Obtención de lista de perfiles para un puesto determinado de trabajo. Venta por obtención de perfiles calificados para un puesto de trabajo solicitado por una empresa.	Propuesta de Valor Brindar una lista de candidatos del rubro tecnológico calificado para el cubrir el puesto tecnológico que solicita una empresa.	Relación con los Clientes Estrategias comerciales de venta Estrategia boca a boca Efectos de la red Comunidad de soporte	Segmento de Clientes Reclutadores Empresas de core tecnológico o en proceso de transformación digital que desarrollan sw
	Recursos Clave Vendedores Comerciales Información de usuarios de la plataforma HAB		Canales Directo Soporte Redes Sociales (Facebook)	
Estructura de Costos Personal Publicidad			Vías de Ingreso Pago (fee) por servicios prestados al solicitar el servicio de filtrado de perfiles para el puesto requerido.	

- Aliados Clave: Son las áreas de recursos humanos de las empresas en su proceso de reclutamiento de perfiles tecnológicos, proveedor de marketing y el personal de ventas que se encargará de hacer conocer el servicio y contactarse con cada empresa.
- Actividades Clave:
 - Selección de una lista de perfiles para un puesto determinado de trabajo.
 - Venta por obtención de perfiles calificados para un puesto de trabajo solicitado por una empresa.
- Propuesta de Valor: Brindar una lista de candidatos del rubro tecnológico calificado para el cubrir el puesto tecnológico que solicita una empresa.
- Relación con el Cliente:
 - Estrategias comerciales de venta a la que tendrá a cargo un personal de ventas
 - Estrategia boca a boca
 - Comunidad de Soporte
- Segmento de Clientes:
 - Reclutadores
 - Empresas de core tecnológico o startups que desarrollan software
- Estructura de Costos: Los costos en los que se incurren son publicidad y el pago al personal de ventas.
- Estructura de Ingresos: Pago (fee) por servicios prestados al solicitar el servicio de filtrado de perfiles para el puesto requerido.

3.2.3 Propuesta de modelo de las tecnologías a usar para cada proceso del desarrollo

La plataforma “Hab” proporcionará búsquedas optimizadas de las habilidades más relevantes en el mercado respecto a los puestos de trabajo requeridos actualmente en base al puesto o sector tecnológico de interés, para ello los datos pasarán por un proceso ETL (extracción, transformación y carga) y , que se describe en la figura 3.13.

Figura 3.13
Proceso ETL



3.2.3.1 Paso 1: Obtención de la información

Por medio de la técnica de web scraping se extraerán todos los anuncios del trabajo relacionados a los puestos de trabajo tecnológicos publicados en las bolsas laborales, para el MVP (producto mínimo viable) se hará la simulación con la plataforma digital LinkedIn y se trabajará con una lista de roles de trabajo definidas.

Este procedimiento contempla de 2 fases: descarga/extracción de datos previamente filtrados y almacenamiento de estos para su posterior exploración, extracción y reducción.

Las publicaciones de trabajo están conformadas por un título y una descripción abierta de texto de los requerimientos que se solicitan. Por medio del proceso de scraping que recibe un conjunto de instrucciones, en este caso el filtro de palabras claves la lista de los puestos predeterminados ej. Backend developer y en base a ellos extraer las descripciones de trabajo para luego almacenarla en una base de datos no relacional como se detalla en la figura 3.14 y 3.15

Figura 3.14

Keywords para extracción de puestos de trabajo

```
keywords = {
  "Frontend": ['Frontend', 'Frontend Developer'],
  "Backend": ['Backend', 'Back End Developer'],
  "QA": ['QA Engineer', 'QA Automation', 'QA Analyst'],
  "Mobile": ['Android Developer', 'IOS Developer'],
  "DevOps": ['DevOps'],
  "Data Scientist": ['Data Science', 'Data Scientist'],
  "Scrum": ['Scrum master', 'Agile coach'],
  "Full Stack": ['Full Stack'],
}
```

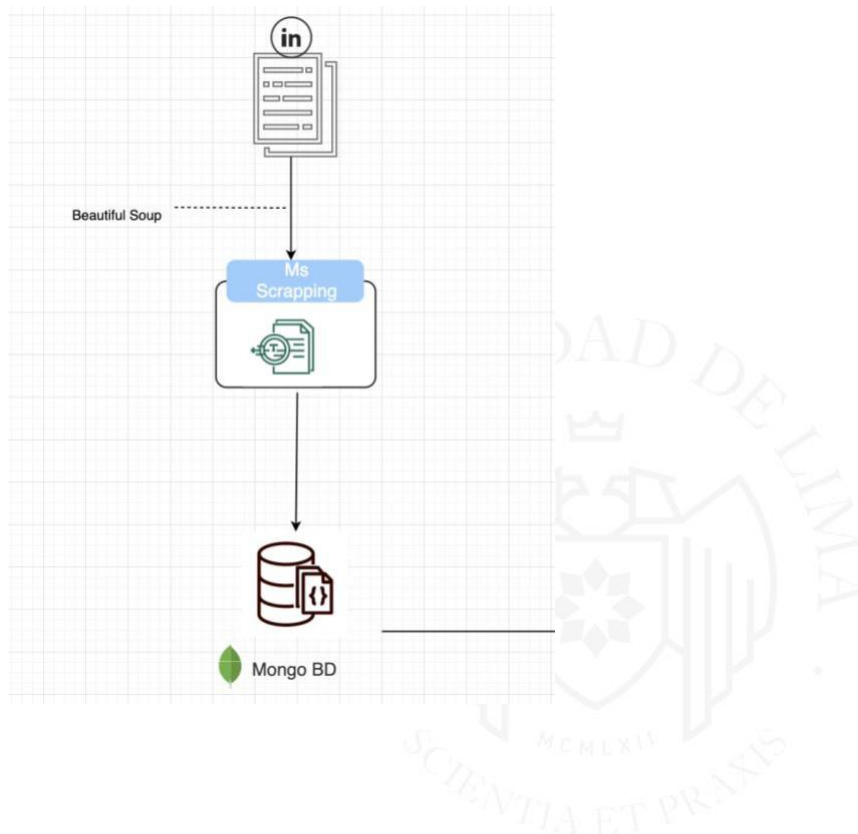
Figura 3.15

Descripción de trabajo extraídos por puestos de trabajo

<pre> _id: ObjectId("5d7fe2f44c9aa79c61449bf9") pk: "40a47d39-0669-4467-9afd-72ba44ab9ade" id: "1489601720" title: "Desarrollador de Python" company: "DSB Mobile" location: "Provincia de Lima, Peru" description: "¡Nuestro equipo en Lima sigue creciendo! En DSB Mobile nos encontramos..." career: "Full Stack" </pre>
<pre> _id: ObjectId("5d7fe2f44c9aa79c61449bfa") pk: "9ce3c1ee-27b7-4b3a-9ab2-64845cfe78b8" id: "1488392418" title: "Developer Android" company: "Belatrix Software" location: "Lima, PE" description: "Lima, LimaBelatrix is a privately held software development firm.Our m..." career: "Mobile" </pre>
<pre> _id: ObjectId("5d7fe2f44c9aa79c61449bfb") pk: "9690c4a6-92b4-4fdb-ab31-3d8d6be597ff" id: "1467220267" title: "Devops" company: "Belatrix Software" location: "Lima, PE" description: "Belatrix is a privately held software development firm. Our mission is..." career: "DevOps" </pre>
<pre> _id: ObjectId("5d7fe2f44c9aa79c61449bfc") pk: "690cf5b1-c30f-4b18-9411-179d0beeea4b" id: "1467220284" title: "Scrum Master" company: "Belatrix Software" location: "Lima, PE" description: "Belatrix is a privately held software development firm. Our mission is..." career: "Scrum" </pre>

De herramienta de web scraping se usó la librería open source disponible para python denominada “Beautiful Soup” que permite analizar documentos HTML y trabajar con una base de datos no relacional MongoDB. El procesamiento estará dentro de un microservicio a cuál se denominó Ms Scrrapping, como se observa en la figura 3.16

Figura 3.16
Extracción y almacenamiento de puestos de trabajo



Por otro lado, se creó la lista de habilidades técnicas para la primera versión del prototipo que se formará a partir de un dataset de competencia extraído de las API ofrecidas de las herramienta stackoverflow , que es referentes en el sector tecnología. ver figura 3.16

Figura 3.17
Dataset de habilidades técnicas

```

"items": [
  {
    "has_synonyms": true,
    "is_moderator_only": false,
    "is_required": false,
    "count": 1830184,
    "name": "javascript"
  },
  {
    "has_synonyms": true,
    "is_moderator_only": false,
    "is_required": false,
    "count": 1561132,
    "name": "java"
  },
  {
    "has_synonyms": true,
    "is_moderator_only": false,
    "is_required": false,
    "count": 1323031,
    "name": "c#"
  },
  {
    "has_synonyms": true,
    "is_moderator_only": false,
    "is_required": false,
    "count": 1291732,
    "name": "php"
  },
  {
    "has_synonyms": true,
    "is_moderator_only": false,
    "is_required": false,
    "count": 1202572,
    "name": "android"
  },
  {
    "has_synonyms": true,
    "is_moderator_only": false,
    "is_required": false,
    "count": 284632,
    "name": ".net"
  },
  {
    "has_synonyms": true,
    "is_moderator_only": false,
    "is_required": false,
    "count": 280488,
    "name": "node.js"
  },
  {
    "has_synonyms": true,
    "is_moderator_only": false,
    "is_required": false,
    "count": 265287,
    "name": "json"
  },
  {
    "has_synonyms": true,
    "is_moderator_only": false,
    "is_required": false,
    "count": 262871,
    "name": "sql-server"
  },
  {
    "has_synonyms": true,
    "is_moderator_only": false,
    "is_required": false,
    "count": 258423,
    "name": "angularjs"
  },
  {
    "has_synonyms": true,
    "is_moderator_only": false,
    "is_required": false,
    "count": 234540,
    "name": "swift"
  },
  {
    "has_synonyms": true,
    "is_moderator_only": false,
    "is_required": false,
    "count": 602056,
    "name": "ios"
  },
  {
    "has_synonyms": true,
    "is_moderator_only": false,
    "is_required": false,
    "count": 591290,
    "name": "css"
  },
  {
    "has_synonyms": true,
    "is_moderator_only": false,
    "is_required": false,
    "count": 564518,
    "name": "mysql"
  },
  {
    "has_synonyms": true,
    "is_moderator_only": false,
    "is_required": false,
    "count": 496798,
    "name": "sql"
  },
  {
    "has_synonyms": true,
    "is_moderator_only": false,
    "is_required": false,
    "count": 347190,
    "name": "asp.net"
  },
  {
    "has_synonyms": true,
    "is_moderator_only": false,
    "is_required": false,
    "count": 307457,
    "name": "ruby-on-rails"
  }
]

```

Nota: De “Stack Exchange API”, por Stackoverflow, 2019 (<https://api.stackexchange.com>)

3.2.3.2 Paso 2: Limpieza y pre procesamiento de la información

Por otro lado, las descripciones almacenadas tienen que pasar por un pre-procesamiento de texto, para ello se usó las herramientas de Procesamiento del Lenguaje Natural con el objetivo de eliminar datos inútiles y obtener una lista de palabras a partir de texto procesado. Para realizarlo se consideró una herramienta adaptable y que tenga capacidad de procesamiento, la herramienta utilizada es la librería NLTK (Natural Language Toolkit) desarrollada en el lenguaje de programación Python, la cual es open source, su facilidad de usar y cuenta con soporte de la comunidad. A continuación, se detalla cada paso seguir:

- Eliminación del ruido.

Las descripciones pasan por el proceso de limpieza (pre procesamiento) para filtrar datos o palabras que en las descripciones de trabajos sean inútiles, y se eliminan para evitar que ocupen un tiempo de procesamiento, así mismo las tareas de eliminación de ruido que incluye eliminar encabezados de archivos de texto, pies de página, así como la eliminación de HTML y la puntuación. También se usó la librería BeautifulSoup, que utiliza expresiones regulares para eliminar las etiquetas HTML y cualquier elemento intermedio.

- Tokenización

Las publicaciones de trabajos son convertidas en oraciones, las oraciones tokenizadas en palabras, etc. Se va a tokenizar cada descripción en una lista de palabras. Esto se hace usando una función de NLTK.

- Normalización

La normalización se refiere a una serie de pasos relacionados a poner todo el texto en igualdad de condiciones, nuestra normalización la hacemos a nivel de palabras, se removió los caracteres NON-ASCII de la lista de palabras, también preposiciones, conjunciones, artículos, adverbios, pronombres a estas palabras se denominan “stop words”.

Figura 3.18

Pre procesamiento de descripciones de trabajo



3.2.3.3 Paso 3: Análisis y Modelado

El siguiente paso, se encontró en las descripciones ya pre procesadas de trabajo las palabras de la ontología de habilidades para poder conocer las más solicitadas y relevantes que coincidan con cada puesto de trabajo. Así mismo como las descripciones de trabajo son insertadas en un campo libre de texto, podría haber habilidades que se refieran a lo mismo pero que están escritas de manera similar en comparación a las del diccionario de datos, por ejm: la tecnología “vue” algunas descripciones la escribe como vue.js o el lenguaje de programación “go” y “goland” se refieren a lo mismo. Es por ello que se evaluaron diferentes algoritmos de coincidencia de cadenas, para ver cual tenía mayor valor de reconocimiento de cadenas.

La métrica usada para la elección es la inversa del valor obtenido del porcentaje de similitud en la comparación de palabras, por lo que el resultado se convierte en porcentaje y es más rápido de leer. Como se observa en la figura 3.23 los tres algoritmos comparados son con la distancia de Levenshtein (que es implementada por la librería fuzzy de Python y una derivación de la misma llamada partial fuzzy) y la distancia de Jaro-Winkler que también es implementada por la librería jaro-winkler de Python) respectivamente. Se implementaron los 3 para poder hacer una comparativa con data de prueba.

Figura 3.19*Resultado comparación de algoritmos de similitud de cadenas*

Word 1	Word 2	Fuzzy	Partial Fuzzy	Jaro Winkler
vue.js	Vue	44	67	83
vue.js	vue	67	100	88.0
vue.js	Vue.js	83	83	100.0
vue.js	VUEJS	0	0	94.0
vue.js	vuejs	91	80	96.0
vue.js	VueJS	36	40	94.0
vue.js	vueJS	55	60	96.0
vue.js	VueJs	55	60	94.0
node.js	Node	55	75	86.0
node.js	node	73	100	91.0
node.js	nodejs	92	83	97.0
node.js	NodeJS	46	50	95.0
node.js	NodeJs	62	67	95.0
node.js	nodeJS	62	67	97.0
node.js	Node.js	86	86	100.0
Ember	EMBERJS	17	20	91.0
Ember	emberjs	67	80	90.0
Ember	ember.js	62	80	88.0
Ember	Ember.js	77	100	92.0
Ember	ember	80	80	100.0
Ember	emberJS	67	80	90.0
Ember	EmberJS	83	100	94.0
Golang	GO	25	50	80.0
Angular	AngularJS	88	100	96.0

Como se demuestra, el algoritmo que mejor representa la similitud de cadenas es el algoritmo de Jaro Winkler , este es el que se utilizó para el análisis de las descripciones de trabajo por cada puesto de trabajo para luego ponderar su frecuencia agrupados por cada uno de ellos.

Como segunda parte del procesamiento, fue la búsqueda de qué otras habilidades están más relacionadas con cada una de las habilidades encontradas como relevantes, para poder conocer qué otras habilidades se solicitan en mayor demanda y comprender qué habilidades profesionales se relacionan entre sí, según nuestros datos. Por ejemplo, si alguien solicita machine learning como una habilidad requerida, es probable que también se solicite que se

conozca habilidades hadoop, spark o python u otras no directamente relacionadas como design thinking.

Para esto se construye una matriz que relaciona el diccionario de habilidades con las descripciones de trabajo, esta matriz presenta una fila por habilidad, junto con una columna para cada descripción de trabajo y la cantidad de veces que aparece la habilidad. Así se forma un vector para cada habilidad (skill) representado en su fila, ahora podemos calcular similitudes entre pares de habilidades calculando la similitud de sus vectores.

Tabla 3.1

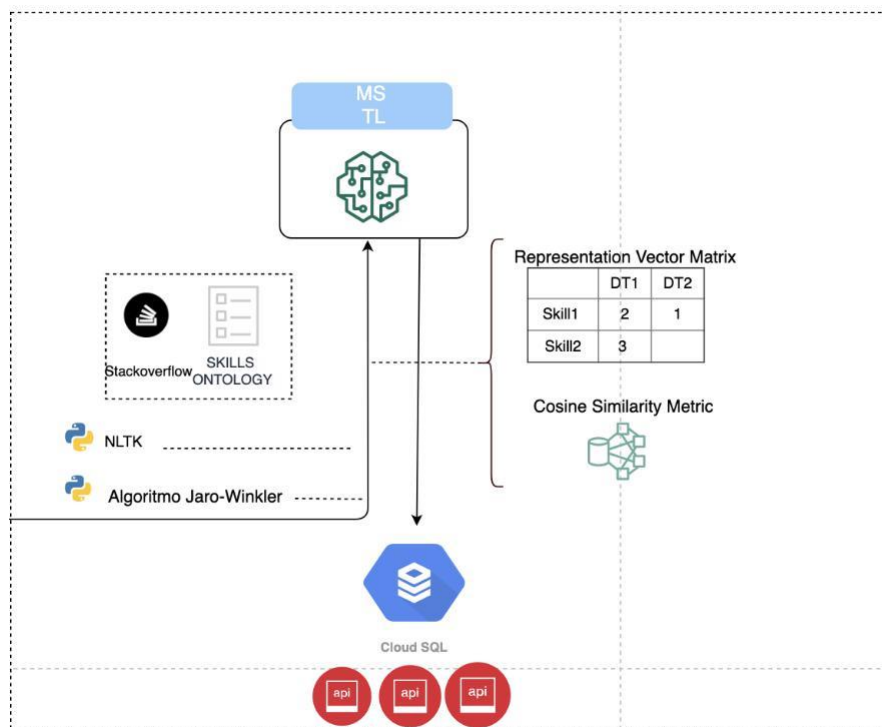
Tabla de matriz de relación de habilidades con descripciones del trabajo

Habilidades/Descripción	Descripción de Trabajo	Descripción de Trabajo 2
Ajax	0	1
Docker	1	0
Kubernetes	2	1

Una vez que tenemos los vectores para cada habilidad, usamos la métrica de similitud de coseno para calcular las habilidades más similares para una habilidad dada.

Toda la información expuesta, procesada y ya modelada se almacenan en una base de datos relacional MySQL, se decidió usar el servicio Google db-n1-standard-1 que simplifica la configuración, escalamiento y copias automáticas de la base de datos.

Figura 3.20
Procesamiento de datos en MS TL

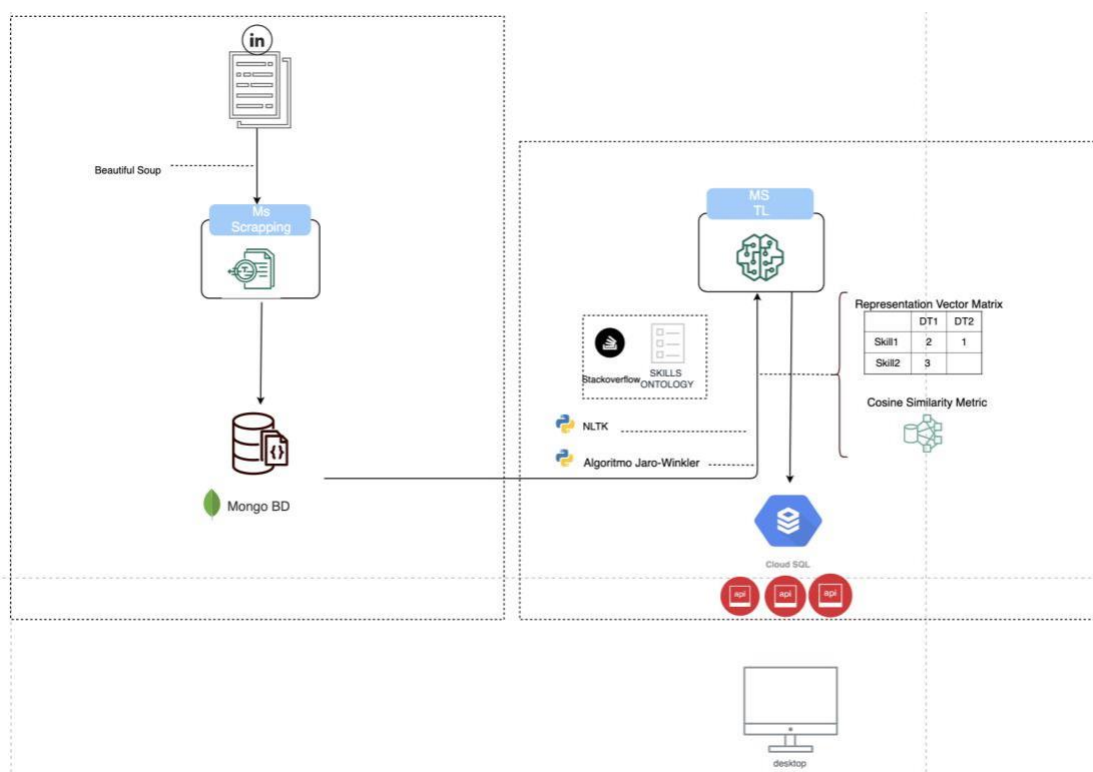


3.2.3.4 Paso 4: Visualización

Finalmente, los datos almacenados serán consumidos por los servicios creados para poder mostrarlos en la aplicación web. La plataforma estará alojada en la nube Google, y cada microservicio usará docker para su creación y despliegue.

El gráfico que resume la arquitectura completa se muestra a continuación:

Figura 3.21
Proceso de Web Scraping y NLP



3.3 Fundamentación de la viabilidad del proyecto

3.3.1 Descripción del proyecto

Mediante una plataforma web, enfocada a los usuarios profesionales en el rubro de tecnología, donde podrán registrar sus datos personales y registrar las habilidades con las que ya cuentan para generar un perfil, para mostrarles la lista de habilidades que son más requeridas en el sector laboral del puesto de trabajo seleccionado de acuerdo al análisis de las descripciones de puestos de trabajo las bolsas laborales online, así mismo se les recomendará cursos en donde pueden adquirirlas o potenciarlas.

El análisis del perfil de los usuarios, luego será utilizado para filtrar una lista de potenciales perfiles buscados en empresas en busca de talento tecnológico y así brindar el servicio pagado por dichas empresas para su proceso de reclutamiento.

3.3.2 Ingresos

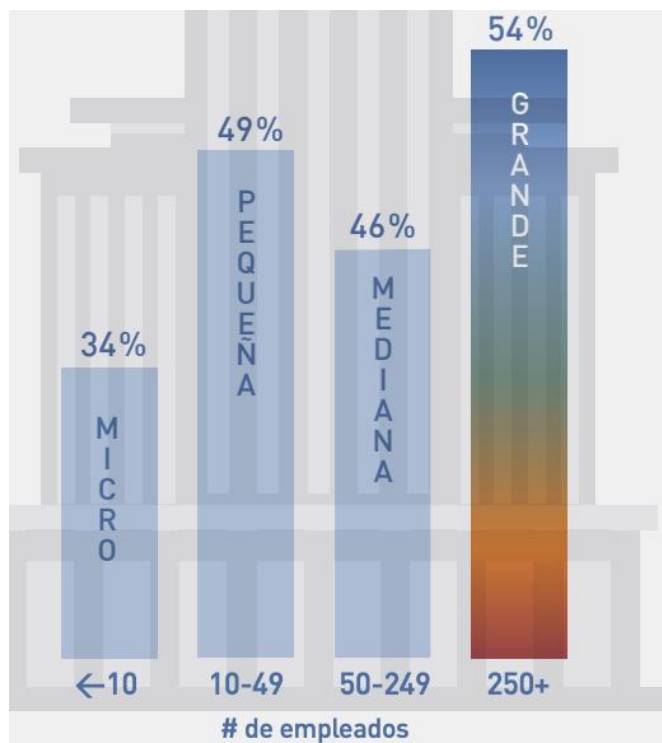
El usuario no se tendrá una ganancia a corto plazo, los primeros años de operación no se tendrá ningún ingreso como se detalló en el canvas.

La generación de ingresos inicial del proyecto es el pago (success fee) por la venta de información de perfiles de talento tech-digital a empresas que necesitan cubrir sus vacantes, que puedan ser el mejor fit a nivel técnico para ayudar en su proceso de reclutación para que sea más rápido y eficiente; ya que se cuenta con la información de los usuarios registrados. Cada pago solo se realiza si se contrata algún perfil de la lista que se le brinda.

Como se demuestra en la encuesta de Encuesta de escasez de talento 2018, llevada a cabo por Manpower, la gran Empresa tiene mayor dificultad para cubrir vacantes: el 54% enfrenta este tipo de retos en la contratación e incluso un 5% dice estar teniendo más dificultades ahora que hace un año, es por ellos que nuestro que nos basaremos en esta información para segmentar nuestra propuesta económica.

Figura 3.22

Dificultad para cubrir vacantes según el tamaño de la empresa



Nota. De “Encuesta de expectativa de empleo”, por ManpowerGroup , 2018

3.3.2.1 Cálculo de la demanda de usuarios y demanda potencial de empresas a las que se les podría vender los perfiles registrados en la web hab.

A continuación, se presenta la demanda potencial considerando:

Promedio del número de empresas en Lima que demandan personal del rubro tecnológico en base a:

- Número de startups registradas en Lima. Innovación, Tecnología, Digitalización y Formalización del Ministerio de la Producción (2018)
- Número de empresas registradas de Software. Apesoft (2017) (Se consideró demanda potencial en el escenario de que este sería el sector donde marketing y ventas sería su mercado objetivo)

- Disposición a usar una nueva herramienta para reclutamiento de talento tecnológico (ver Anexo 4)
- Disposición a pagar una comisión por precio success fee (ver Anexo 4)

Figura 3.23

Demanda potencial de empresas

	Startups	Empresas Software grandes
Total de número de Empresas Lima (Startup Perú - Asociación Peruana de Productores de Software)	309	500
% de empresas que utilizarían el servicio de perfiles para reclutación	95%	95%
% de empresas que pagarían por success fee /membresía	100%	100%
Total de demanda que usaría el servicio de perfiles para reclutación de talento de tecnología	294	475

Promedio del número potencial de demanda de personas del rubro tecnología.

- Número promedio de personas que se demandarán para puestos de trabajo del sector tecnología entre técnicos y profesionales. (Ministerio de trabajo y promoción del empleo, 2018) (Ver anexo 6)

Figura 3.24

Demanda potencial de personas

	Número de personas
DEMANDA OCUPACIONAL DE PROFESIONALES TÉCNICOS	9 252
DEMANDA DE PROFESIONALES TÉCNICOS POR CONSULTORÍA O SERVICIOS PRESTADOS	2419
DEMANDA DE TECNICOS EN PROGRAMACIÓN POR COMPUTADORA	1485
SUB TOTAL DEMANDA TECNICOS	1485
DEMANDA OCUPACIONAL DE PROFESIONALES UNIVERSITARIOS	6 723
DEMANDA PROFESIONALES UNIVERSITARIOS POR CONSULTORÍA O SERVICIOS PRESTADOS	1 477
DEMANDA DE PROFESIONALES EN INFORMATICA	554
DEMANDA PROGRAMACIÓN POR COMPUTADORA	397
SUB TOTAL DEMANDA PROFESIONALES	951
TOTAL DEMANDA TECNICOS + TOTAL DEMANDA PROFESIONALES UNIVERSITARIOS	2436

3.3.2.2 Dimensionamiento del mercado

De acuerdo con las encuestas realizadas a 15 empresas entre ellas startups y empresas enfocadas en el desarrollo de software de acuerdo a la venta de perfiles para las vacantes requeridas en el rubro tecnológico, todas indican que tienen una inversión en la búsqueda y contratación de perfiles de tecnología y la inversión se presenta bajo las siguientes modalidades: (Ver anexo 5).

Del resultado de las encuestas en un mayor porcentaje, las empresas medianas realizan un pago de membresías anuales en las principales bolsas de trabajo por su rotación de personal.

- El 43% solo utiliza la plataforma de LinkedIn con un precio aproximado de S/ 30000 soles anuales.
- El otro 15% sólo utiliza la plataforma Boomerang con un gasto promedio de s/ 20000 soles anuales.
- El 29% usa más de una bolsa laboral como LinkedIn, Boomerang y Computrabajo con gasto promedio de 45000 soles anuales.
- Por otro lado, la mayoría de los startups hacen un pago a demanda a empresas que se encargan del reclutamiento del personal .
- El 65% usa empresas especializadas en reclutamiento que brinda perfiles con las habilidades técnicas requeridas como Krowdy y Open Talent, de acuerdo con la encuesta, en su mayoría estas empresas cobran un porcentaje del sueldo aproximadamente el 80% de la persona a reclutar que varía por el rango del puesto laboral. Como se puede visualizar en la figura 3.25 la referencia de los montos a invertir en la empresa Open Talent.

Figura 3.25

Precios de consultoría Open Talent

PRICING		
Candidates level	Success Fee Pre -Platform's Launch: Corp.	Success Fee Pre -Platform's Launch: Startup
Junior	\$850.00	\$350.00
Mid / Semi Senior	\$1,164.00	\$479.00
Senior	\$1,865.00	\$719.00
Architect	\$2,550.00	\$985.00
Boss Architect	\$3,499.00	\$1,349.00

Nota: De "Tarifas por consultoría", por Open Talent , 2019

- El 35% de las empresas contrata las bolsas laborales, pero a demanda. Luego se establece los montos basados en las condiciones del mercado y la competencia.

Existen 2 tipos de pagos dentro de los cuáles se toman en cuenta 2 variables:

- Pago por persona
- Nivel de candidatos: Se establece 3 categorías como nivel de candidatos detalle en la tabla 3.2

Tabla 3.2

Tabla de descripción de nivel de puestos de trabajo

Nivel	Años de Experiencia
Junior	1-2 años de experiencia
Semi- senior	3-4 años de experiencia
Senior	> 5 años de experiencia

- Porcentaje de sueldo por puesto.

Estas estimaciones en promedio del sueldo se muestra en la tabla 3.3 (Tu Salario - Perú , 2019).

Tabla 3.3

Descripción de nivel de puestos de trabajo

Nivel	Sueldo aproximado
Junior	2000
Semi- senior	4000
Senior	6000

Así la fijación de montos respecto a los roles es la misma para las empresas grandes como para los startups, se cobrará a los startups por reclutación de personas solo el 80% del sueldo de acuerdo al nivel al que corresponde ver tabla 3.4. Exceptuando que como parte de la estrategia respecto a las empresas competidoras.

Tabla 3.4*Detalle de la proyección de sueldo/cantidad*

			Precio de Servicio		
		<i>Sueldo</i>	<i>Cantidad</i>	Pago x 1 mes	Pago x 1 año
Grande	<i>Junior</i>	2000	1	2000	24000
	<i>Semiseniior</i>	5000	1	5000	60000
	<i>Senior</i>	8000	1	8000	96000
			TOTAL INGRESO EMP GRANDE	15000	180000
			Precio de Servicio		
Porcentaje de sueldo a cobrar	80.00%				
		<i>Sueldo</i>	<i>Cantidad</i>	Pago x 1 mes	Pago x 1 año
Mediana o Startups	<i>Junior</i>	2000	1	1600	19200
	<i>Semi-seniior</i>	5000	1	4000	48000
	<i>Senior</i>	8000	1	6400	76800
			TOTAL INGRESO EMP MED/STARTUP	12000	144000

3.3.2.3 Precios

Se invertirá S/ 356,604,000 para poder cubrir el capital necesario para el gasto operativo y la adquisición de activos fijos. La inversión se financiará con el aporte capital de los 3 socios fundadores con una tasa de retorno del 6%. Los supuestos considerados fueron los siguientes

- Tipo de cambio de dólar 3.365
- Sólo se contratará a partir de los 2 últimos meses del 1er año al vendedor ya que el primer año solo será para ganar usuarios que generen perfiles.

Tabla 3.5
Activos Fijos

Activos fijos	Monto
MacBook Pro x 5 miembros del equipo	S/ 35000.00
Equipo de oficina	S/ 500.00
Total	S/ 40000.00

Tabla 3.6
Detalle del gasto único

Activos fijos	Monto
Constitución de la empresa	S/ 1,500.00
Registro de la marca	S/ 750.00
Total	S/2250.00

Tabla 3.7*Detalle de gastos mensuales*

Gastos Mensuales	Pago Mensual	Pago Anual
Infraestructura	S/ 967	S/ 11604
Suscripción Gitlab (control de versiones) para 4 desarrolladores	S/ 54	S/ 648.00
Compute Engine	S/ 242	S/ 2904,00
Cloud SQL	S/ 175	S/ 2100
Kubernetes Engine	S/246	S/2952
Servicio correos	S/ 250	S/3000
Personal	S/ 31000	S/ 372000
Personal Ventas Sueldo (1er año solo 2 últimos meses)	S/ 6000	S/ 72000
Personal Desarrollo Aplicativo (Front, Back, QA,UI)	S/20000	S/ 240000
Personal Administrativo	S/5000	S/ 60000
Publicidad	S/ 2750	S/ 33000
Publicidad Facebook Ads	S/ 1500	S/ 18000
Publicidad LinkedIN Ads	S/ 700	S/ 8400
Publicidad Youtube Ads	S/ 550	S/ 6600
Total	S/ 34717	S/ 6604

3.3.2.4 Precios

Cálculo de la venta e ingresos proyectada

A continuación, para el resumen de la venta e ingresos proyectada por año, se han considerado los siguientes supuestos:

- La demanda de perfiles anuales de acuerdo al nivel del puesto tiene una relación 3 a 1 del puesto con más experiencia al de la menor experiencia-
- De acuerdo con la encuesta realizada se infieren los siguientes puntos

- De las empresas cuyo sector no es el desarrollo de software la demanda de perfiles tecnológicos es de un promedio de 6 perfiles al año.
- De las empresas cuyo sector es el desarrollo de software la demanda de perfiles tecnológicos duplica a las no tecnológicas siendo de un promedio de 12 perfiles al año.

Tabla 3.8

Demanda anual de puestos de tecnología por tipo de empresa

		<i>Sueldo</i>	<i>Cantidad</i>	<i>Anual</i>
Grande Tec	<i>Junior</i>	2000	6	12000
	<i>Semisenior</i>	4000	4	24000
	<i>Senior</i>	6000	2	12000
				48000
Startup	<i>Junior</i>	2000	1	2000
	<i>Semisenior</i>	4000	1	4000
	<i>Senior</i>	6000	1	6000
				12000
Grande No Tec	<i>Junior</i>	2000	3	6000
	<i>Semisenior</i>	4000	2	12000
	<i>Senior</i>	6000	1	6000
				24000

- El primer año no se generarán ingresos por medio de las empresas que necesiten reclutamiento, debido a que los esfuerzos serán para generar tráfico y usuarios a la página web para el registro de su información y mapeo de habilidades para recolectar perfiles.

- De los 294 startups actualmente existentes en promedio en el Perú, se planea en los 5 años contar con el 17% del sector, aproximado en promedio 50 startups. Para el segundo año con una sola persona orientada a las propuestas comerciales se considera la adquisición de 5 startup al año , duplicando el número hasta el 3er año y luego creciendo a un 20%
- De las 475 empresas dedicadas al desarrollo de software se planea en los 5 años contar con el 17% del sector, aproximando en promedio 80 empresas, para el segundo año se consideran la adquisición de 6 empresas anualmente, duplicando el número hasta el 4to año y luego creciendo a un 20%.
- Finalmente, las empresas que no son se centran en el desarrollo de software tendrán un crecimiento menos acelerado, manteniéndose en un 2% anual.

A continuación, se presenta los ingresos proyectados:

Tabla 3.9
Ingresos proyectados por año y tipo de empresa

		Año 1	Año 2	Año 3	Año 4	Año 5
Cantidad de Empresas	<i>Startups</i>	0	5	10	12	15
	<i>Empresas de SW</i>	0	6	12	24	30
	<i>Empresas Grandes</i>	0	2	3	4	5
Ingresos	<i>Startups</i>	0	60000	120000	144000	180000
	<i>Empresas de SW</i>	0	150000	300000	600000	750000
	<i>Empresas Grandes</i>	0	50000	75000	100000	125000
Total, Ingresos		0	260000	495000	844000	1055000

3.3.3 Flujo de caja

En la tabla 3.10 se muestra el detalle final del flujo de caja para los 5 años siguientes en donde se puede observar que a partir del tercer año los flujos se vuelven positivos.

Tabla 3.10
Flujo de caja

		Año 0	Año 1	Año 2	Año 3	Año 4	Año 5
Ingresos	<i>Startups</i>		0	60000	120000	144000	180000
	<i>Empresas de SW</i>		0	150000	300000	600000	750000
	<i>Empresas Grandes</i>		0	50000	75000	100000	125000
Total, Ingresos		0	0	260000	495000	844000	1055000
Egresos	<i>Infraestructura</i>	11604	11604	11604	11604	11604	11604
	<i>Personal*</i>	312000	372000	372000	372000	372000	372000
	<i>Publicidad</i>	0	33000	33000	33000	33000	33000
Total, Egresos		323604	416604	416604	416604	416604	416604
Utilidad		-323604	-416604	-156604	78396	427396	638396
Saldo Inicial			-416,604	-416604	-156604	78396	427396
Flujo Neto		-323604	-833,208	-573208	-78208	505792	1065792
*Nota: En el año 0 solo se considera el sueldo del asistente comercial los 2 últimos meses							

Por último, se presenta en la tabla 3.9 los cálculos de la VAN y TIR correspondientes:

Tabla 3.11*Cálculo del VAN y el TIR*

VAN	25,403.44
Tasa de interés para rentabilidad (anual)	6%
TIR	7%

Basado en el análisis del VAN del proyecto, se puede inferir que se llega a generar una ganancia de S/ 25,403,4 al 6% de tasa de rentabilidad. La inversión produciría ganancias por encima de la rentabilidad exigida .

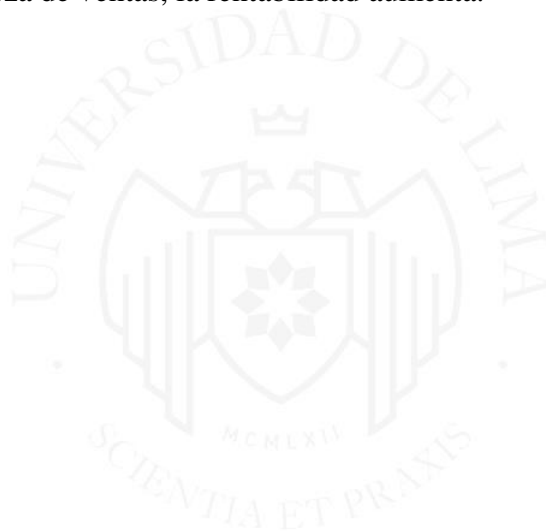
En el cálculo del TIR se concluye que el proyecto generará una rentabilidad mayor a la esperada por los inversores y el proyecto es viable. La TIR es la máxima tasa de descuento que un proyecto puede tener para ser rentable y se calculó tasa interna en función de flujos de efectivo, el proyecto brinda 7% entonces se trata de un proyecto atractivo.

Se han realizado escenarios de evaluación, disminuyendo la tasa de crecimiento de usuarios en la plataforma en -5%. Los resultados para el escenario de disminución nos indica que el proyecto no es viable debido a que el VAN es negativo como se observa en la tabla 3.10

Tabla 3.12*Análisis del VAN disminuyendo 5% los ingresos por parte de las empresas*

VAN	-\$79,791.31
Tasa de interés para rentabilidad (anual)	6%

Por lo que , se concluye que el escenario propuesto tiene riesgos, los cuales buscaremos mitigar en una segunda etapa del proyecto, asegurando las características mínimas para hacer rentable la propuesta de negocio considerando que, si se genera una mejor fuerza de ventas, la rentabilidad aumenta.



CAPITULO IV: DEFINICIÓN DEL PROYECTO

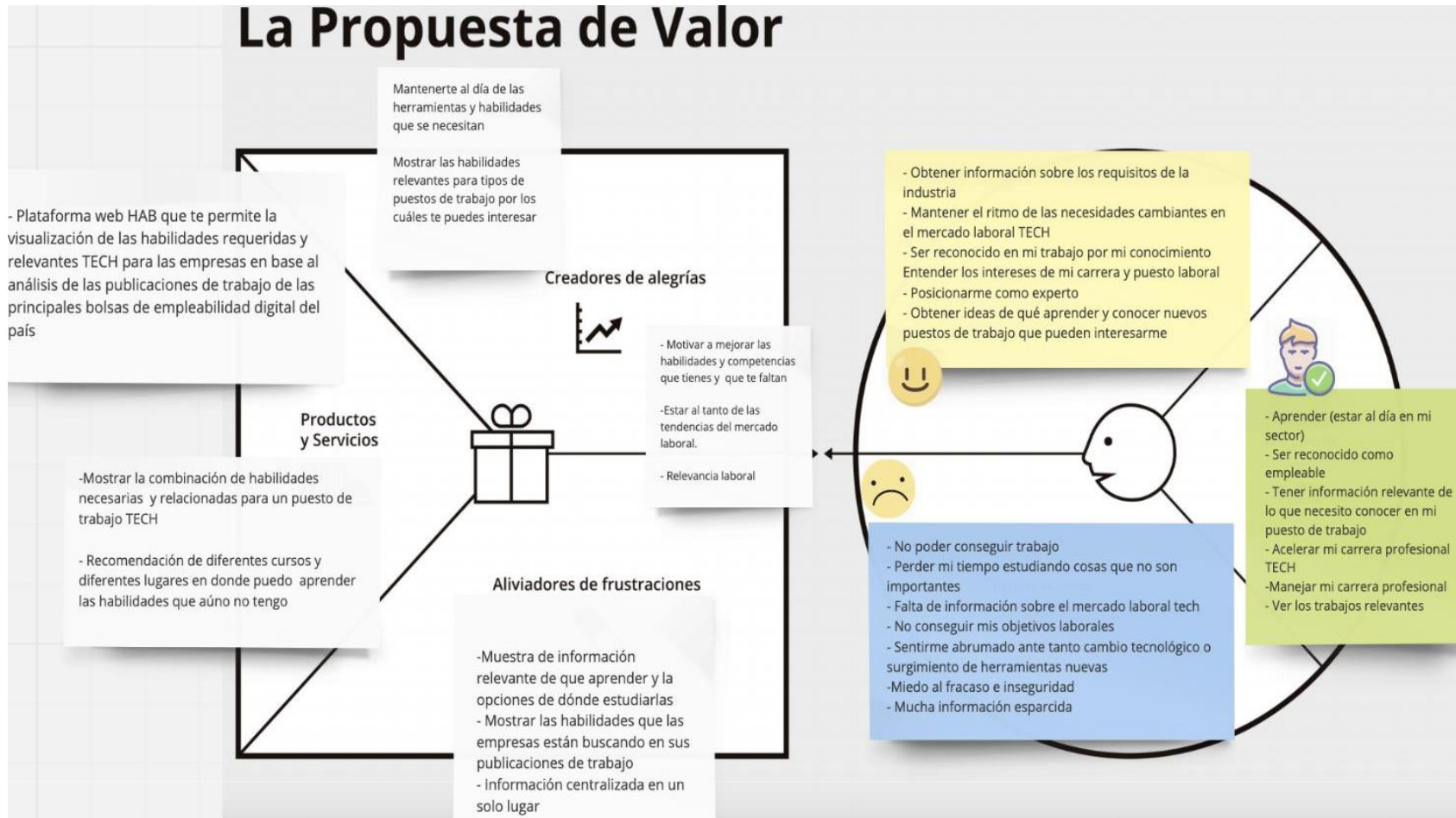
4.1 Diseño de propuesta de valor

Se diseñó una propuesta de valor que cumpla con las necesidades específicas de cara a los usuarios que utilizarán la plataforma. Además, se definieron las características del segmento de clientes y el mapa de valor, donde se especifica cómo se pretende crear valor con nuestra solución para ellos. Así como las funcionalidades que trabajan sobre los pain-points de los usuarios e incluye como se soluciona.

Como se puede visualizar en la figura 4.1 el modelo del lienzo que consta “Mapa de valor” y “Perfil del cliente” y “Encaje” para visualizar los beneficios que el proyecto propuesto ofrece al nuestro de cliente

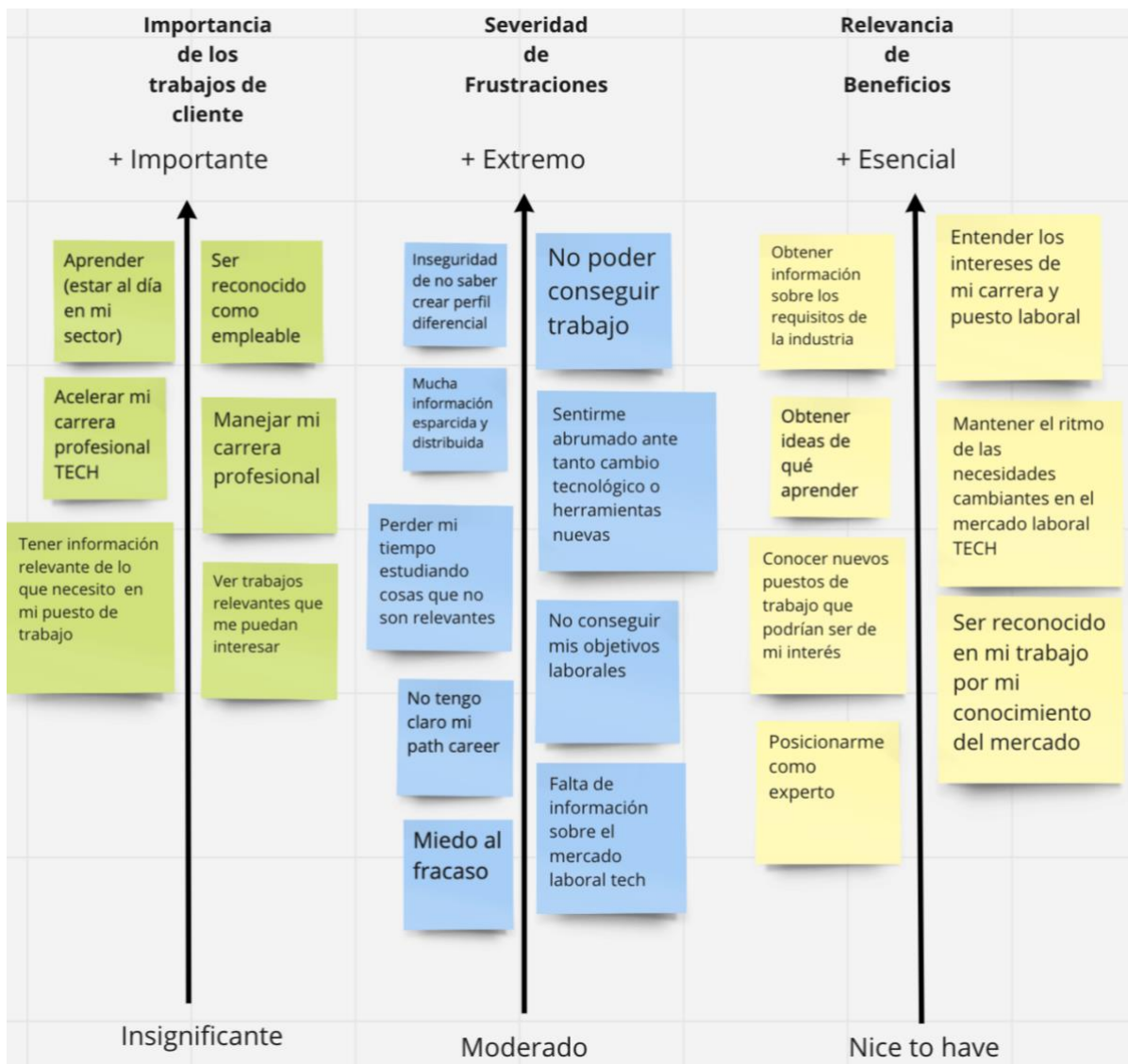


Figura 4.1
Definición de la propuesta de valor



Luego se prioriza las necesidades de los perfiles de cliente por importancia, severidad de frustración o cuán esencial o relevante son sus beneficios, ver figura 4.

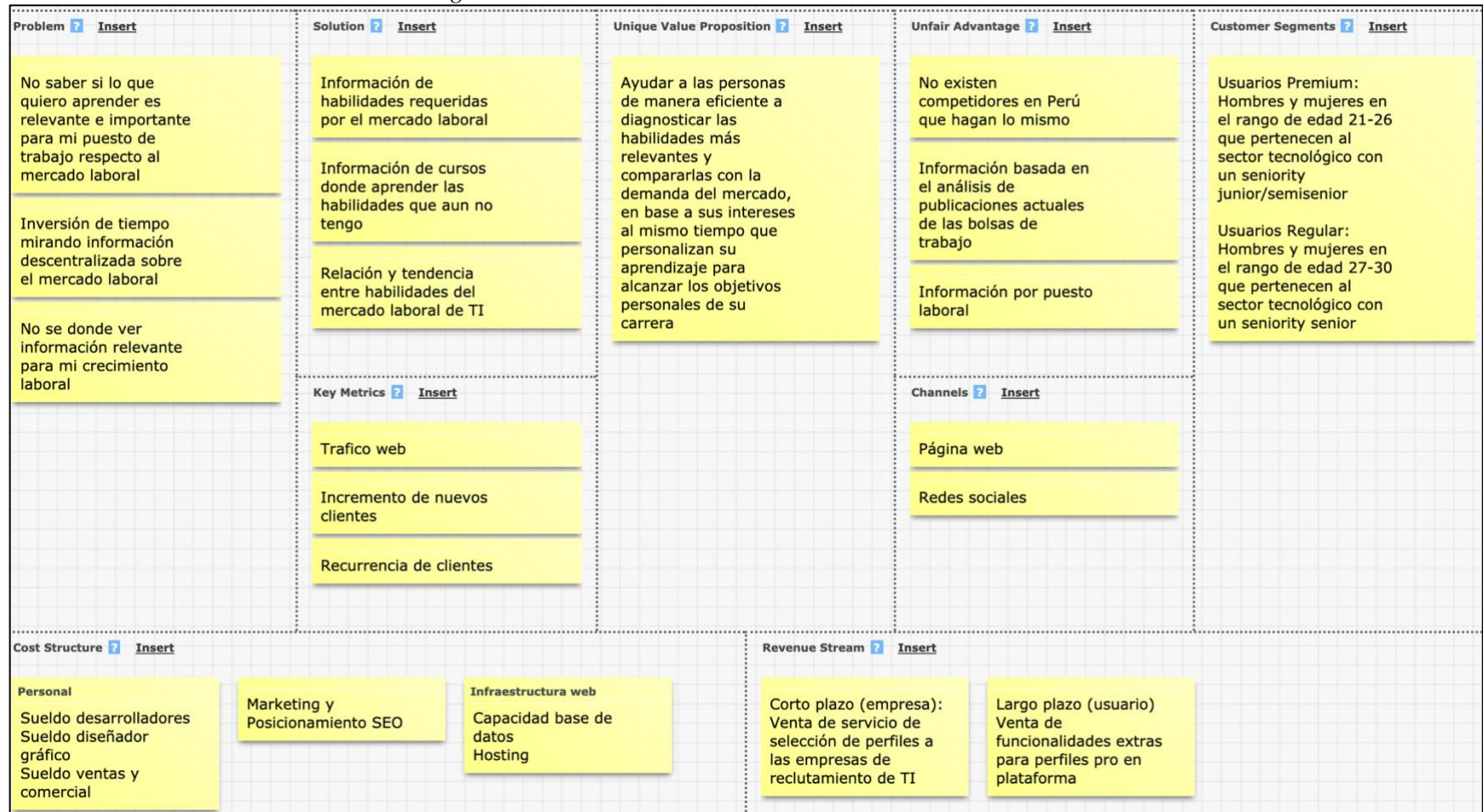
Figura 4.2
Priorización de necesidades del perfil de cliente



En base a esto se generó la versión del Lean Canvas en base al modelo de negocio propuesto, ver figura 4.3



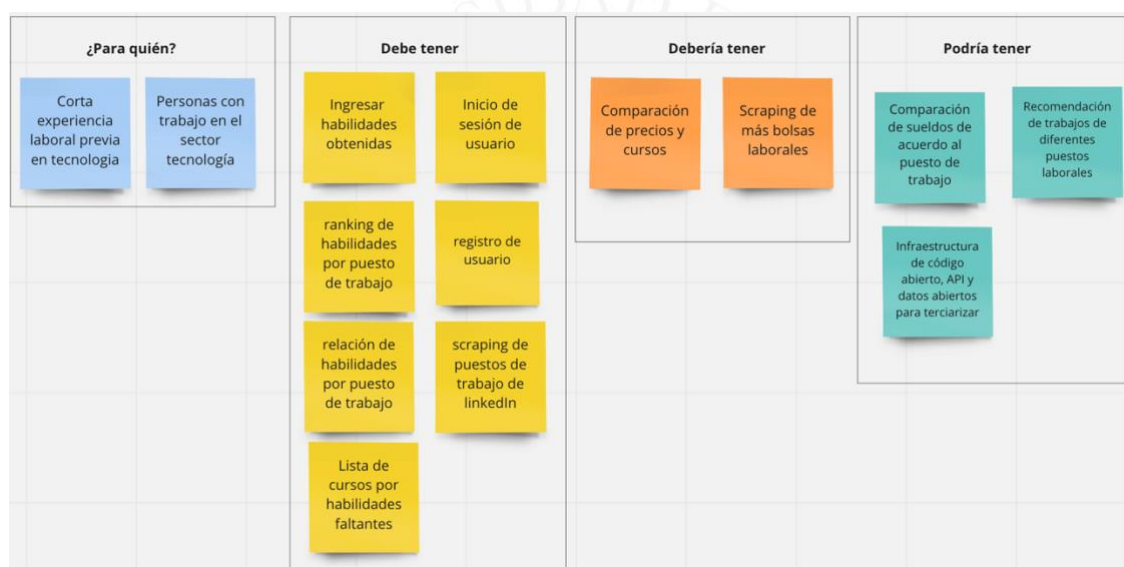
Figura 4.3
Lean Canvas actualizado del modelo de negocio



El modelo de negocio final se da por la unión del valor generado por los usuarios que ingresarán sus perfiles en la web y las habilidades que ya tiene y las que aún deben ser mejorada y la utilización de la información de estos perfiles para poder generar la lista de candidatos más óptimos para las empresas y el pago de una cuota por esta información, la cual se detalla en la viabilidad económica.

A partir de ello se definió el mínimo producto viable (MVP) y las funcionalidades que debería tener el mismo, como se visualiza en la Figura 4.4

Figura 4.4
Definición de MVP de HAB



4.2 Alcance

El alcance del proyecto es una plataforma donde los usuarios (profesionales de TI entre 22 y 30 años) al registrarse podrán generar su perfil y así poder visualizar un dashboard, el cual tendrá la lista de las herramientas/habilidades técnicas más demandadas así como entre cuales se relacionan de acuerdo al puesto ingresado en base a los descripciones de trabajo de la plataforma LinkedIn que pasaron por el proceso de scraping y ETL (previamente explicados en el capítulo III), como parte adicional a las habilidades el

usuario puede ver los cursos previamente cargados por habilidad acerca de las plataformas donde puede aprenderlas .

4.3 Objetivos generales y específicos

4.3.1 Objetivo general

Brindar una visión general eficiente a las personas que trabajan en el rubro tecnológico de las habilidades técnicas que son potencialmente demandadas por la mayoría de empresas según un puesto de trabajo determinado para estar actualizadas sobre las necesidades cambiantes del mercado laboral .

4.3.2 Objetivos específicos

- Mapeo de habilidades tecnológicas a partir de datos abiertos mediante el uso de lenguaje procesamiento natural .
- Recomendación de lugares o centros de capacitación donde adquirir las habilidades que no cumplen las personas que buscan cierto puesto de interés.
- Identificar habilidades de trabajo que más buscadas por los empleadores, para ayudar a las personas a encontrar trabajos que se ajusten mejor a ellos respecto al agrupamiento de sus habilidades técnicas.
- Analizar e informar el estado del mercado laboral para la toma de decisiones de planificación y desarrollo profesional.

4.4 Modelo de negocio propuesto

4.4.1 Modelo de negocio

El objetivo del modelo de negocio es describir que la plataforma de nombre Hab puede resolver y aliviar los problemas descritos anteriormente ofreciendo un valor agregado al cliente de manera efectiva con las siguientes características:

- Novedad:

Es una plataforma que no existe en Perú, no se cuenta con alguna herramienta expresamente orientada a aprovechar los datos del sector laboral en convertirlo en una oportunidad para cada persona, proporcionando una visibilidad completa y precisa de las publicaciones de trabajo utilizado por cualquier persona, la cual esté motivada y en mantener la competitividad y la empleabilidad a través del aprendizaje permanente.

- Precio y Mercado:

Usuarios: El precio inicial para las personas que se registren en HAB será gratuito, para generar un vínculo de soporte entre los usuarios, no se pretende vender el servicio desde el comienzo, sino que se adoptará un modelo de negocio en donde se regalan el servicio básico.

El modelo de negocio de cara al usuario se basará en el modelo Freemium, es decir que se recibe el producto de manera gratuita, pero con la opción de adquirir una versión con más funcionalidades por una suscripción a la plataforma, como por ejemplo generación de perfiles pro, acceder a descuentos especiales en cursos, capacitaciones o conferencias, en base al costo de afiliación.

Empresas: Las empresas harán un pago para el acceso a los perfiles de acuerdo con los roles que estén necesitando en su proceso de reclutación tecnológica.

Se manejará una economía de escalas, esto quiere decir que mientras más usuarios se registren en la plataforma esto aportará más valor en la comunidad y así crear viralidad.

- **Utilidad:** Enfatizar el diseño y usabilidad como elemento diferenciador para obtener rápida adaptación y familiaridad del usuario.
- **Promoción:** Por posicionamiento orgánico y pagado tanto en medios de búsqueda y redes sociales. Se necesita empuje por medio de los early adopter (primeros usuarios) e ir acaparando posicionamiento.
- **Estrategia:** La estrategia consiste en utilizar la tecnología y la innovación tecnológica como principal factor en el desarrollo de una plataforma digital que permita lograr a las personas tener acceso a información relevante

En el corto plazo (primer año)

Implementación de la página web gratuita de cara a los usuarios que le da visibilidad de las habilidades que son relevantes para su desarrollo profesional y las oportunidades que tiene para poder desarrollar las que no posee.

A mediano plazo (a partir del segundo año)

Ampliar el ecosistema inicial para poder validar las habilidades técnicas que poseen por medio de retos y dar visibilidad a las empresas de que los poseen y a qué nivel.

Visión a largo plazo

Generar un ecosistema completo de acompañamiento laboral para las personas del rubro tecnológico y su desarrollo profesional, así como su aprendizaje continuo. Lograr alianzas con entidades para generar ayudas financieras para las capacitaciones o adquisición de cursos por medio de la herramienta para poder potenciar sus habilidades en diferentes rubros

4.4.2 Roles de equipo de proyecto

La organización del recurso humano se está proponiendo los siguientes roles y responsabilidades ver tabla 4.1

Tabla 4.1

Roles y responsabilidades

Rol	Responsabilidades
CEO	Liderar el desarrollo de la estrategia de la empresa
Programador Backend/FrontEnd	Encargados de convertir las especificaciones utilizando lenguajes de programación, así como herramientas de software de apoyo a la programación y de análisis de data
Diseñador y UX	El encargado de generar el diseño de la plataforma e interacción de la misma.

4.4.3 Stack Tecnológico

Para el desarrollo de la plataforma se ha tomado en cuenta los siguientes componentes:

Tabla 4.2

Descripción Stack Tecnológico

Componente Backend	Descripción
Ms Scrapping Web Scraping y ETL	Proceso de Scraping y Transformación de Datos en lenguaje Python con librerías Scipy, NLTK y Jarowinkler
Desarrollo de servicios	Desarrollo con framework loopback en lenguaje Node Js
Componente Frontend	Descripción
Desarrollo de Interfaces	Desarrollo en lenguaje Javascript, html y css
Google Cloud - Google Kubernetes Engine	Servicio donde se alojarán los microservicios de la aplicación

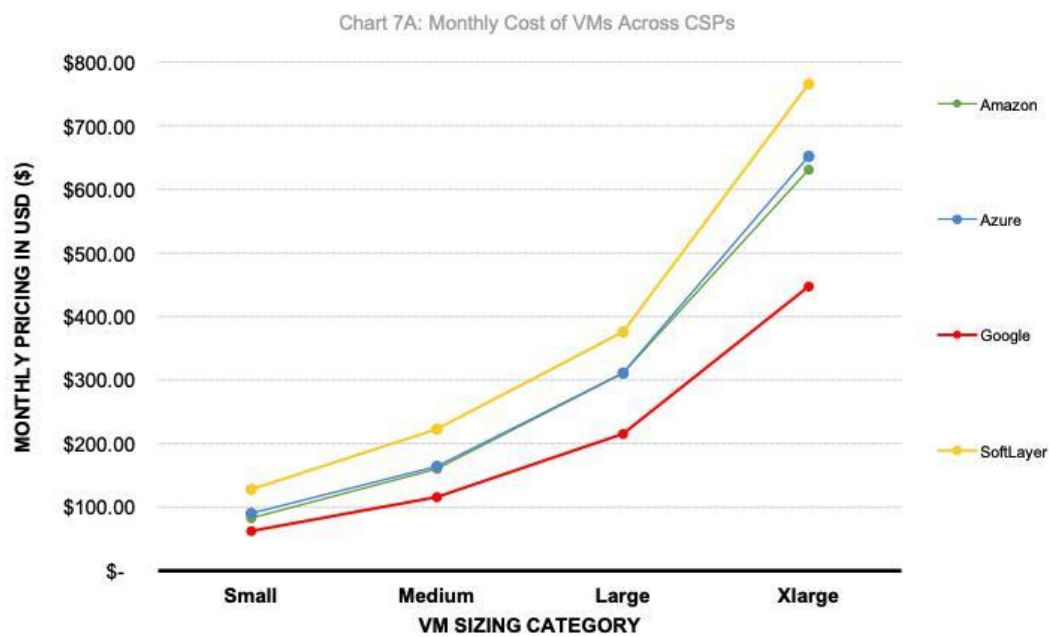
La administración de toda la infraestructura se encuentra automatizada con el framework de infraestructura como código Terraform, lo cual permite crear, actualizar y eliminar todos los componentes de la infraestructura.

La arquitectura de la plataforma está orientada a microservicios y soportada en una infraestructura cloud (Google Cloud) basada en contenedores y orquestada con “kubernetes” para permitir el escalado y la réplica a medida que el proyecto va teniendo un crecimiento escalonado tanto vertical como horizontal.

Se decidió alojar la plataforma en Google Cloud Platform para el prototipo debido a que tiene una prueba gratuita de 12 meses con \$300 de crédito, la prueba gratuita te ofrece un bono con el que pagar los recursos que usas a medida. Así mismo , para la evaluación se hizo en comparación con otras soluciones de nubes, dicho análisis se basa en el costo beneficio detallado de proveedores de nube, como Google Cloud Platform, Amazon, Azure. El informe detalla que Google Cloud es el proveedor menos costoso (Cloud Spectator , 2018). Ver figura 4.5 .

Además, se vio beneficioso utilizar el stack de Google para el servicio de base de datos como servicio , que cumple con los requisitos de disponibilidad y fácil administración y “Kubernetes” para la orquestación de contenedores. Debido a que los mismos se desarrollan internamente dentro de Google, y se integra con otros servicios de Google y su curva de aprendizaje es más corta. (Ver anexo 7)

Figura 4.5 Comparación de tarifas de Cloud Computing



Nota: De “Cloud Performance report , por Cloud Spectator”, 2018 (<https://cloudspectator.com/cloud-performance-reports>)

Figura 4.6
Diagrama de Infraestructura de aplicación

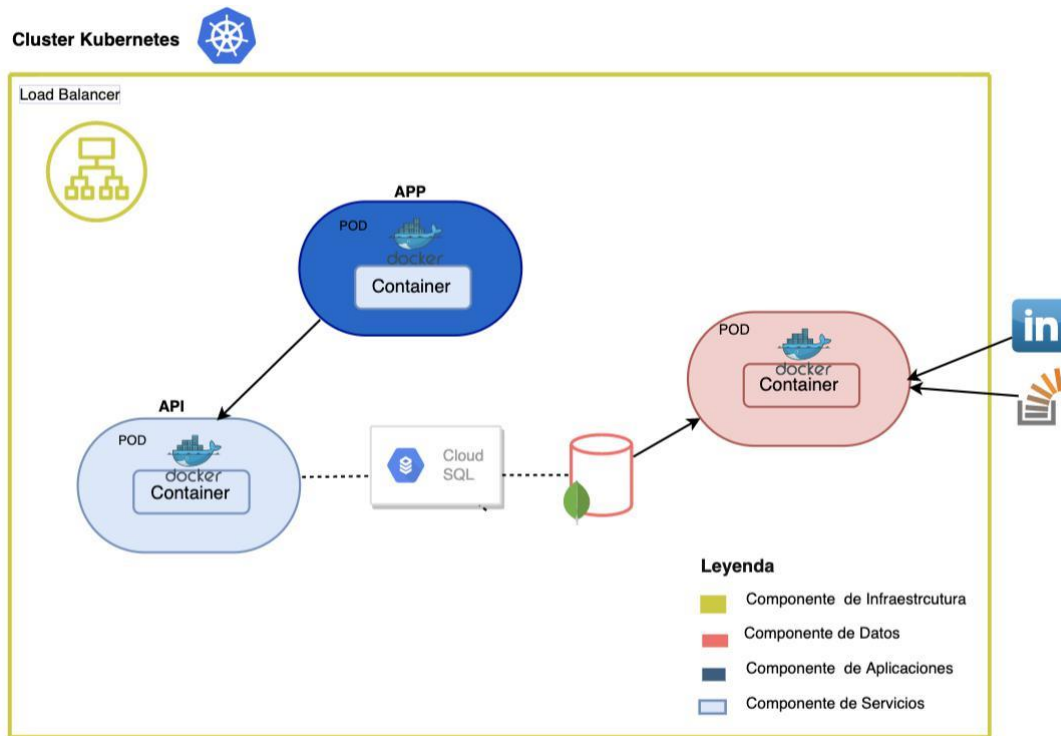


Figura 4.7
Diagrama de Despliegue UML

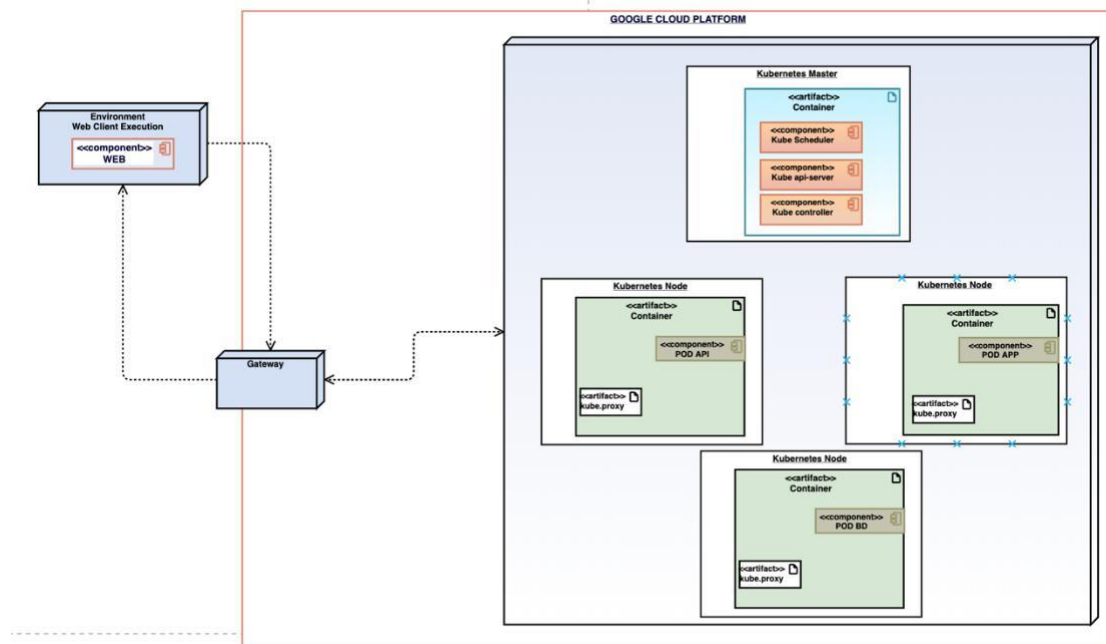


Figura 4.8
Diagrama de Componentes UML

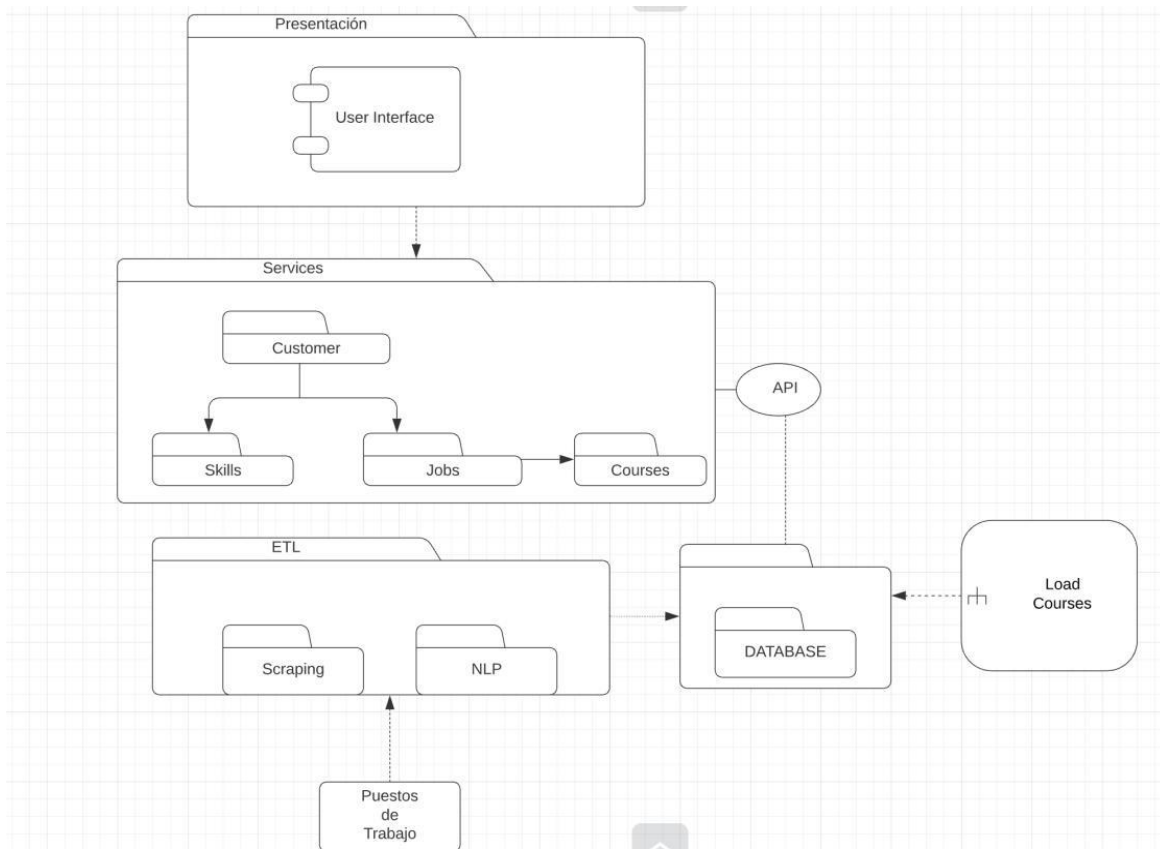
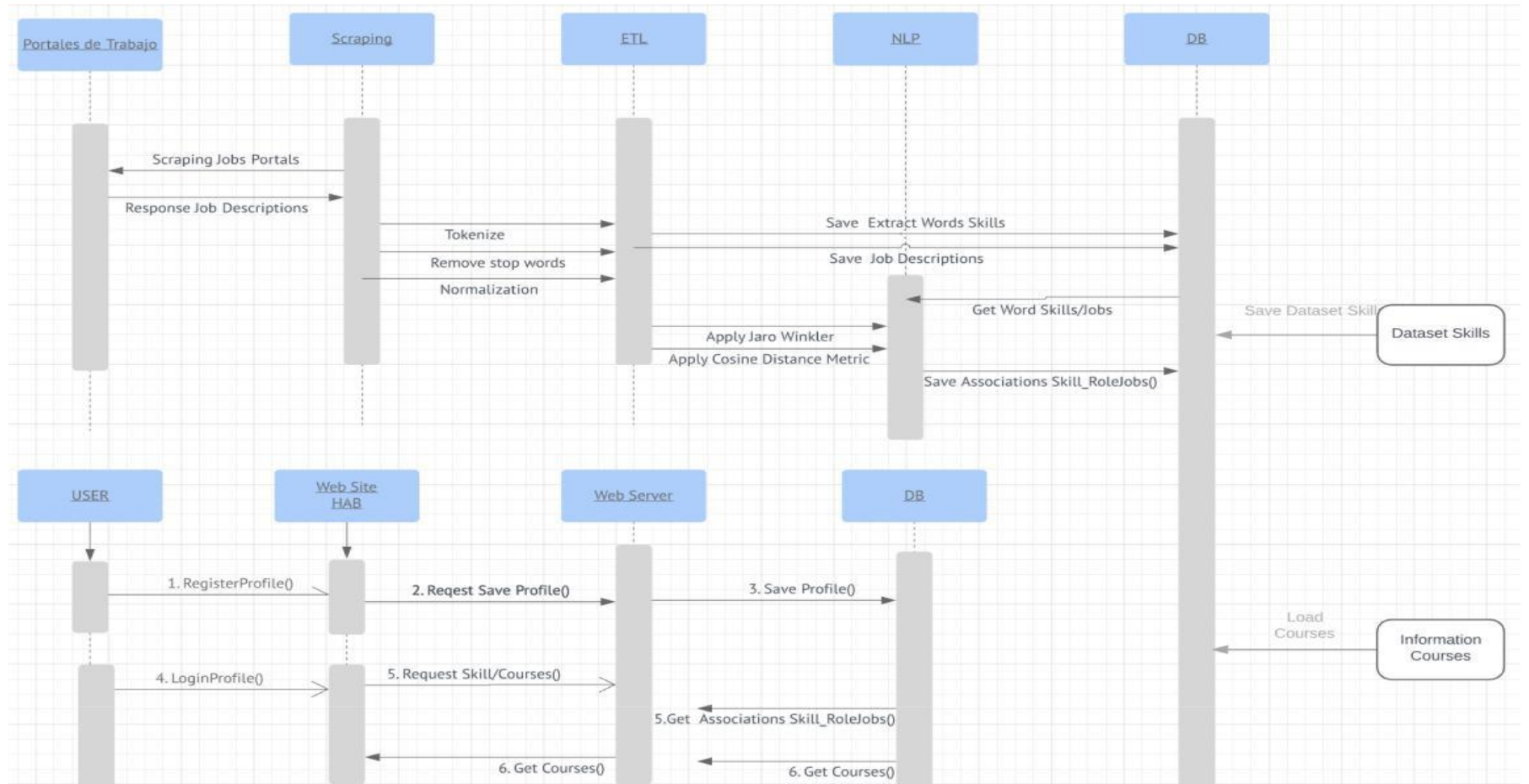


Figura 4.9
Diagrama de Flujo UML

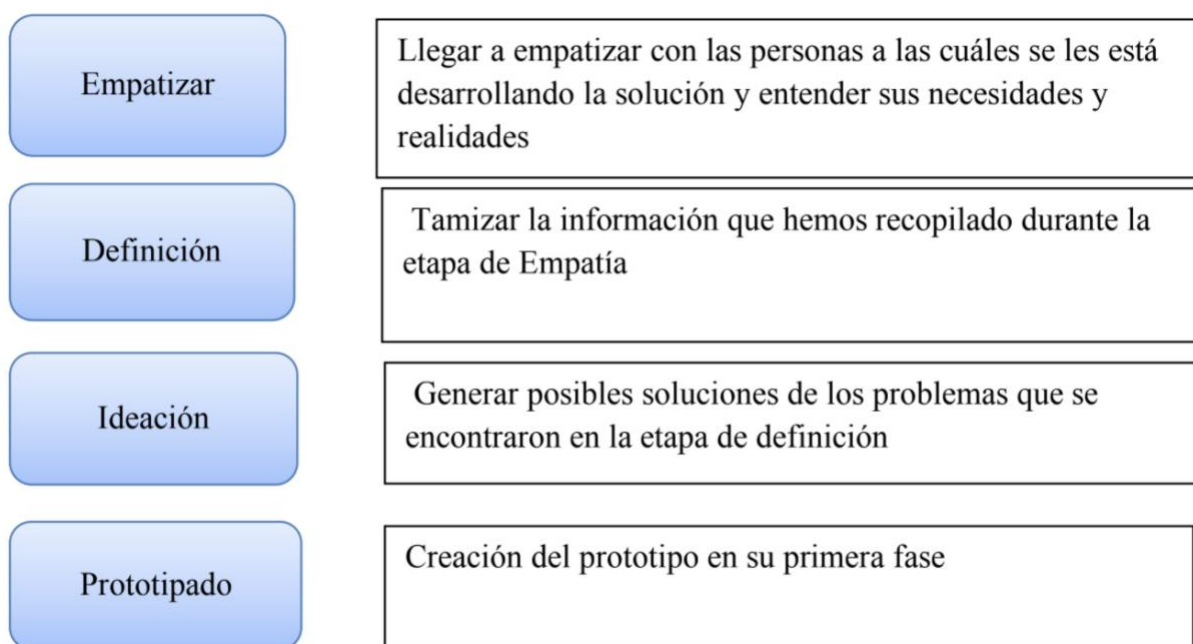


CAPITULO V : DESARROLLO DEL PROTOTIPO

Para el desarrollo de este proyecto se utilizó la metodología de design thinking. Las fases y los objetivos de cada fase de esta metodología se muestran a continuación:

Figura 5.1

Objetivos por etapas del proceso design thinking



5.1 Empatizar

Se creó un plan para poder realizar un seguimiento del propósito, es decir, descubrir cómo nuestro grupo objetivo se considera respecto a ¿Cómo mejorar sus habilidades técnicas en el sector tecnológico? ¿Cómo consiguen ser personas empleables? o si ¿Sienten que tienen acceso a recomendaciones personalizadas de capacitación superior orientadas al mercado? Así podríamos conocer las características nuestro grupo objetivo.

5.1.1 Propósito

- **Objetivo general de las entrevistas realizadas:**

Identificar las necesidades por las cuales los usuarios pasan al obtener nuevas habilidades técnicas en base a las expectativas que pueden encontrar en el mercado laboral para ser considerados adecuados y calificados.

- **Objetivos específicos de las entrevistas realizadas:**

- Identificar las principales motivaciones de las personas para aprender alguna nueva habilidad técnica.
- Comprender las limitaciones que han encontrado en el proceso de conocer qué nuevas competencias adquirir.
- Entender las causas por las que las personas puedan ser considerados o no empleables.
- Conocer, los inconvenientes que han encontrado las personas al intentar buscar un nuevo puesto de trabajo de acuerdo a las habilidades técnicas que poseen.
- Analizar la respuesta de los entrevistados.

5.1.2 Método

Se realizaron 20 entrevistas de acuerdo a una guía establecida (ver anexo 2) con personas del sector tecnológico y se pidió que describan las experiencias en su proceso de empleabilidad y aprendizaje de habilidades técnicas por los que han pasado.

Instrucciones: Las preguntas están enfocadas en 4 partes :

Parte 1 Ruta de habilidades: Conocer al usuario a nivel personal, profesional y saber cómo mantiene a sus competencias.

Parte 2 Perfiles diferenciales: Comprender las motivaciones y los factores que influyen en aprender algo nuevo que es relevante para el mercado.

Parte 3 Aceleración de carrera: Conocer sobre cuáles son las características principales en su toma de decisiones para o potenciar su perfil laboral.

Parte 4: Finalmente indagar sobre las expectativas que tiene el usuario con respecto al acompañamiento en el desarrollo de su habilidad

5.1.3 Participantes

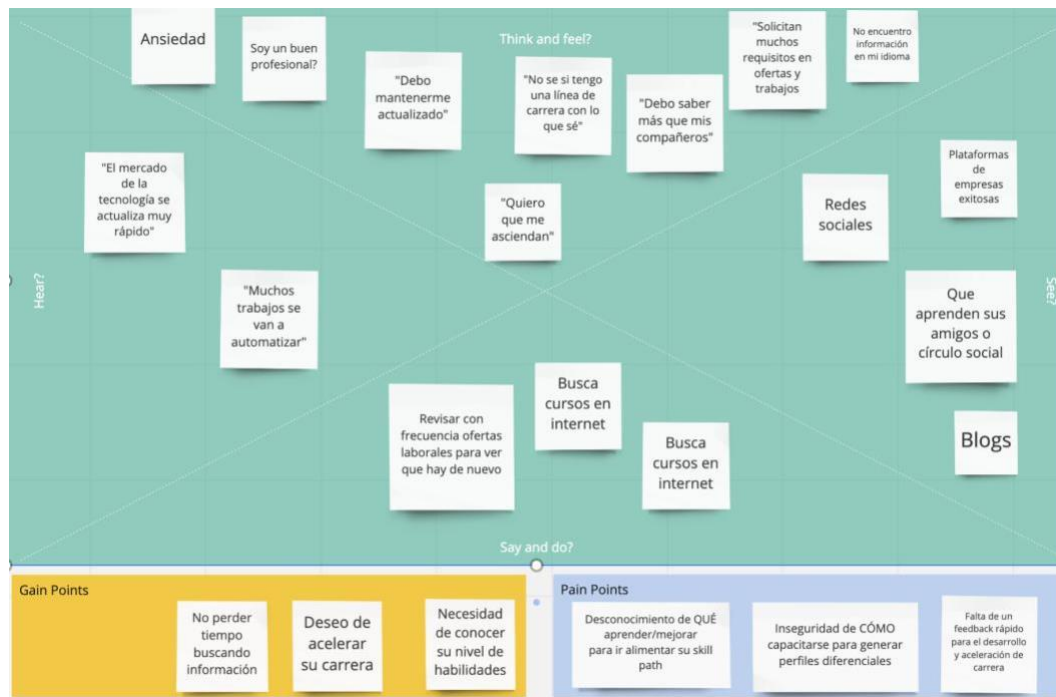
Los 20 participantes fueron entre personas de entre 21 y 30 años, que trabajan en el sector tecnología o en área de TI.

5.1.4 Ubicación

Las entrevistas a las 20 personas se dieron en áreas fuera de su entorno de trabajo de manera que se los participantes se sienten cómodos y pueden expresarse.

Las respuestas fueron transcritas a post-its y luego se agruparon de acuerdo con su similitud o si definían una misma idea y así encontrar patrones que resuman lo que el usuario estaba sintiendo o pensando y que sea representativo. Luego se representaron los posts it para crear el mapa de empatía en una herramienta online llamada Miró, siendo el resultado final en mapa de la figura 5.2

Figura 5.2
Mapa de empatía



5.2 Definición

Se encontró un patrón de 2 perfiles de usuarios, los cuáles se detallan a continuación en las tablas 5.1, 5.2 y figura 5.3

Tabla 5.1
Definición usuario - necesidad

Usuario	Necesidad	Insight
Una persona que trabaja en el rubro de tecnología entre 22-25 años de edad	Ayuda para encontrar de manera fácil y centralizada las habilidades que necesita y que demanda el mercado laboral	Las personas invierten tiempo buscando ofertas de trabajo del extranjero para anticipar qué habilidades o nuevas herramientas deben aprender y sentirse empleables.

Tabla 5.2*Definición de usuario - necesidad*

Usuario	Necesidad	Insight
Una persona que trabaja en el rubro de tecnología entre 26-30 años .	Tener un diagnóstico de las habilidades que tiene y cuáles son importantes y relevantes actualmente en el para tener una referencia de su valor de empleabilidad en el mercado laboral.	Las personas cuando se ven ante el escenario de tener que aprender algo (de manera reactiva) empiezan a aprender las habilidades solicitadas. Tienen un aprendizaje informal "en el trabajo" y eso significa que muchos tienen habilidades que no han reconocido. Para que los usuarios obtengan el tipo correcto de capacitación, necesitan identificar qué capacidades ya tienen y comprender qué habilidades son importantes, pero aún no las ha adquirido.



Figura 5.3
User Journey Map



5.3 Ideación

Se realizó una lluvia de ideas o brainstorming de acuerdo con el descubrimiento de los insights previamente explicados, para plantear nuestra solución y también con la ayuda de referencias para poder conocer cómo otras empresas resolvieron el problema, se investigaron páginas con objetivos similares como se muestra en la figura 5.6 el portal Job Tech

Figura 5.4
Brainstorming y clusterización de ideas en herramienta Miró

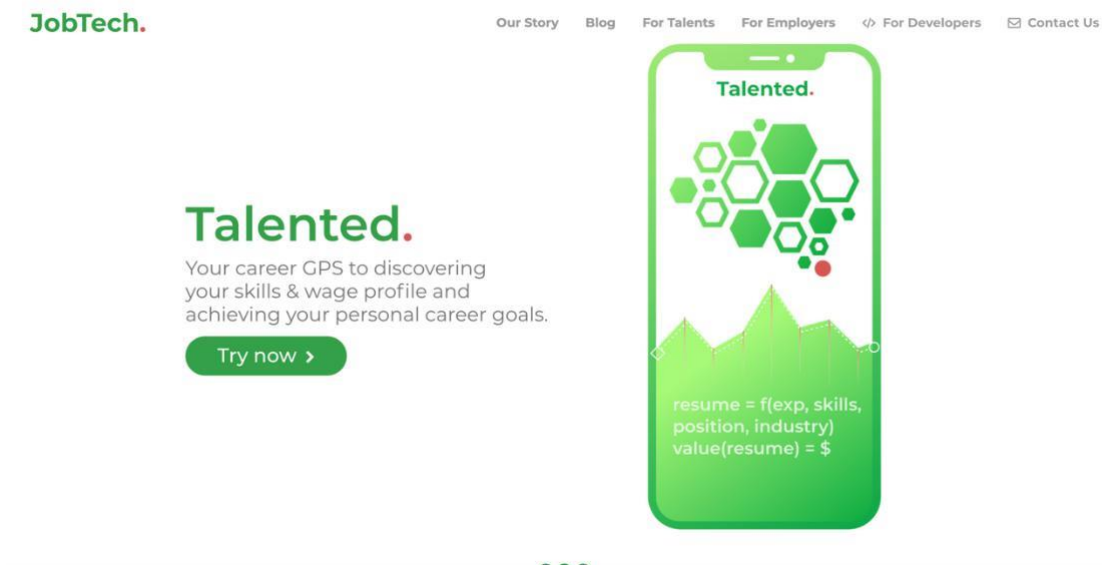


5.4 Prototipado

En esta etapa se realizó la creación del prototipo en su primera fase luego de la definición de lo que el producto debería, tener realizado en el paso de ideación.

Figura 5.5

Página Job Tech de referencia

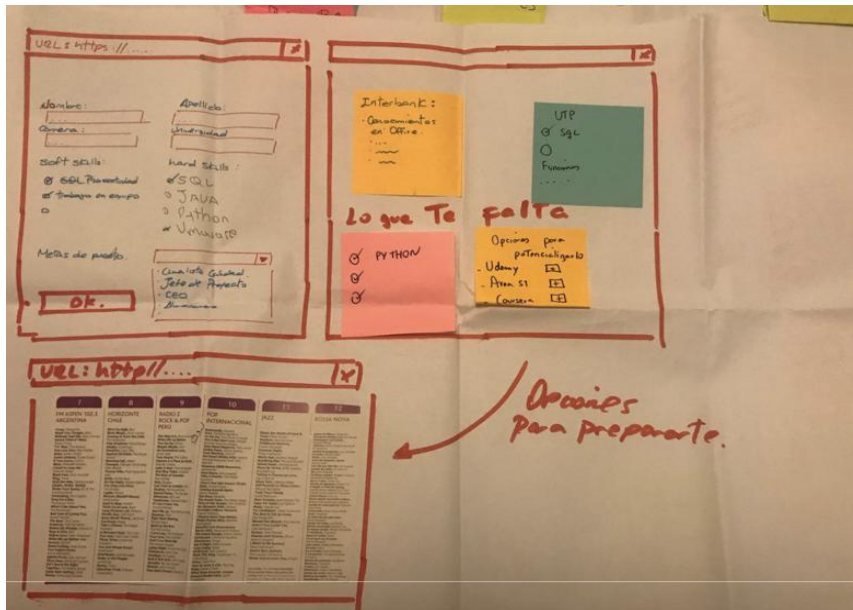


Nota: De “Talented” , por JobTech, 2019 (<https://jobtech.co>)

Finalmente se tuvo una primera versión del prototipo con las principales funcionalidades que debería tener, entre las cuales están permitir una visión de las principales herramientas técnicas que actualmente se demanda en el mercado.

Figura 5.6

Versión 0 a mano alzada



En la versión 1, se decidió hacer el prototipo, los colores utilizados son en tonalidad azul ya que como se explica en Bathi (2018) el azul es solicitado en el 59% de los logos de tecnología y en el 61% de los logos de las compañías líderes del sector.

Las pantallas fueron las siguientes

- Pantalla de inicio bienvenida donde muestra el nombre de la plataforma “HAB” y una frase sobre la propuesta de valor de la plataforma
- La pantalla de registro de datos personales que consta de nombre, email y password.
- La pantalla de login con email y password solamente para el acceso simple y fácil de los usuarios.
- La pantalla de registro de habilidades técnicas y el puesto de trabajo actual.
- La pantalla de visualización de ranking habilidades técnicas.

Figura 5.7

Versión 1: Pantalla de inicio, login y registro en herramienta “Marvel”

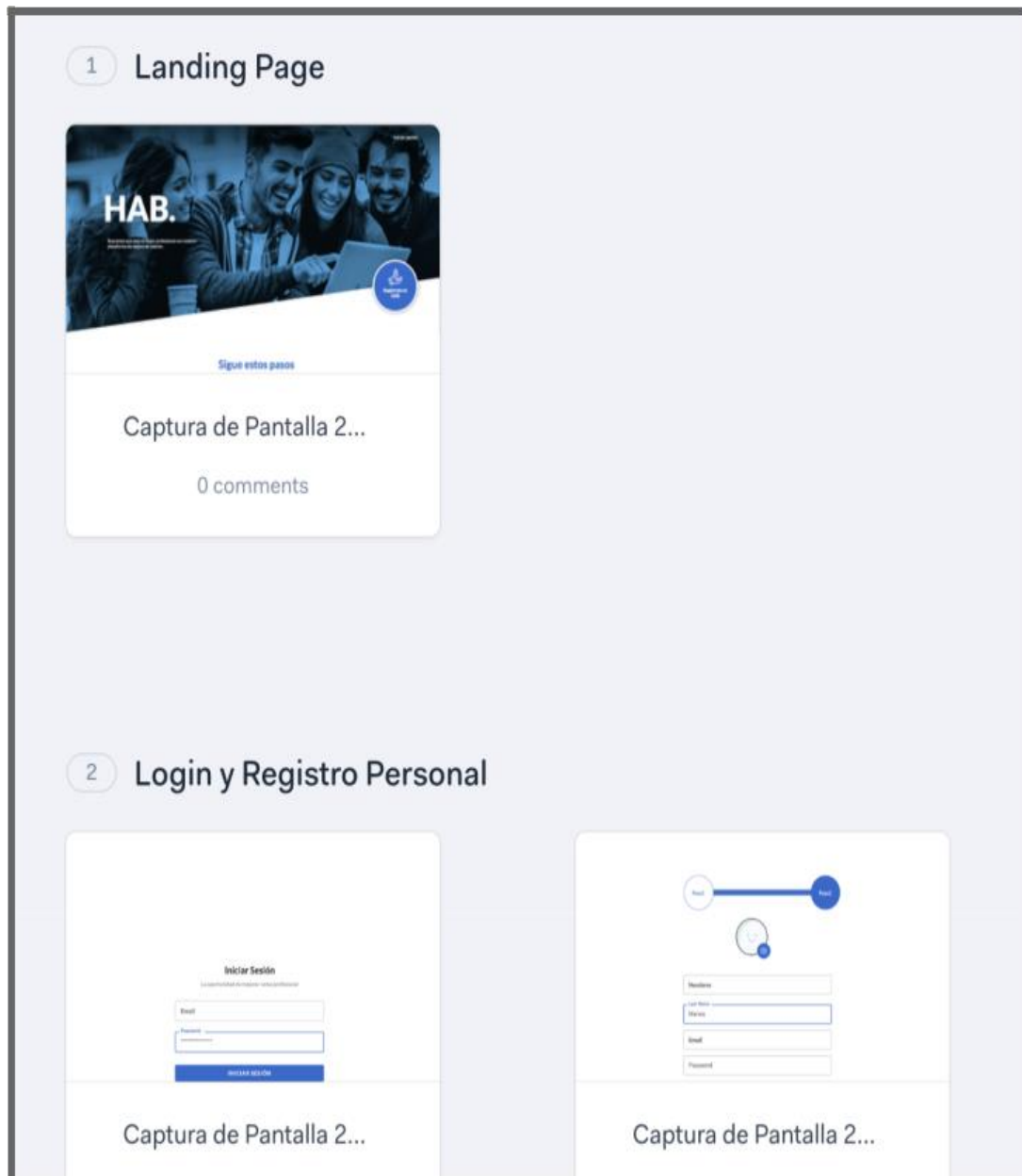
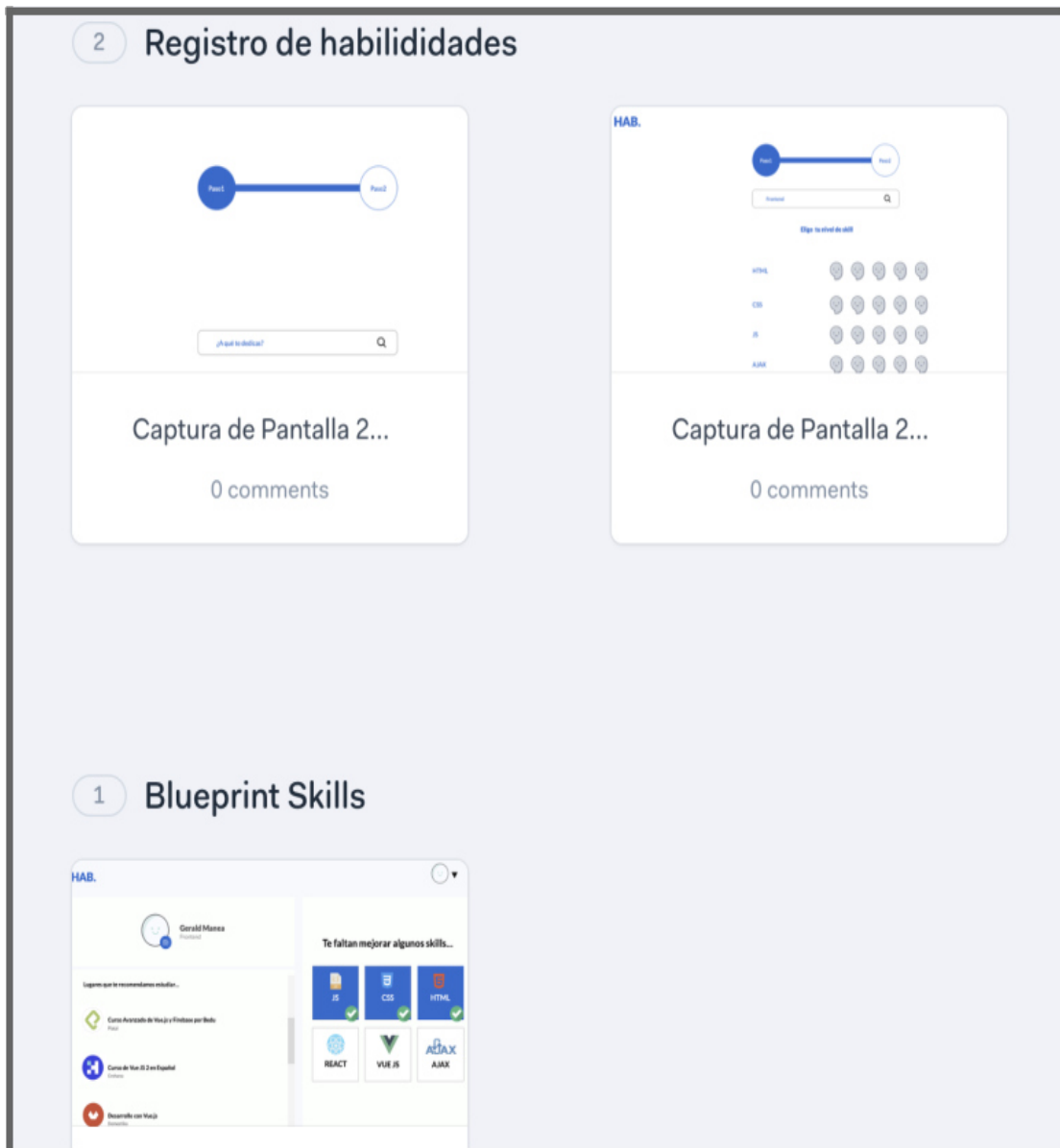


Figura 5.8

Versión 1: Pantalla habilidades y cursos en herramienta “Marvel”



5.5 Testeo

5.5.1 User Test Versión 1

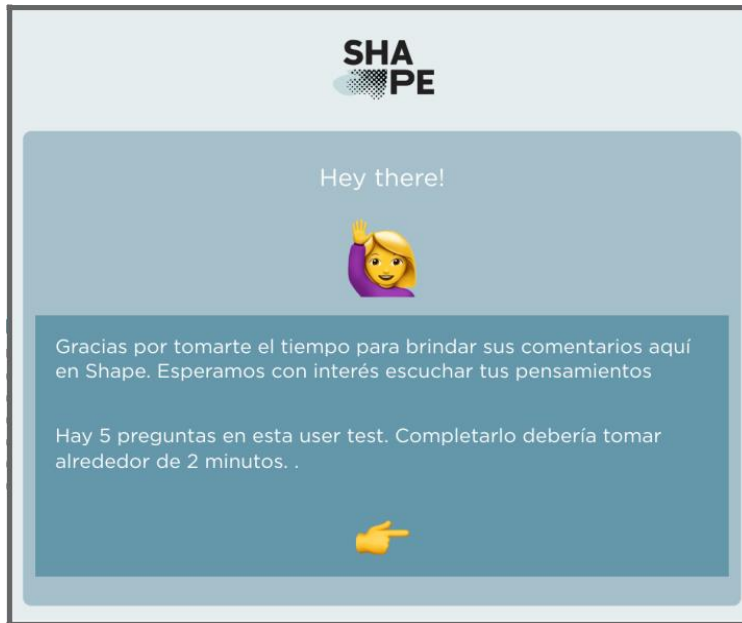
Se realizó un user test a 8 personas, a quienes se le mostró los primeros prototipos de la solución que abarcaban la pantalla de bienvenida, pantalla de login, pantalla de registro, pantalla de visualización de datos, para llevar a cabo se utilizó la herramienta Shape , que es una herramienta colaborativa desarrollada por IDEO, la cual permitió la recopilación de comentarios directamente de los usuarios para una retroalimentación efectiva.

Se elaboró un cuestionario por pantalla que tenían opciones múltiples de respuesta, así como preguntas abiertas, para conocer en base a las pantallas mostradas, el interés que generaba en los usuarios, así como validar si la idea que se trataba transmitir estaba clara

La primera pantalla que se mostró era la de bienvenida y se les preguntó a los usuarios acerca de la claridad de la idea que se llevaban al ver la página de bienvenida, tuvieron 4 opciones que representados por emojis para las opciones de muy claro, claro, algo claro, y nada claro. Así mismo se les dejó una pregunta abierta para que expresaran la idea que tenían acerca de lo presentado. Ver figura 5.10

Figura 5.9

User Test en la herramienta Shape



Nota: Herramienta libre para la generación de user test

Las siguientes preguntas estuvieron relacionadas sobre la propuesta de valor que la página deseaba brindar, así como cuán usable y emocionante era para ellos, donde tuvieron 4 opciones de respuesta: Muy usable, usable, poco usable, nada usable, ver figura 5.10

Figura 5.10

Pregunta 1 y 2 del user test versión 1 realizada en la herramienta Shape



Nota: Herramienta libre para la generación de user test

La última pregunta, fue abierta para poder recopilar mejoras o nuevas funcionalidades esperaban los usuarios de la plataforma, ver figura 5.11

Figura 5.11

Pregunta 3 y 4 del user test versión 1 realizado en la herramienta Shape

Una plataforma web que te muestre los skills mas relevantes de tecnología de acuerdo al puesto de trabajo de tu interés basado en el análisis de las ofertas de trabajo de las principales bolsas de trabajo online

Cúan emocionante es esta idea para ti?

Selecciona tu respuesta abajo

🤔 😐 😊 😍

Cúan usable es esta idea par ti

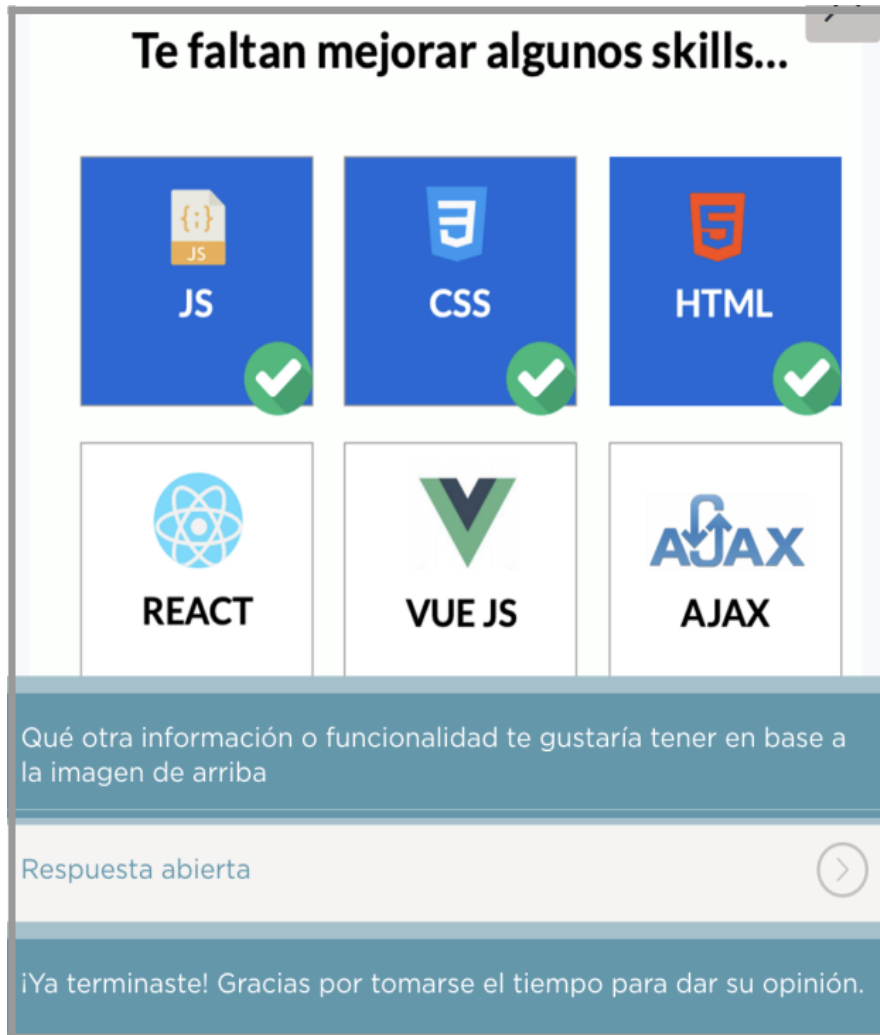
Selecciona tu respuesta abajo

👎 👏 👍 👊

Nota: Herramienta libre para la generación de user test

Figura 5.12

Pregunta 5 del test de usuario versión 1 realizada en la herramienta Shape



Nota: Herramienta libre para la generación de user test

Los resultados encontrados en base a las respuestas en el user test, mostraron información sobre el interés de las personas al presentarles la idea y sobre si la idea mostrada se transmitía de manera clara, es decir si entendían al ingresar a la página inicial de qué trataba la plataforma o tenían dudas (ver figura 5.15)

Sin embargo, al evaluar la pantalla inicial (landing page) sin hacer ninguna explicación a los usuarios, un 45% y 27% de usuarios respondió que no le dejaba claro o no se dejaba atender de lo que se trataba la página web. Así mismo, al responder cada uno de los 8 usuarios sobre cuál era la percepción que tenían, (ver figura 5.14). La mayoría infirió que era sobre cursos online. Al explicarles a los usuarios sobre una plataforma que tiene como propuesta, brindarles la oportunidad de crearse un perfil donde registren las habilidades que tienen adquiridas para así contrarrestar y mostrarles

las habilidades más requeridas en base al análisis de las descripciones de trabajo de los portales de trabajo como LinkedIn de igual manera mostrarle las opciones de donde podrían aprenderlos para potenciar su perfil de empleabilidad, más del 50% de usuarios respondió que le parecía muy emocionante la existencia de una plataforma como HAB, así como útil.

Figura 5.13

Respuestas de los 8 usuarios a la pregunta abierta sobre la percepción que tenían del propósito de la página respecto a las pantallas mostradas realizada en el user test versión 1

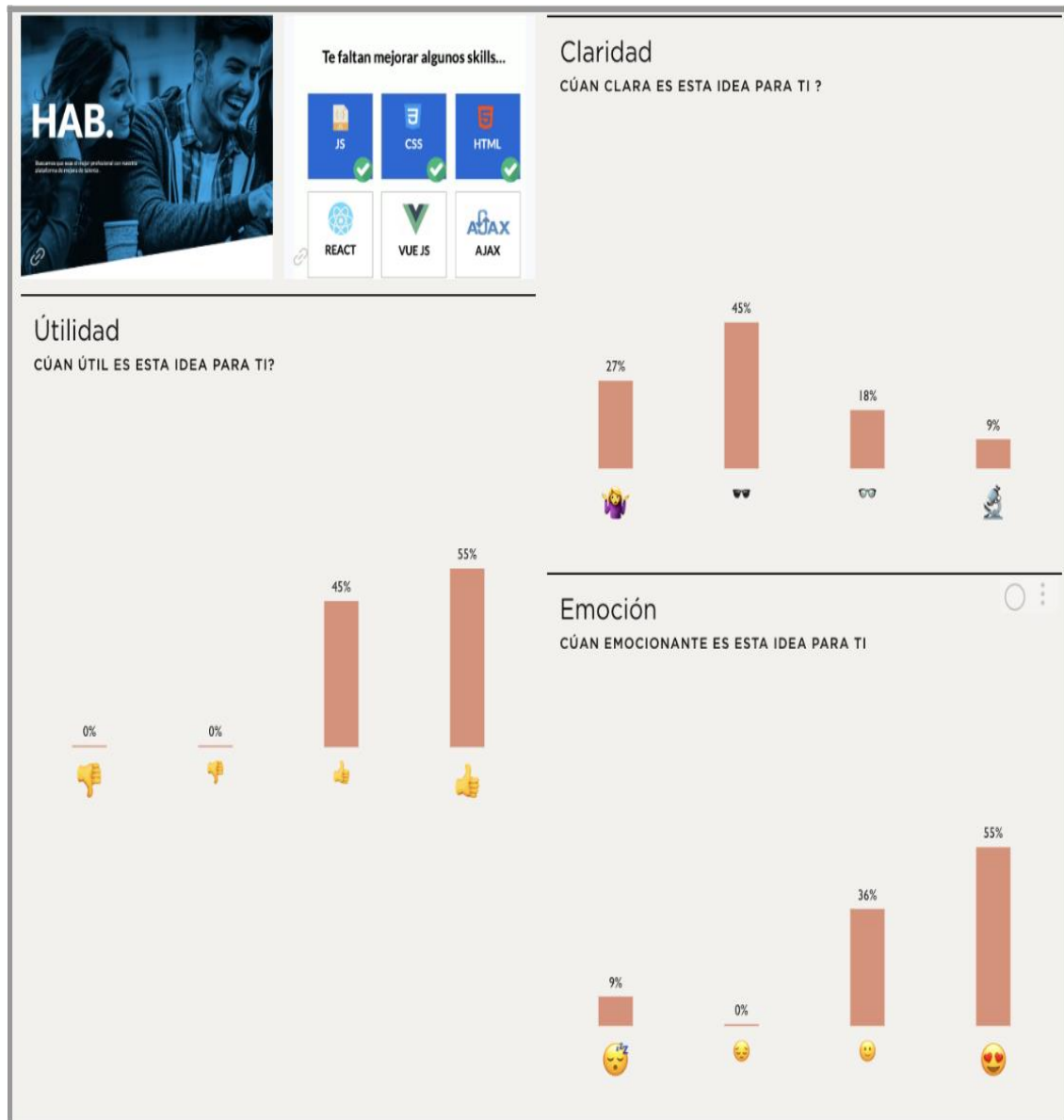
¿Qué producto y/o servicio crees que te ofrece esta web al ver la imagen de arriba? Responses			
Enseñanza virtual de cursos o habilidades blandas	Couching de carrera	Cursos online	Reclutamiento
Una plataforma como udemy?	Asesoría de cursos	Enseñanza de habilidades para mejorar profesionalmente	Cursos para potenciar mis habilidades

Nota: Herramienta libre para la generación de user test

Así mismo, las respuestas abiertas dónde se les preguntó acerca de qué otras características les gustaría que tuviera la herramienta donde la mayoría de los usuarios hizo hincapié en que les era importante la información de los cursos donde poder estudiar ciertas habilidades técnicas. Las respuestas se pueden visualizar en la figura 5.14

Figura 5.14

Resultados de respuestas de usuarios realizada en el user test versión 1. (%)



Nota: Herramienta libre para la generación de user test

Donde:

Claridad: Al ver la página de bienvenida sin alguna explicación, entender de qué trata la plataforma

- 🙄 Nada claro
- 🕶️ Algo claro
- 👁️ Muy claro
- 🔬 Totalmente claro

Figura 5.15
Resultados de User Test versión 1

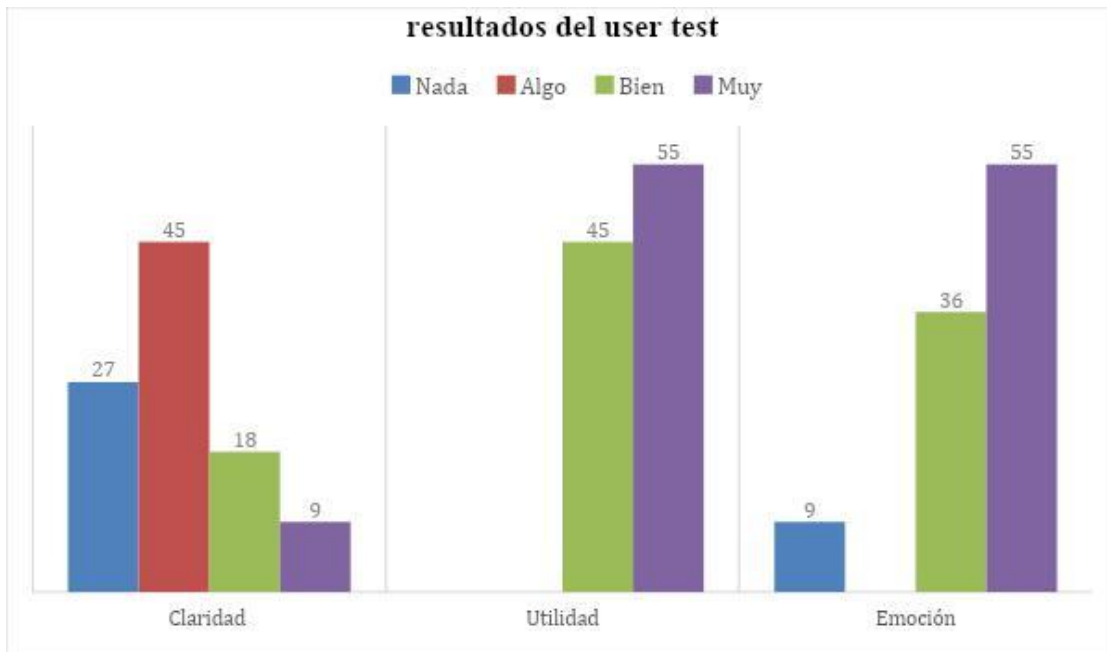


Figura 5.16
Respuestas de usuarios realizada en el user test acerca de nuevas funcionalidades

Que otra información o funcionalidad te gustaría tener en torno a la imagen de arriba Responses

Cómo puedo mejorarlos o donde	Tendencias, salarios, una ruta de aprendizaje de la habilidad.	Recomendación de lugares de estudio para las habilidades faltantes , descuentos	El orden de relevancia de los skills
En donde puedo aprender los skills que me hacen falta. Recomendaciones sobre los cursos: pueden ser gratuitos o pagados.	INFORMACION DE CURSOS	relación de tecnologías	EXPLICACION SOBRE QUE TRATA LA TECNOLOGÍA o con que otras se relaciona

Nota: Herramienta libre para la generación de user test

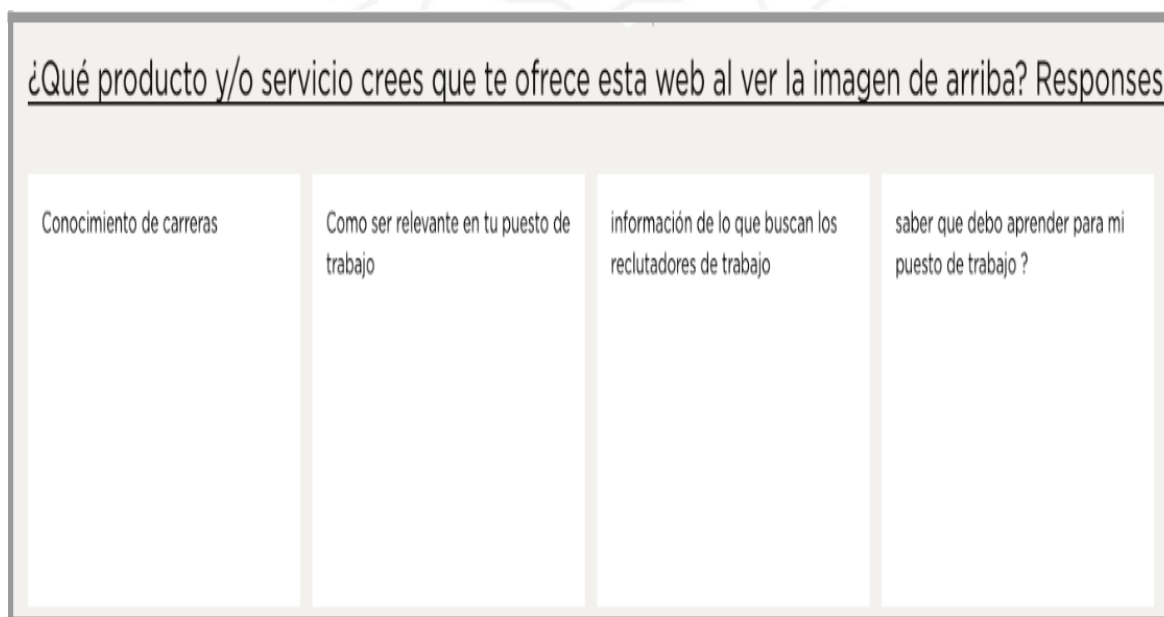
5.5.2 User test versión 2

Se consideraron por la recurrencia de respuestas cambios, en torno al diseño y comunicación en las pantallas sobre el propósito de la página web, así como agregar información de lugares dónde aprender y relación de habilidades.

Se utilizó la misma herramienta para realizar user test versión al mismo grupo de personas, pero con los nuevos diseños elaborados que se pueden observar en la figura 5.17. Se formularon las mismas preguntas que el user test versión 1 con las mismas opciones de respuesta. Las respuestas sobre la primera impresión al ver la página sin ninguna introducción mejoraron y se orientaron más al propósito de la página web.

Figura 5.17

Respuestas sobre percepción del producto de acuerdo a las pantallas mostradas, realizadas en el user test versión 2



¿Qué producto y/o servicio crees que te ofrece esta web al ver la imagen de arriba? Responses			
Conocimiento de carreras	Como ser relevante en tu puesto de trabajo	información de lo que buscan los reclutadores de trabajo	saber que debo aprender para mi puesto de trabajo ?

Nota: Herramienta libre para la generación de user test

Así mismo la emoción y usabilidad de la herramienta se mantuvo en casi todos los usuarios, ver figura 5.18 y 5.19

Figura 5.18

Dashboard de respuestas de usuarios realizada en el test de usuario versión 2 (%)



Nota: Herramienta libre para la generación de user test

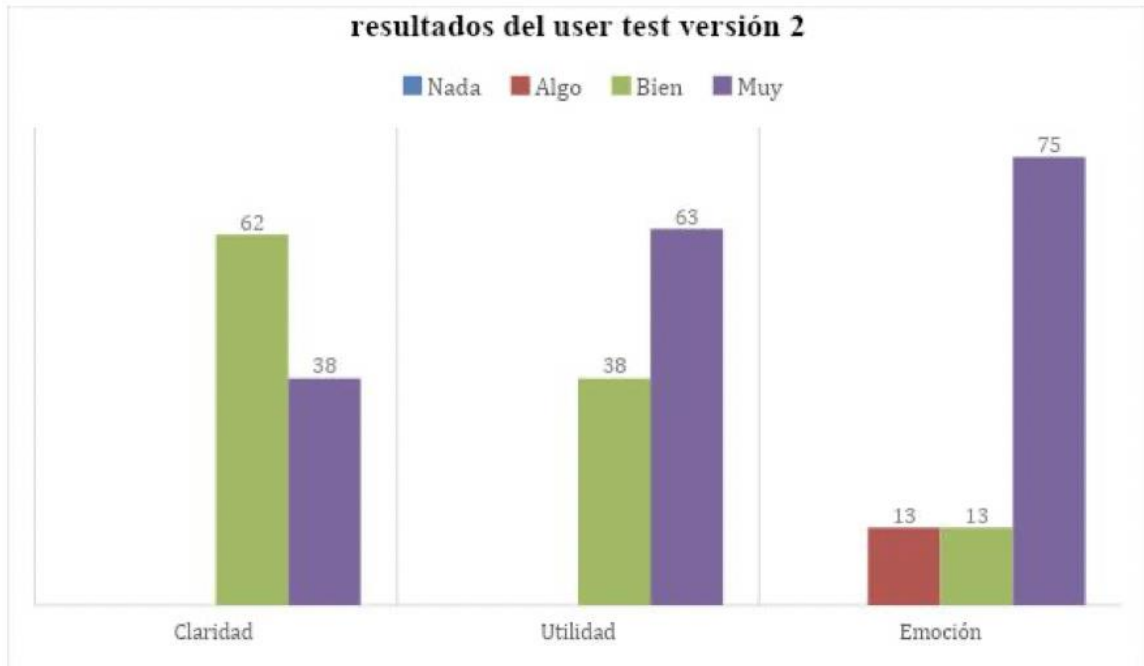
Donde:

Claridad: Al ver la página de bienvenida sin alguna explicación, entender de qué trata la plataforma

- 🤔 Nada claro
- 🕶️ Algo claro
- 👁️ Muy claro
- Totalmente claro

Figura 5.19

Resultados del user test versión 2



Los diseños mejorados a partir del user test versión 2 que servirán para el desarrollo de la página web para el MPV definido incluyen:

- Diseño página de bienvenida y registro

Figura 5.20

Página de Bienvenida



Figura 5.21
Registro de perfil

The image shows a registration profile page for H.A.B. The page has a blue gradient background. At the top left, there is a navigation bar with 'EQUIPO' and 'INICIAR SESIÓN / REGISTRO'. At the top right, there is the H.A.B. logo. In the center, there is a progress indicator for 'Paso 01' with a camera icon and the text 'Añade una foto'. Below this, there is a list of registration steps, each with a radio button:

- Sergio Lan...
- Ingresa tu correo
- Contraseña
- Confirmar contraseña

At the bottom, there is a 'Siguiete' button.

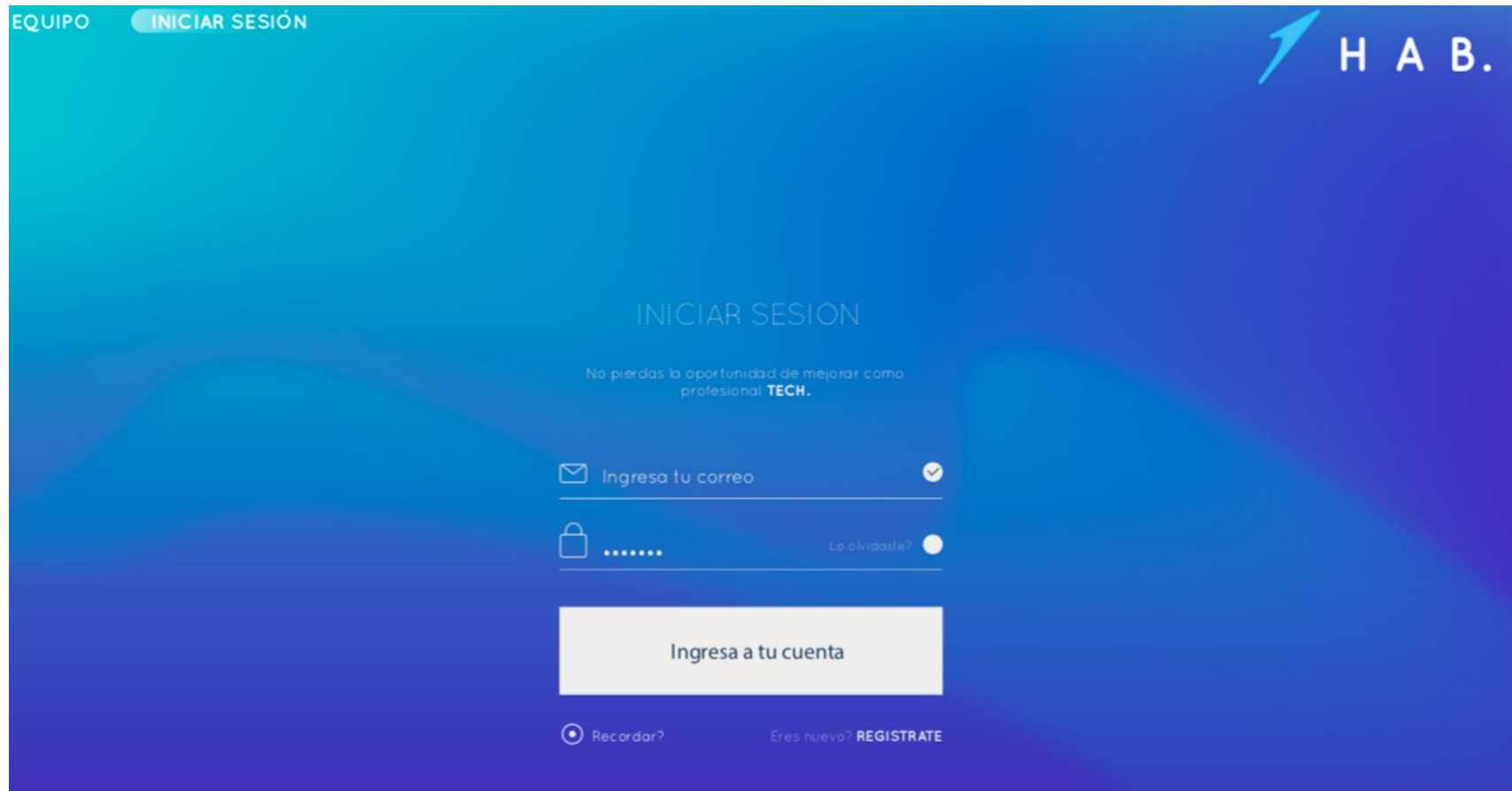
Figura 5.22
Registro de habilidades

The image shows a registration form for 'HAB.' on a blue gradient background. At the top left, there are navigation links: 'EQUIPO', 'INICIAR SESIÓN', and 'REGISTRO'. The 'REGISTRO' link is highlighted. In the top right corner is the 'HAB.' logo, which consists of a lightning bolt icon and the letters 'HAB.'. The main content area is titled 'Paso 02' with a checklist icon and two progress indicators (one white, one purple). Below this, the form asks the user to 'Elige el puesto al que quieres llegar.' with a dropdown menu. The next section is 'Selecciona que habilidades posees.' with three radio buttons, each followed by a dropdown menu. At the bottom, there is a checked checkbox for 'Acepto terminos y condiciones' and a large blue button labeled 'REGISTRARSE'.

- Inicio de Sesión

Figura 5.23

Página de inicio de sesión



- Blueprint de habilidades

Figura 5.24

Página principal de habilidades y cursos

EQUIPO **SERGIO VALENCINI**

H A B.

Sergio Valencini Sánchez
Front End Developer

SKILLS
Habilidades registradas:

React Javascript Ionic

PERÚ **EXTRANGERO**

Front End Developer

Resultados actualizados
Septiembre 2019.
Teniendo en cuenta los
puestos de trabajo y
requerimientos del mercado
laboral actual.

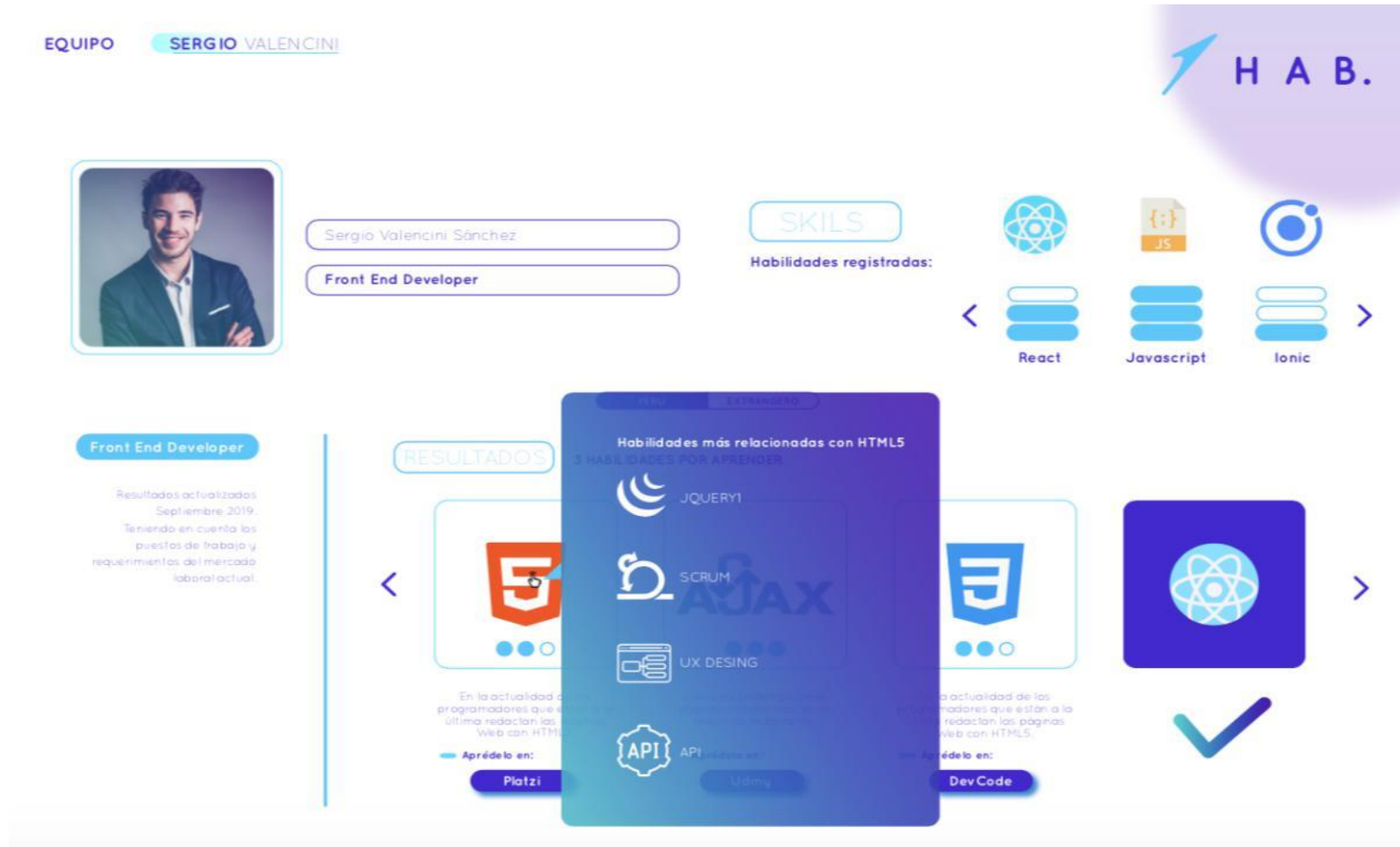
RESULTADOS **3 HABILIDADES POR APRENDER**

En la actualidad de los programadores que están a la última redactan las páginas Web con HTML5.
— Apréndelo en: **Platzi**

Con ajax podemos crear páginas interactivas, es un requisito importante.
— Apréndelo en: **Udmy**

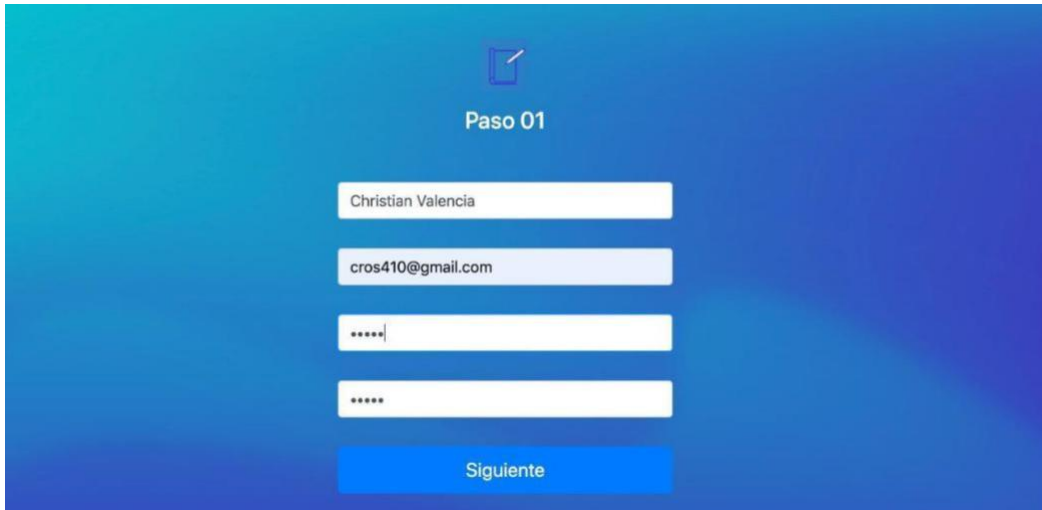
En la actualidad de los programadores que están a la última redactan las páginas Web con HTML5.
— Apréndelo en: **DevCode**

Figura 5.25
Página de listado de habilidades relacionadas



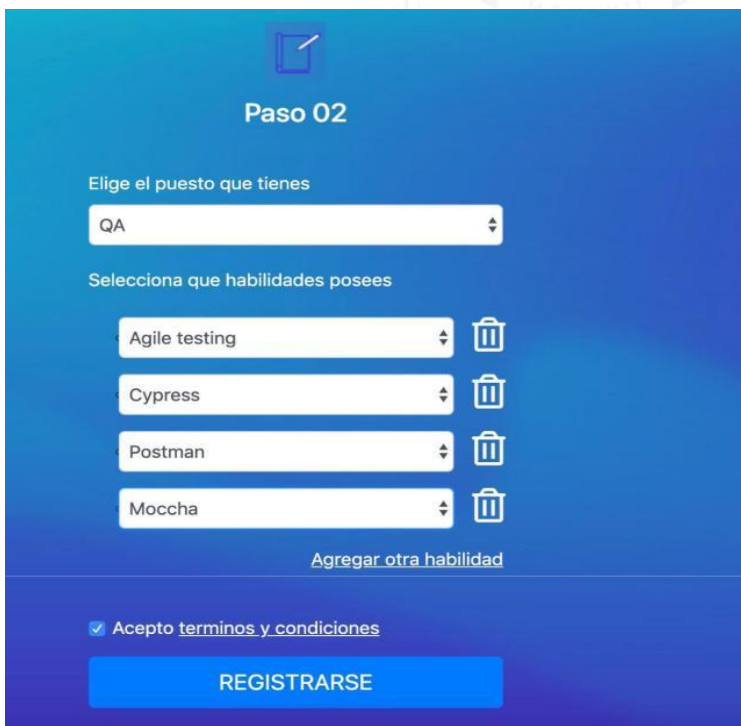
5.5.3 Casos de Uso 1

Figura 5.26
Registro Perfil



The screenshot shows a registration form titled "Paso 01" on a blue gradient background. At the top center is a small icon of a document with a pencil. Below the title are four input fields: the first contains "Christian Valencia", the second contains "cros410@gmail.com", and the next two contain masked characters "*****". A blue button labeled "Siguiente" is positioned at the bottom center.

Figura 5.27
Registro de habilidades técnicas



The screenshot shows a registration form titled "Paso 02" on a blue gradient background. At the top center is a small icon of a document with a pencil. Below the title, the text "Elige el puesto que tienes" is followed by a dropdown menu with "QA" selected. Below that, the text "Selecciona que habilidades posees" is followed by four rows, each with a skill name in a dropdown menu and a trash icon to its right: "Agile testing", "Cypress", "Postman", and "Moccha". Below these rows is the text "Agregar otra habilidad". At the bottom, there is a checked checkbox labeled "Acepto terminos y condiciones" and a large blue button labeled "REGISTRARSE".

Figura 5.28

Datos guardados del perfil en base de datos

```
select id, name, password, email from customer where email = 'cros410@gmail.com'
```

name	realm	username	password	email
12 Christian Valencia	NULL	NULL	\$2a\$10\$gDKYA3jRT.mKKF9F77fdQeOnpdfjrsW9/4Y6zS2WKM4amfHRPcA6	cros410@gmail.com


```
select customerId, skillId, name from customerSkill cs inner join skill s on cs.skillId = s.id where customerId = '12'
```

customerId	skillId	name
12	024d2ae4-5c10-473d-8246-7ff024922774	Junit
12	b93821eb-01bd-426a-b64c-404c5174999d	Moccha
12	c5ae2e24-ab58-4272-b3d8-f41282515ce3	Postman
12	d8bd194e-1551-4a76-a146-d85a77856348	Cypress

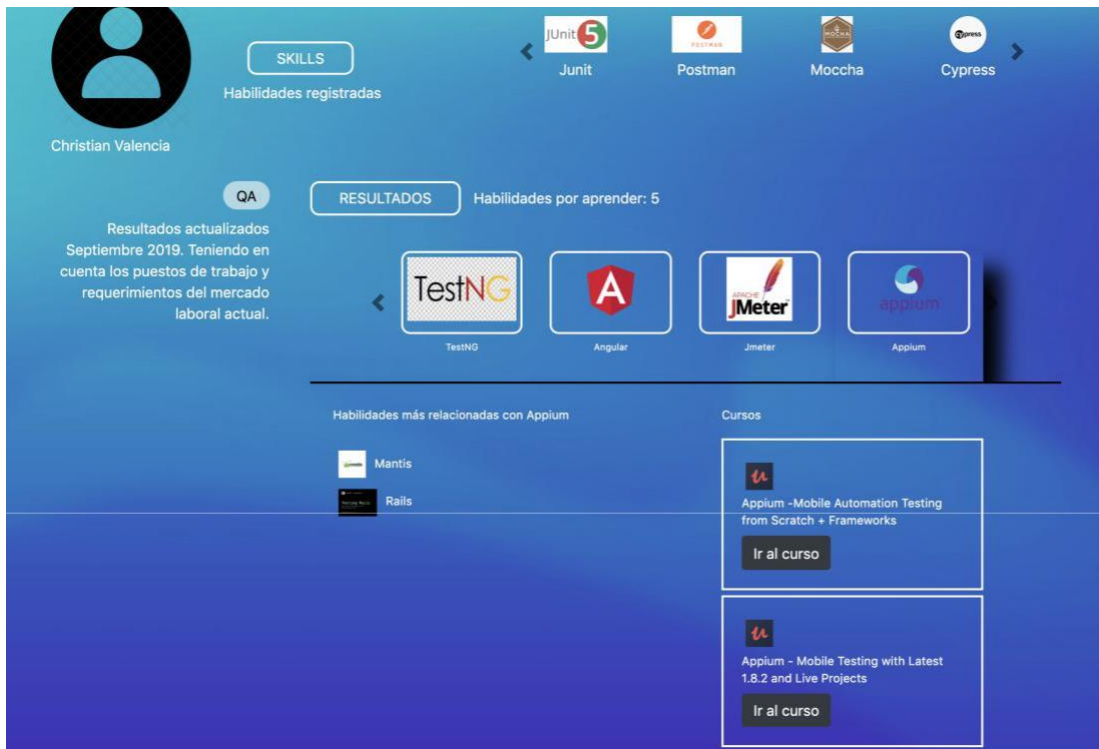
Figura 5.29

Habilidades y pesos relacionados para el skill "Appium"

```
select * from skillSkill ss inner join skill s on ss.secondarySkill = s.id where jobId = 'be34c016-dc83-4264-a4c4-4d8...' and primarySkill in ('961710a1-691c-472d-bb32-a10c14848917') order by primarySkill, weight desc
```

jobId	primarySkill	secondarySkill	weight	id	name	icon	
13252	be34c016-dc83-4264-a4c4-4d8...	961710a1-691c-472d-bb32-a10...	d21537b3-ac3e-4913-8039-eb3...	0.707107	d21537b3-ac3e-4913-8039-eb3...	Mantis	https://encrypted-tbn0.gstatic.co...
13254	be34c016-dc83-4264-a4c4-4d8...	961710a1-691c-472d-bb32-a10...	5d65bd34-49b4-4e0f-9ca7-023...	0.707107	5d65bd34-49b4-4e0f-9ca7-023...	Rails	https://encrypted-tbn0.gstatic.co...
13256	be34c016-dc83-4264-a4c4-4d8...	961710a1-691c-472d-bb32-a10...	14b81bc9-e581-44b9-b506-7fc...	0.288675	14b81bc9-e581-44b9-b506-7fc...	BDD	https://encrypted-tbn0.gstatic.co...
13253	be34c016-dc83-4264-a4c4-4d8...	961710a1-691c-472d-bb32-a10...	f6f5fe52-ce0a-4af5-ba16-86863...	0.213201	f6f5fe52-ce0a-4af5-ba16-86863...	Angular	https://upload.wikimedia.org/wik...
13255	be34c016-dc83-4264-a4c4-4d8...	961710a1-691c-472d-bb32-a10...	d953b761-a5d6-4122-8601-9b2...	0.150756	d953b761-a5d6-4122-8601-9b2...	TestNG	https://encrypted-tbn0.gstatic.co...

Figura 5.30
Dashboard de habilidades técnicas y cursos




5.5.4 Caso de Uso 2

Figura 5.31
Registro Perfil



Figura 5.32
Registro de Habilidades


Paso 02

Elige el puesto que tienes

DevOps

Selecciona que habilidades posees

CircleCI

Heroku

Travis

Weblogic

[Agregar otra habilidad](#)

Acepto [terminos y condiciones](#)

REGISTRARSE

Figura 5.33

Datos guardados del perfil del usuario que se registró en DB

```
select id, name, password, email from customer where email = 'paulo@gmail.com'
```

	name	password	email
14	Paulo Torres	\$2a\$10\$/PnJG8FUiWjIBYrrrTnX3uZNXreXuxnFr724mQF4pt2fDUdB3GbES	paulo@gmail.com


```
select customerId, skillId, name from customerSkill cs inner join skill s on cs.skillId = s.id where customerId = '14'
```

customerId	skillId	name
14	4419ec16-7719-4460-ac6c-f817e993a2d0	Travis
14	5abaeda7-918d-4a8b-a697-8f1efc4761a4	CircleCI
14	8a985f06-8232-4597-af2c-d1820515c066	Heroku
14	9458ffec-b28a-4987-b645-17a2cb806e84	Weblogic

Figura 5.34

Skill y pesos relacionados para el skill "BDD"

```
select * from skillSkill ss inner join skill s on ss.secondarySkill = s.id where jobId = '333febac-2f8b-43f8-ac5e-4a2ea45facbe' and primarySkill in ('f6f5fe52-ce0a-4af5-ba16-86863f9b9696', 'bd7d5476-b5f0-416c-8e5c-dfb2695e08cf', 'e9d4d71a-ed93-493f-9992-g904a92165ac', '14b81bc9-e581-44b9-b506-7fca93ec1111') order by primarySkill, weight desc
```

jobId	primarySkill	secondarySkill	weight	id	name
13419	333febac-2f8b-43f8-ac5e-4a2ea...	14b81bc9-e581-44b9-b506-7fc...	0.481125	024d2ae4-5c10-473d-8246-7ff0...	JUnit
13421	333febac-2f8b-43f8-ac5e-4a2ea...	14b81bc9-e581-44b9-b506-7fc...	0.288675	e52f733d-7a17-47b5-9274-586...	Genexus
13422	333febac-2f8b-43f8-ac5e-4a2ea...	14b81bc9-e581-44b9-b506-7fc...	0.288675	c5ae2e24-ab58-4272-b3d8-f412...	Postman
13424	333febac-2f8b-43f8-ac5e-4a2ea...	14b81bc9-e581-44b9-b506-7fc...	0.288675	c2665659-25aa-4fb1-8c58-b92c...	Cobol
13425	333febac-2f8b-43f8-ac5e-4a2ea...	14b81bc9-e581-44b9-b506-7fc...	0.288675	adf39f46-d8d2-4083-99d7-d0a8...	OWASP
13426	333febac-2f8b-43f8-ac5e-4a2ea...	14b81bc9-e581-44b9-b506-7fc...	0.204124	a8fba368-0d64-442c-9e8b-f6f7...	Golang
13418	333febac-2f8b-43f8-ac5e-4a2ea...	14b81bc9-e581-44b9-b506-7fc...	0.111111	f6f5fe52-ce0a-4af5-ba16-86863...	Angular
13420	333febac-2f8b-43f8-ac5e-4a2ea...	14b81bc9-e581-44b9-b506-7fc...	0.0912871	69c1feeb-d76d-46fc-b50d-d8b1...	POO
13423	333febac-2f8b-43f8-ac5e-4a2ea...	14b81bc9-e581-44b9-b506-7fc...	0.0912871	2700ada1-5ff1-4118-b333-3fb1...	Sonarqube
13427	333febac-2f8b-43f8-ac5e-4a2ea...	14b81bc9-e581-44b9-b506-7fc...	0.0745356	d953b761-a5d6-4122-8601-9b2...	TestNG

Figura 5.35
Dashboard de habilidades y cursos

The dashboard is titled "Habilidades registradas" (Registered Skills) and "Habilidades por aprender: 5" (Skills to learn: 5). The user's name is Paulo Torres. The dashboard is divided into several sections:

- SKILLS:** A horizontal list of registered skills: Heroku, Travis CI, Weblogic, and CircleCI.
- RESULTADOS:** A section for learning skills, showing "Habilidades por aprender: 5". It includes a "DevOps" filter and a "Resultados actualizados Septiembre 2019. Teniendo en cuenta los puestos de trabajo y requerimientos del mercado laboral actual." (Results updated September 2019. Taking into account job positions and market requirements). Below this are four skill cards: Containers (Docker), Azure (Azure DevOps), BDD (BDD), and TestNG (TestNG).
- Habilidades más relacionadas con BDD:** A section showing related skills, with JUnit listed.
- Cursos:** A section for recommended courses, including:
 - "BDD with Python (Cucumber/Gherkin for Python)" with a button "Ir al curso".
 - "Fundamentals of BDD (Java)" with a button "Ir al curso".

CONCLUSIONES

- El estudio de este proyecto permite identificar qué habilidades están buscando los empleadores, y las brechas de habilidades y las oportunidades para la creación de empleos. Las personas podrían detectar cuándo falta su experiencia en determinadas habilidades, así como qué habilidades son relevantes aprender y donde se podrían estudiarlas.
- En el contexto de la necesidad cambiante de habilidades técnicas, desarrollar las habilidades correctas que responden a las necesidades del mercado laboral y garantizar que estas habilidades sean plenamente utilizadas por individuos y empleadores, se vuelve cada vez más importante y son una importante fuerza impulsora del crecimiento a través de su efecto sobre la productividad laboral.
- HAB plantea la iniciativa para construir también un ecosistema del mercado laboral donde se utilizan datos abiertos para el desarrollo de las habilidades y así construir un mapa de habilidades para mostrar dónde y qué tan bien se conectan, la combinación de habilidades necesarias para un trabajo. Esto nos ayuda a proporcionar una correspondencia precisa de requisitos de trabajo y perfiles de talento en escala, y para monitorear el crecimiento rápido y nuevo.

RECOMENDACIONES

A continuación, detallaremos las recomendaciones:

- Aplicar para descubrir rutas de reevaluación y oportunidades de transición de trabajo. Utilizar una variedad de taxonomías de requisitos de trabajo y fuentes de datos encontrando una oportunidad para crear una infraestructura digital o servicios externos (B2B) que sirvan para la utilización de los datos abiertos para el desarrollo de servicios digitales para identificar las oportunidades laborales.
- Luego que se tenga una cantidad de datos significativo podría aplicarse un modelo de predicción para predecir las necesidades de habilidades a corto plazo mediante la combinación de datos estadísticos.
- Escalar a también tener la funcionalidad de un sistema de recomendación de trabajos inteligente de acuerdo a las habilidades obtenidas.
- Para el reclutamiento de perfiles, el proceso puede tener una versión 2.0 de cara a integrar en la plataforma el desarrollo de retos cargados por las empresas y generar desarrollos de casos donde se evidencie cuantitativamente y cualitativamente el manejo las competencias que necesiten y obtener respuestas y retroalimentación inmediata.
- Es necesario investigar en profundidad la relación entre las habilidades y el mercado laboral, esto quiere decir los vínculos entre las habilidades y los resultados clave del mercado laboral, así como la evaluación y la respuesta a las necesidades cambiantes de habilidades. Organizaciones tan reconocidas como OCDE actualmente también está extendiendo su trabajo sobre las habilidades a las economías y algunos países de bajos ingresos y ha desarrollado una base de datos internacional sobre indicadores de empleo y habilidades. Esto podría servir como referencias para futuros pasos.

GLOSARIO DE TÉRMINOS

Diccionario de los términos especializados empleados en el proyecto.

TÉRMINO	SIGNIFICADO
API	Acrónimo de Application Programming Interfac3
AWS	Hace referencia a la plataforma de Amazon Web Services
Back-End	Capa de procesamiento de los datos de la aplicación con los sistemas de soporte
Skill Path	Ruta de habilidades
Framework	Marco de trabajo para el desarrollo
Front-End	Capa de presentación al usuario final
IDC	International Data Corporation (IDC),
Machine Learning	Término que hace referencia a la disciplina científica del ámbito de la Inteligencia Artificial que crea sistemas que aprenden automáticamente.
MVP	Acrónimo de Minimum Viable Product
NLP	Acrónimo de natural language processing
OCDE	Organización para la Cooperación y el Desarrollo Económicos
SEO	Acrónimo de search engine optimization
TIR	Acrónimo de Tasa Interna de Retorno
UI	Acrónimo de User Interface
UIT	Unidad Impositiva Tributaria
UX	Acrónimo de User Experience
VAN	Acrónimo de Valor Actual Neto
Wireframe/Mockup	Se refiere a un esquema de página o plano de pantalla, es una guía visual que representa el esqueleto o estructura visual de un sitio web.

REFERENCIAS

- Adecco. (2018). Perú en puesto 83 del Índice de Competitividad del Talento Global 2017. *Adecco*. Recuperado del sitio de Internet de Adecco: <https://www.adecco.com.pe/peru-ranking-mundial-gtci/>
- Appaloosa Store. (2018). String Similarity Algorithms Compared. Recuperado de <https://medium.com/@appaloosastore/string-similarity-algorithms-compared-3f7b4d12f0ff#targetText=Jaro%2DWinkler%20Algorithm>
- Baeza-Yates, . G. N. R. (February 01, 1999). Faster Approximate String Matching. *Algorithmica : an International Journal in Computer Science*, 23, 2, 127-158
- Blavin, J. (8 de agosto del 2016). Caso LinkedIn vs Doe. Recuperado del sitio de internet del Información Jurídica Santa Clara Law Digital Commons <https://digitalcommons.law.scu.edu/cgi/viewcontent.cgi?referer=https://benbernardblog.com/&httpsredir=1&article=2261&context=historical>.
- Branded. (2018). Qué tienen los modelos ‘success fee’, que sientan tan bien a tu cuenta de resultados. Recuperado de <https://www.merca2.es/que-tienen-los-modelos-success-fee-que-sientan-tan-bien-a-tu-cuenta-de-resultados-1>
- Bykov (2015). Dice Data: How Tech Skills Connect. Recueperado de <https://insights.dice.com/2015/07/01/dice-data-how-tech-skills-connect/>
- Carnevale, A., Jayasundera, T., Repnikov, D. (2014). Recuperado del sitio de Internet de Georgetown University: https://cew.georgetown.edu/wp-content/uploads/2014/11/OCLM.Tech_.Web_.pdf
- Castillo, A. (2016). Régimen Laboral Especial de la Micro y Pequeña Empresa. Apuntes iniciales, a propósito de las modificaciones introducidas por la Ley N° 30056 y la aprobación del T.U.O. de la Ley MIPYME mediante Decreto Supremo N° 013-2013-PRODUCE. Recuperado de <http://blog.pucp.edu.pe/blog/agustinacastillo/>.
- Condor, J. (2019). Inversión en startups de Perú se triplicó apuntando a 8 emprendimientos. Recuperado de <https://gestion.pe/economia/empresas/peru-inversion-startups-triplico-8-emprendimientos-proyectos-268425-noticia/>

- Dehaze, A. (2017). Digitization in workforce solutions: why does it matter?. Recuperado del sitio de Internet de LinkedIn: <https://www.linkedin.com/pulse/digital-dexterity-alain-dehaze/>
- Deming, D., Kahn, L (2017) Skill requirements across firms and labor markets evidence from job postings for professionals. *Journal of Labor Economics*, 36(S1) . 337-369. <https://doi.org/10.1086/694106>
- Djumaliev, J., Lima, A. y Sleeman C. (2018). Classifying Occupations According to Their Skill Requirements in Job. Recuperado del sitio de Internet del Economic Statistics Centre of Excellence: <https://www.escoe.ac.uk/wp-content/uploads/2018/03/ESCoE-DP-2018-04.pdf>.
- Eranl. (2017). A Judge Just Ordered LinkedIn to Allow Scraping. Recuperado de: <https://webhose.io/blog/technology/a-judge-just-ordered-linkedin-to-allow-scraping-heres-why/>
- Falco. A (2003). Productividad del trabajador del conocimiento: El gran desafío del siglo XXI. Recuperado de: <http://www.cema.edu.ar/productividad/download/2003/Falco.pdf>
- Ganova, A. (2017). Business model for design thinking a case study for the evolution 6² model (tesis de maestría). Facultad de Ingeniería de la Universidad de Porto. Tesis
- Gestión. (2017). ¿Cómo está la industria de Software en Perú? Recuperado de <https://gestion.pe/tecnologia/apesoft-entidad-dedicada-tic-duplicaria-crecimiento-sector-software-132697-noticia/>
- Huong Nguyen, Hailey (6 de febrero de 2019) Algorithms for Text Classification- Part . Recuperado de <https://towardsdatascience.com/algorithms-for-text-classification-part-1-naive-bayes-3ff1d116fdd8>
- Ilyankou, I (2014) Comparison of Jaro-Winkler and Ratcliff/Obershelp algorithms in spell check Candidate. Recuperado de <https://ilyankou.files.wordpress.com/2015/06/ib-extended-essay.pdf>
- Junior, E (03 de marzo del 2018) COMPUTING THE LEVENSHTAIN (EDIT) DISTANCE OF TWO STRINGS USING C#. Recuperado de <https://www.eleamarjr.com/en/2018/03/computing-the-levenshtein-distance-of-two-strings-using-c/>

Li, Susan. (11 de octubre 2018). Natural Language Processing for Fuzzy String Matching with Python, 41, 233-242. Recuperado de <https://towardsdatascience.com/natural-language-processing-for-fuzzy-string-matching-with-python-6632b7824c49>.

Lossio, D. (2016). Plan de negocio para una consultora de pequeñas y medianas empresas en el Perú (Tesis de Máster en Dirección de Empresas). Universidad de Piura. (Tesis)

Manpower (2018). DIGITIZATION AND WHY SKILLS Recuperado del sitio de Internet del grupo Manpower: https://www.manpowergroup.com/wps/wcm/connect/5943478f-69d4-4512-83d8-36bfa6308f1b/MG_Skills_Revolution_lores.pdf

Ministerio de Trabajo y Promoción de Empleo. (2019) . Portal de Transparencia - III Trimestre 2019 - Información de Personal Activo, Ministerio de Trabajo y Promoción del Empleo. Recuperado del sitio de Internet de Ministerio de Trabajo: <https://www.gob.pe/institucion/mtpe/informes-publicaciones/344436-informacion-de-personal-activo>

Ministerio de Trabajo y Promoción de Empleo. (2019) Resultado de la Encuesta Nacional de variación Mensual del Empleo (ENVME) a Setiembre de 2019. Recuperado del sitio de Internet de Ministerio de Trabajo: <https://www.gob.pe/institucion/mtpe/informes-publicaciones/344654-resultado-de-la-encuesta-nacional-de-variacion-mensual-del-empleo-envme-a-setiembre-de-2019>

Mondragón, P. (19 de marzo del 2018). De Edison a Tim Brown: algunas ideas sobre Design Thinking [mensaje en un blog]. Recuperado de <https://blog.antropologia2-0.com/es/edison-tim-brown-design-thinking/>.

Rodriguez, N. (2019). El mercado laboral para profesionales de Tecnologías de la Información. Recuperado de <https://rpp.pe/columnistas/nadiarodriguez/el-mercado-laboral-para-profesionales-de-tecnologias-de-la-informacion-noticia-1179382>

OECD. (2018). Skills for Job. Recuperado del sitio de Internet de Organización para la Cooperación y el Desarrollo Económicos: https://www.oecdskillsforjobsdatabase.org/data/Skills%20SfJ_PDF%20for%20WEBSITE%20final.pdf

OECD. (2016). Getting Skills Right: Assessing and Anticipating Changing Skill Needs. Recuperado del sitio de Internet de Organización para la Cooperación y el Desarrollo Económicos: <https://www.oecd.org/employment/getting-skills-right-assessing-and-anticipating-changing-skill-needs-9789264252073-en.htm>

OECD. (2012). OECD SOCIAL, EMPLOYMENT AND MIGRATION WORKING PAPERS No. 120 FOR THE JOB: ¿OVER-QUALIFIED OR UNDER-SKILLED Recuperado del sitio de Internet de Organización para la Cooperación y el Desarrollo Económicos: <https://www.oecd.org/els/48650012.pdf#targetText=A%20worker%20is%20classified%20as,his%20Fher%20occupation%20is%20positive.&targetText=Situation.>

Orlik, J, Casasbuenas, J & Helkkula, K. (2018). Designing inclusive skills policy for the digital age. *Digital Frontrunners*. Recuperado de https://media.nesta.org.uk/documents/Digital_Frontrunners_Web.pdf.

Pichihua, S. (2018). Startups en el Perú se incrementarán en 60% en el 2018. Recuperado de <https://andina.pe/agencia/noticia-startups-el-peru-se-incrementaran-60-el-2018-684507.aspx>

Quintos, M. (2017). ¿Cómo está la industria de Software en Perú? Recuperado de <http://www.timov.la/article/como-esta-la-industria-de-software-en-peru>

Randstad. (2018). Global report randstad workmonitor Q4 2018. Recuperado del sitio de Internet de Randstad: <https://cdn2.hubspot.net/hubfs/481927/Workmonitor/Randstad%20Workmonitor%20global%20report%20Q4%20-%20December%202018.pdf>.

Santa-Olalla, B. (2018). Design thinking aplicado al storytelling. Recuperado del sitio de Internet <https://innovacionaudiovisual.com/2018/01/19/design-thinking-aplicado-al-storytelling/>

Yordanov, V. (2018). Recuperado de: Introduction to Natural Language Processing for Text

Wang, Shuheng & Li, Guohao & Bao, Yifan. (2017). A novel improved fuzzy support vector machine based stock price trend forecast model. Recuperado 10 agosto, 2019

Wu, W. (2016). A Method for Fuzzy String Matching. *2016 International Computer Symposium (ICS)*, 380-383.

Zapata, C. (2019). *Talento tech*, [Correo electrónico].



BIBLIOGRAFÍA

Department of Health, Education, and Welfare, National Institute of Education,
(2001). *Resources in Education*.

Kouzis-Loukas, D. (2016). *Learning Scrapy*. Birmingham: Packt Publishing,
Limited.





ANEXOS

Anexo 1 Apelación HIQ LABS V. LINKEDIN por caso Scraping

FOR PUBLICATION

UNITED STATES COURT OF APPEALS FOR THE NINTH CIRCUIT

HIQ LABS, INC.,
Plaintiff-Appellee,

v.

LINKEDIN CORPORATION,
Defendant-Appellant.

No. 17-16783

D.C. No.
3:17-cv-03301-EMC

OPINION

Appeal from the United States District Court
for the Northern District of California
Edward M. Chen, District Judge, Presiding

Argued and Submitted March 15, 2018
San Francisco, California

Filed September 9, 2019

Before: J. Clifford Wallace and Marsha S. Berzon, Circuit
Judges, and Terrence Berg,* District Judge.

Opinion by Judge Berzon;
Concurrence by Judge Wallace

* The Honorable Terrence Berg, United States District Judge for the Eastern District of Michigan, sitting by designation.

Anexo 1 Apelación HIQ LABS V. LINKEDIN por caso Scraping

infringement, misappropriation, unjust enrichment, conversion, breach of contract, or breach of privacy, may also lie. *See, e.g., Associated Press v. Meltwater U.S. Holdings, Inc.*, 931 F. Supp. 2d 537, 561 (S.D.N.Y. 2013) (holding that a software company’s conduct in **scraping** and aggregating copyrighted news articles was not protected by fair use).

D. Public Interest

Finally, we must consider the public interest in granting or denying the preliminary injunction. Whereas the balance of equities focuses on the parties, “[t]he public interest inquiry primarily addresses impact on non-parties rather than parties,” and takes into consideration “the public consequences in employing the extraordinary remedy of injunction.” *Bernhardt v. Los Angeles Cty.*, 339 F.3d 920, 931–32 (9th Cir. 2003) (citations omitted).

As the district court observed, each side asserts that its own position would benefit the public interest by maximizing the free flow of information on the Internet. HiQ points out that data scraping is a common method of gathering information, used by search engines, academic researchers, and many others. According to hiQ, letting

2004) (holding that the use of a scraper to glean flight information was unauthorized as it interfered with Southwest’s use and possession of its site, even if the scraping did not cause physical harm or deprivation), *with Ticketmaster Corp. v. Tickets.Com, Inc.*, No. 2:99-cv-07654-HLH-VBK, 2003 WL 21406289, at *3 (C.D. Cal. Mar. 7, 2003) (holding that the use of a web crawler to gather information from a public website, without more, is insufficient to fulfill the harm requirement of a trespass action); *Intel Corp. v. Hamidi*, 30 Cal. 4th 1342, 1364 (2003) (holding that “trespass to chattels is not actionable if it does not involve actual or threatened injury” to property and the defendant’s actions did not damage or interfere with the operation of the computer systems at issue).

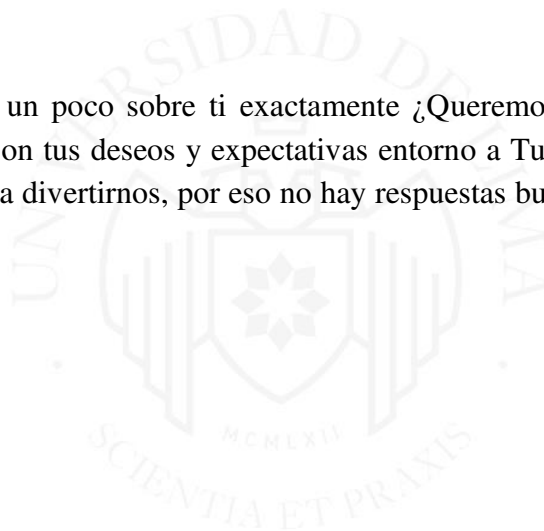
Anexo 2 Guía de Entrevista para profesionales

Presentación

Hola mi nombre es **Diandra Sánchez** actualmente estoy en un un proceso de titulación por la modalidad de innovación, y estoy realizando una investigación para conocer un poco más sobre ti y tu crecimiento laboral a lo largo del tiempo. Quisiera comenzar por agradecer por el tiempo brindado y la información obtenido. Así mismo valoro mucho tu opinión honesta. Si no sabes o no quieres responder a cualquier pregunta, por favor no dudes en comentarnos para pasar a otro punto.

Primero, déjame contarte un poco sobre mí, ya te había comentado mi nombre es Diandra, estude la carrera de sistemas y me desempeño actualmente como Analista de Calidad en el centro de innovación de la UTP y llegué aquí (se cuenta historia personal)

Ahora cuéntame un poco sobre ti exactamente ¿Queremos entender lo que te motiva, ¿cuáles son tus deseos y expectativas entorno a Tu experiencia laboral? Estamos aquí para divertirnos, por eso no hay respuestas buenas ni malas. ¡Todo vale!



Anexo 2 Guía de Entrevista para profesionales

Parte 1: Conocer al usuario a nivel personal y profesional y cómo mantiene a sus competencias.
Antes que nada, queremos conocerte un poco mejor
1. Cuéntanos acerca de ti: ¿cómo es un día normal en tu vida? (no dejar pasar por alto: edad, ocupación, trabajo, domicilio...)
2. ¿Cuéntanos un poco sobre tu familia? ¿A qué se dedican tus padres? ¿Tienes hermanos? ¿Vives con ellos?
3. ¿Hay alguna persona que dependa de ti? ¿De qué manera depende de ti? (puede ser económica, emocionalmente...)
4. ¿Cómo llegaste a este trabajo? ¿Hace cuánto trabajas en esta área? ¿Qué puesto de trabajo tienes actualmente y qué cosas realizas? .
5. ¿Podrías contarnos un poco acerca de tus sueños y metas profesionales?
Parte 2: Comprender las motivaciones y los factores que influyen en aprender algo nuevo que es relevante para el mercado.
6. ¿Cómo sabes que es momento de aprender algo nuevo referente a tecnología?
7. ¿Cuál es el medio por el cual sabes que debes actualizarte o aprender algo que está en tendencia y no conocías?
7. ¿Sabes qué habilidades son requeridas para el rol que desempeñas en el mercado laboral?
Parte 3 Aceleración de carrera: Conocer sobre cuáles son las características principales en su toma de decisiones para o potenciar su perfil laboral.
8. ¿Consideras que tienes herramientas que te muestren las oportunidades de crecimiento para cada etapa de tu carrera?
9. Ahora realizaremos una dinámica, imagina que tu empresa te diría que te dará una capacitación o curso de alguna habilidad técnica que desees ¿Qué es lo que harías?, ¿Cómo tomarías esta decisión?
10. Ahora pensemos que quieres cambiarte de trabajo o que es momento de buscar otro trabajo - ¿Cómo empiezas esa búsqueda? - ¿Buscas en base a las habilidades que tienes o solo por el puesto de trabajo a fin que tienes? ¿Cómo te ayudaría conocer lo que las empresas están pidiendo? - ¿Cómo afrontas cuando te enfrentas a que este puesto de trabajo te pida alguna habilidad que tú no tienes?

Anexo 3 Factores concurrentes de categorización empresarial: Número de trabajadores y el nivel de ventas anuales

2.2.1. Según la Ley MYPE (vigente hasta el 02-07-2019). Factores concurrentes de categorización empresarial: el número de trabajadores y el nivel de ventas anuales

El artículo 5° del Decreto Supremo N° 007-2008-TR (Ley MYPE) establecía las características concurrentes de las MYPES. De modo esquemático, teníamos lo siguiente:

Características..... TrabajadoresVentas Anuales

Microempresa.....De 1 hasta 10.....Hasta 150 UIT (S/.630,000)

Pequeña Empresa...De 1 hasta 100 ...Hasta 1700 UIT (S/.7'400,000)

Donde: Valor de la UIT para el Ejercicio 2019 era de S/.4,200.

Microempresa: Tenía de uno (1) hasta diez (10) trabajadores inclusive y un nivel de ventas anuales no mayor a 150 UIT. En este sentido, no había variación –en relación con lo que señalaba la Ley N° 28015– en los límites para calificar como microempresa.

Pequeña Empresa: Tenía de uno (1) hasta cien (100) trabajadores inclusive y un nivel de ventas anuales no mayor a 1700 UIT. Por ende, en comparación con lo que señalaba la Ley N° 28015, el número máximo de trabajadores subía de 50 a 100 y, además, en cuanto al volumen de ventas, se eliminaba el límite inferior de 150 UIT y se elevaba el límite superior de 850 UIT a 1700 UIT.

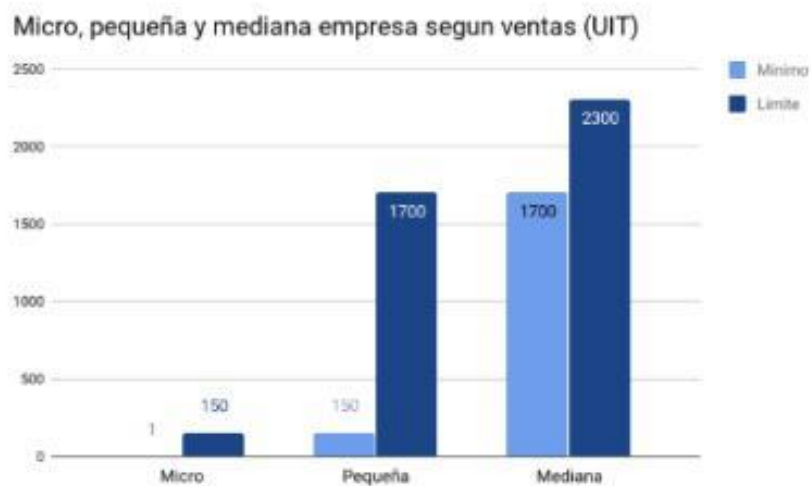
Fuente: Ministerio del trabajo y promoción del empleo (2016)

Anexo 3 Factores concurrentes de categorización empresarial: Número de trabajadores y el nivel de ventas anuales

El método para clasificar el tamaño de las empresas es según sus ventas

La forma de clasificar a las empresas en el Perú es en base a las ventas anuales que haya tenido.

- **La Micro empresa:** Tiene un nivel de venta anual menor a 150 UIT (607 500 soles para el 2017).
- **La Pequeña empresa:** Tiene un nivel de venta anual a partir de 150 UIT hasta 1 700 UIT (Mayor a 607 500 hasta 6 885 000 Soles para el 2017).
- **La Mediana empresa:** Tiene un nivel de ventas anuales superior a 1 700 UIT hasta 2 300 UIT (Mayor 6 885 000 hasta 9 315 000 Soles para el 2017).
- **La Gran empresa:** Tiene un nivel de ventas anual superior a 2 300 UIT (Mayor 9 315 000 Soles para el 2017).



Fuente: Ministerio del trabajo y promoción del empleo (2016)

Anexo 4 Valores de equivalencia de UIT por año

VALORES DE LA UIT

AÑO	VALOR (S/.)	BASE LEGAL	OBSERVACIONES
2019	4,200	D.S. N° 298-2018-EF	
2018	4,150	D.S. N° 380-2017-EF	
2017	4,050	D.S. N° 353-2016-EF	
2016	3,950	D.S. N° 397-2015-EF	
2015	3,850	D.S. N° 374-2014-EF	
2014	3,800	D.S. N° 304-2013-EF	
2013	3,700	D.S. N° 284-2012-EF	
2012	3,650	D.S. N° 233-2011-EF	
2011	3,600	D.S. N° 252-2010-EF	
2010	3,600	D.S. N° 311-2009-EF	
2009	3,550	D.S. N° 189-2008-EF	
2008	3,500	D.S. N° 209-2007-EF	
2007	3,450	D.S. N° 213-2006-EF	

Fuente: Superintendencia Nacional de Aduanas y de Administración Tributaria (2019)

Anexo 5 Cuestionario a empresas

¿Con qué frecuencia suelen reclutar perfiles tecnológicos durante el año en tu empresa?	¿Cuánta es la inversión de tiempo promedio para cubrir un puesto de trabajo del área de tecnología en tu empresa?	¿Cuánta es la inversión de dinero promedio para cubrir un puesto de trabajo del área de tecnología en tu empresa?	¿Qué bolsas de trabajo usan actualmente para su proceso de reclutamiento?	¿Han tenido deserción de personal en el área de tecnología, con cuánta frecuencia se da estos casos? De ser así, indicar un promedio	¿Estarías dispuesto a usar un servicio de reclutación de perfiles que se especialice en reclutación de perfiles solamente del rubro tecnología?	¿Estarías dispuesto a pagar alguna herramienta que se especialice en reclutación de perfiles solamente del rubro tecnología?
Todo el año	Mucho ... Hemos automatizado el proceso pero igual lleva tiempo .. el problema de raíz es la falta de propuesta de valor para la retención de talentos	Un sueldo adicional por cada búsqueda	LinkedIn	En InDigital no porque venimos manejándolo muy bien y con una propuesta de valor mejor, pero en Intercorp Retail si hay rotación constante.	Sí	Sí
2 veces al año	Desde la entrevista hasta la inducción, unas 80 horas	No hay inversión de por medio	LinkedIn, Grupos de Interés, Búmeran	No, porque no invertimos	No	
1 a 3 veces al año	4000 a 5000 soles por mes estándar	4000 a 5000 soles por mes estándar	Linkedin	No que yo sepa	Sí	Sí
4 veces	2 meses	8000 soles	LinkedIn, referencias	No todavía	Sí	Sí
2 anuales	Se mantiene un pipeline de talento, no tenemos una búsqueda reactiva sino proactiva	Depende del nivel que se necesita su sueldo aprox	Para perfiles senior LinkedIn, para perfiles más juniors referencias	Si, pero por que se fueron a estudiar al extranjero	Sí	Sí
cada 2 meses	1.5 meses	sueldo+publicaciones en bolsas	Linkedin, Aptitus, Bumeran, Universidades	No	Sí	Sí

Anexo 5 Cuestionario a empresas

¿Con qué frecuencia suelen reclutar perfiles tecnológicos durante el año en tu empresa?	¿Cuánta es la inversión de tiempo promedio para cubrir un puesto de trabajo del área de tecnología en tu empresa?	¿Cuánta es la inversión de dinero promedio para cubrir un puesto de trabajo del área de tecnología en tu empresa?	¿Qué bolsas de trabajo usan actualmente para su proceso de reclutamiento?	¿Han tenido a deserción de personal en el área de tecnología, con frecuencia da estos casos? De ser así, indicar un costo promedio	¿Estarías dispuesto a usar un servicio de reclutación de perfiles que se especialice en reclutación de perfiles de tecnología?	¿Estarías dispuesto a pagar alguna herramienta que se especialice en reclutación de perfiles de tecnología?
2	Contratamos a una empresa de reclutamiento	80% del sueldo del developer requerido	LinkedIn para validar perfil	Si, 2 vez al año normalmente... Nos cuesta 2 sueldos el reemplazo	Sí	Sí
5	2 semanas	750	Universia, Krowdy	No	No	-
2 x mes	3 a 4 semanas de búsqueda, entrevistas y contratacion	30 horas por persona + 50 soles en servicios a terceros	Bumeran, LinkedIn y redes sociales regulares	1 al mes	Sí	Sí
2	3 meses	Depende	Se encarga el área de GDH	2 al mes	No	tendría que ver su eficacia y rapidez
10	30 DIAS	850	LN	SI 3 VECES	Sí	Sí
4	20 dias	4000	Bumeran	No	Sí	Sí
Todos los meses	2 a 3 semanas	No manejamos presupuesto	El empleo.com de Colombia	Cada 2 meses	Sí	Sí
Las fechas más fuertes son enero y agosto	3 meses	e invierte en hunting y consultoría externa con un sueldo de promedio del puesto requerido	LinkedIn, búmeran y bolsas de universidades.	No	Sí	Sí
Trimestral	2 meses	3500 aprox	Bumeran, LinkedIn	Mínimo 3500. La rotación en el área de sistemas es alta.	Sí	Sí

Anexo 6 Número promedio de personas que se demandarán para puestos de trabajo del sector tecnología entre técnicos y profesionales. MINISTERIO DE TRABAJO Y PROMOCIÓN DEL EMPLEO (2018)

GRÁFICO N° 2.25
PERÚ: PERSONAL PROFESIONAL UNIVERSITARIO A CONTRATAR, SEGÚN PRINCIPALES SECTORES ECONÓMICOS, 2018
(Absoluto)



1/ Incluye establecimientos financieros y seguros (452); electricidad, gas y agua (215); minería (104); y, bienes inmuebles (4).

Fuente: MTPE - DGPE - Encuesta de Demanda Ocupacional a empresas de 20 a más trabajadores.

Elaboración: MTPE - DGPE - Dirección de Investigación Socio Económico Laboral (DISEL).

Las ocupaciones más requeridas para los trabajadores profesionales universitarios en los sectores económicos analizados durante el 2018, serían lideradas principalmente por los profesores de educación superior (707), ingenieros civiles (686), profesionales de la informática (554), profesores de enfermería (423) y técnicos en programación por computadora (397). Las diez ocupaciones que más se requerirían concentrarían el 60,4% del total de trabajadores profesionales universitarios demandados a nivel nacional.

GRÁFICO N° 2.26
PERÚ: OCUPACIONES MÁS REQUERIDAS DEL PERSONAL PROFESIONAL UNIVERSITARIO A CONTRATAR, 2018
(Absoluto)



Fuente: MTPE - DGPE - Encuesta de Demanda Ocupacional a empresas de 20 a más trabajadores.

Elaboración: MTPE - DGPE - Dirección de Investigación Socio Económico Laboral (DISEL).

El nivel de remuneración promedio mensual para los trabajadores profesionales universitarios en los sectores estudiados a nivel nacional ascendería a S/ 3 073. De las ocupaciones más requeridas, los técnicos en programación por computadora serían la ocupación mejor remunerada con S/ 4 954 promedio al mes. Asimismo, le sigue los médicos (S/ 4 882), ingenieros civiles (S/ 4 308), profesionales de la informática (S/ 3 515), ingenieros industriales (S/ 2 924) y auxiliares en contabilidad, del derecho y administración (S/ 2 271).

Fuente: Ministerio del trabajo y promoción del empleo (2019)

Anexo 6 Número promedio de personas que se demandarán para puestos de trabajo del sector tecnología entre técnicos y profesionales. MINISTERIO DE TRABAJO Y PROMOCIÓN DEL EMPLEO (2018)

PERSONAL A CONTRATAR Y REMUNERACIÓN PROMEDIO MENSUAL

ANEXO N° 1.3
PERÚ: PERSONAL A CONTRATAR Y REMUNERACIÓN PROMEDIO MENSUAL,
SEGÚN GRUPO OCUPACIONAL Y PRINCIPALES OCUPACIONES, 2018
(Absoluto y soles)

Grupo Ocupacional	Trabajadores	Remuneración Mínima Mensual	Remuneración Promedio Mensual	Remuneración Máxima Mensual
Total	103 653	407	1 389	35 000
Gerente y directivo	21	1 800	13 950	25 000
Gerentes de ventas y comercialización	16	5 000	7 456	25 000
Gerentes de planta y proyectos	2	10 000	20 000	30 000
Gerentes generales de comercio mayorista	1	7 000	7 000	7 000
Productores de televisión	1	4 500	4 500	4 500
Gerentes de administración	1	6 000	6 000	6 000
Profesional	5 092	850	2 587	9 000
Profesores de educación superior	708	850	1 523	6 000
Ingenieros civiles	686	1 500	4 308	15 000
Profesionales de la informática	554	1 500	3 515	20 000
Profesores de educación primaria	454	850	1 488	1 700
Profesionales de enfermería	423	900	1 781	4 000
Otros	2 267	-	-	-
Técnico	12 684	850	1 606	20 000
Instructores de entrenamiento funcional	2 620	850	1 498	1 500
Agentes técnicos de ventas y representantes de firmas	2 498	850	1 045	3 000
Técnicos en programación por computadora	1 882	1 000	2 472	5 000
Técnicos en electricidad, electrónica y telecomunicaciones	1 746	900	1 595	4 429
Técnicos en ingeniería mecánica	934	900	2 245	5 600
Otros	3 004	-	-	-
Empleado	12 446	850	1 014	8 000
Empleados de centro de llamadas	6 878	850	885	1 600
Empleados de aprovisionamiento y almacenaje	1 663	850	1 087	3 723
Cajeros y cobradores de ventanilla	1 151	850	938	1 800
Recepcionistas y empleados de oficina	789	407	1 029	2 000
Empleados de servicios financieros	480	850	1 574	3 000
Otros	1 485	-	-	-
Trabajador de los servicios personales	15 191	850	1 194	2 000
Personal de los servicios de seguridad	5 203	850	1 424	2 800
Vendedores al por menor	3 625	850	1 102	3 000
Demostradores	2 495	407	1 053	1 800
Bármanes y mozos	1 348	525	953	2 200
Cocineros calificados	1 130	450	1 096	2 144
Otros	1 390	-	-	-
Agricultor	905	1 100	1 100	1 100
Agricultores	620	1 000	1 006	1 500
Avicultores	108	950	988	1 000
Cultivadores o recolectores de crustáceos y mariscos	95	995	1 001	1 020
Jefes de explotación agrícola	66	2 000	2 000	2 000
Criadores de ganado	6	850	925	1 000
Otros	10	-	-	-
Obrero	57 314	850	1 109	3 400
Peones de labranza y agropecuarios	17 611	850	1 013	2 000
Peones de obras públicas	4 062	1 000	1 904	3 000
Obreros de la construcción	3 925	1 100	2 055	5 000
Limpiadores de establecimientos	3 505	850	891	1 500
Peones de la industria manufacturera	3 189	850	961	1 700
Otros	25 022	-	-	-

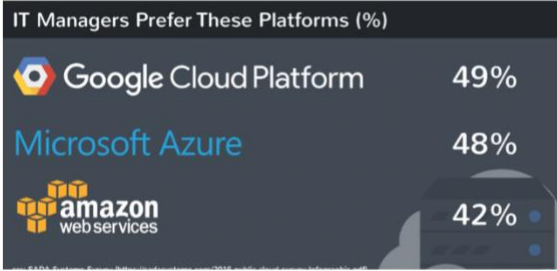
Nota: La suma de las partes puede no coincidir con el total debido al redondeo de las cifras.

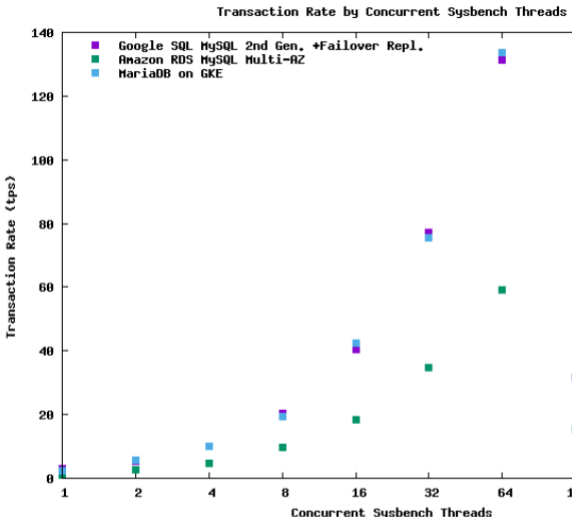
Fuente: MTPE - DGPE - Encuesta de Demanda Ocupacional a empresas de 20 a más trabajadores.

Elaboración: MTPE - DGPE - Dirección de Investigación Socio Económico Laboral (DISEL).

Fuente: Ministerio del trabajo y promoción de empleo (2019)

Anexo 7 Análisis de Stack Tecnológico

Infraestructura		Google	Amazon	Azure	
Cloud	Precio	- \$300 dólares para usarlo durante 1 año	Free Tier (uso por capacidad)	- \$200 dólares para usarlo durante 1 año	
		<table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%; vertical-align: top;"> <p>AWS - Virginia</p> <p>c4.4xlarge</p> <ul style="list-style-type: none"> - 16 CPUs - 30 GB RAM <p>=====</p> <p>+ \$613.42 instance</p> <p>500GB SSD EBS Volume</p> <p>=====</p> <p>+ \$62.50 disk</p> <p>+ \$1040.00 IOPS (16K IOPS)</p> <p>=====</p> <p>= \$1715.92/month on-demand</p> <p>= \$1102.50/month 3yr (\$8,580.00)</p> </td> <td style="width: 50%; vertical-align: top;"> <p>GCP - South Carolina (East)</p> <p>Custom</p> <ul style="list-style-type: none"> - 16 CPU - 30 GB RAM <p>=====</p> <p>+ \$357.25 instance (sustained-use)</p> <p>667GB SSD PD Volume</p> <p>=====</p> <p>+ \$113.39 disk</p> <p>+ \$800.00 IOPS (20K IOPS)</p> <p>=====</p> <p>= \$478.64/month</p> </td> </tr> </table> <p style="text-align: center;">Fuente: HFT Guy (2017)</p>			<p>AWS - Virginia</p> <p>c4.4xlarge</p> <ul style="list-style-type: none"> - 16 CPUs - 30 GB RAM <p>=====</p> <p>+ \$613.42 instance</p> <p>500GB SSD EBS Volume</p> <p>=====</p> <p>+ \$62.50 disk</p> <p>+ \$1040.00 IOPS (16K IOPS)</p> <p>=====</p> <p>= \$1715.92/month on-demand</p> <p>= \$1102.50/month 3yr (\$8,580.00)</p>
<p>AWS - Virginia</p> <p>c4.4xlarge</p> <ul style="list-style-type: none"> - 16 CPUs - 30 GB RAM <p>=====</p> <p>+ \$613.42 instance</p> <p>500GB SSD EBS Volume</p> <p>=====</p> <p>+ \$62.50 disk</p> <p>+ \$1040.00 IOPS (16K IOPS)</p> <p>=====</p> <p>= \$1715.92/month on-demand</p> <p>= \$1102.50/month 3yr (\$8,580.00)</p>	<p>GCP - South Carolina (East)</p> <p>Custom</p> <ul style="list-style-type: none"> - 16 CPU - 30 GB RAM <p>=====</p> <p>+ \$357.25 instance (sustained-use)</p> <p>667GB SSD PD Volume</p> <p>=====</p> <p>+ \$113.39 disk</p> <p>+ \$800.00 IOPS (20K IOPS)</p> <p>=====</p> <p>= \$478.64/month</p>				
	Comunidad	 <p style="text-align: center;">FUENTE : SADA Systems (2017)</p>			
Base de Datos		Cloud SQL	RDS	SQL Database	
	Comunidad	<p>All products with 50+ points are designated as a Leader in their category.</p> <ol style="list-style-type: none"> 1 SQL Azure <div style="display: flex; align-items: center; margin-top: 5px;"> <div style="width: 20px; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 20px; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 20px; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 20px; height: 10px; background-color: cyan; margin-right: 5px;"></div> <div style="width: 20px; height: 10px; background-color: blue; margin-right: 5px;"></div> </div> <p style="font-size: small; margin-top: 2px;">2,868 views 2,368 comparisons 10 reviews 416 words/review 8.5 ave</p> 2 Oracle Database as a Service <div style="display: flex; align-items: center; margin-top: 5px;"> <div style="width: 20px; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 20px; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 20px; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 20px; height: 10px; background-color: cyan; margin-right: 5px;"></div> <div style="width: 20px; height: 10px; background-color: blue; margin-right: 5px;"></div> </div> <p style="font-size: small; margin-top: 2px;">2,016 views 1,480 comparisons 9 reviews 640 words/review 8.3 ave</p> 3 Google Cloud SQL <div style="display: flex; align-items: center; margin-top: 5px;"> <div style="width: 20px; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 20px; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 20px; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 20px; height: 10px; background-color: cyan; margin-right: 5px;"></div> <div style="width: 20px; height: 10px; background-color: blue; margin-right: 5px;"></div> </div> <p style="font-size: small; margin-top: 2px;">2,740 views 2,233 comparisons 4 reviews 571 words/review 9.0 ave</p> 4 Amazon RDS <div style="display: flex; align-items: center; margin-top: 5px;"> <div style="width: 20px; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 20px; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 20px; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 20px; height: 10px; background-color: cyan; margin-right: 5px;"></div> <div style="width: 20px; height: 10px; background-color: blue; margin-right: 5px;"></div> </div> <p style="font-size: small; margin-top: 2px;">2,784 views 2,510 comparisons 1 reviews 425 words/review 8.0 ave</p> <p style="text-align: center;">Fuente: IT Central Station - Report (2020)</p>			

	Características	 <p>Transaction Rate by Concurrent Sysbench Threads</p> <p>Google Cloud SQL has higher TPS (transactions per second) than RDS for MySQL. With the performance goes rapidly down. Our MariaDB performs mostly better than the rest laa.</p> <p>Fuente: Marek Bartík (2017)</p>		
Kubernetes	Alta Disponibilidad	Sí	No	En Desarrollo
	Integración	Kubernetes Engine se desarrolló internamente dentro de Google, y se integra con otros servicios de Google y así se tiene una curva de aprendizaje más corta,	Elastic Container Service para Kubernetes (EKS)	Azure Kubernetes Services (AKS)
	Precio	Pago solo por las máquinas virtuales que ejecutan los nodos de Kubernetes.	Cobra \$ 0.20 por hora por cada clúster implementado	Pago solo por las máquinas virtuales que ejecutan los nodos de Kubernetes