REPUBLIC OF TURKIYE
SAKARYA UNIVERSITY
GRADUATE SCHOOL OF BUSINESS

ONTOLOGY DRIVEN, ARTIFICIAL INTELLIGENCE BASED CAREER PLANNING SYSTEM FOR INDIVIDUALS

PHD THESIS

Bahadır AKTAŞ
ORCID: 0000-0002-3650-6471

Department of the Institute: Management Information Systems

Thesis Supervisor: Assoc. Prof. Adem AKBIYIK
ORCID: 0000-0001-7634-4545

JULY – 2024

This thesis entitled "Ontology Driven, Artificial Intelligence Based Career Planning System For Individuals", prepared by Bahadır AKTAŞ, was found successful as a result of the Thesis Defense Examination held on 24/06/2024 in accordance with the relevant articles of Sakarya University Graduate Education and Training Regulation and was accepted as a Doctoral Thesis by our jury.

Thesis Supervisor: Assoc. Prof. Dr. Adem AKBIYIK

Sakarya University

Jury Members: Prof. Dr. Yasemin ÖZDEMİR

Sakarya University

Assoc. Prof. Dr. Halil İbrahim CEBECİ

Sakarya University

Prof. Dr. Çiğdem TARHAN

Dokuz Eylül University

Asst. Prof. Dr. Aysun BOZANTA HAKYEMEZ

Boğaziçi University

SAKARYA ÜNİVERSİTESİ	**T.C. SAKARYA ÜNİVERSİTESİ İŞLETME ENSTİTÜSÜ TEZ SAVUNULABİLİRLİK VE ORİJİNALLİK BEYAN FORMU**	Sayfa: 1/1

Öğrencinin

Adı Soyadı	:	Bahadır Aktaş
Öğrenci Numarası	:	D189054002
Enstitü Anabilim Dalı	:	Yönetim Bilişim Sistemleri
Enstitü Bilim Dalı	:	Yönetim Bilişim Sistemleri
Programı	:	Doktora Programı
Tezin Başlığı	:	Ontology Driven, Artificial Intelligence Based Career Planning System for Individuals
Benzerlik Oranı	:	%7

Sakarya Üniversitesi İşletme Enstitüsü Lisansüstü Tez Çalışması Benzerlik Raporu Uygulama Esaslarını inceledim. Enstitünüz tarafından Uygulama Esasları çerçevesinde alınan Benzerlik Raporuna göre yukarıda bilgileri verilen tez çalışmasının benzerlik oranının herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi beyan ederim.

...... / ...... / 20….
İmza
Öğrenci

Sakarya Üniversitesi İşletme Enstitüsü Lisansüstü Tez Çalışması Benzerlik Raporu Uygulama Esaslarını inceledim. Enstitünüz tarafından Uygulama Esasları çerçevesinde alınan Benzerlik Raporuna göre yukarıda bilgileri verilen öğrenciye ait tez çalışması ile ilgili gerekli düzenleme tarafımca yapılmış olup, yeniden değerlendirilmek üzere gsbtez@sakarya.edu.tr adresine yüklenmiştir.

Bilgilerinize arz ederim.

...... / ...... / 20….
İmza
Danışman

Uygundur

Danışman
Unvanı / Adı-Soyadı:
Tarih: ...... / ...... / 20....

İmza:

☐ Kabul Edilmiştir ☐ Reddedilmiştir EYK Tarih ve No: …… / …… / 20…. …… / ……	**Enstitü Birim Sorumlusu Onayı**

PREFACE

I sincerely thank my advisor, Assoc. Prof. Dr. Adem AKBIYIK for his valuable contributions and leadership at every stage of my thesis and guiding me through my academic career. I am very grateful to Prof. Dr. Yasemin ÖZDEMİR and Assoc. Prof. Dr. Halil İbrahim CEBECİ, who contributed to the shaping of my thesis and generously shared their ideas. I thank Prof. Dr. Çiğdem TARHAN and Asst. Prof. Dr. Aysun BOZANTA HAKYEMEZ, who are part of my thesis defense jury and provided valuable ideas.

I thank Prof. Dr. Giancarlo GUIZZARDI and Dr. Luis FERREIRA PIRES for letting me be a part of the team at Twente University and for their guidance, and I thank Rogrigo F. CALHAU for being a great friend and a classmate.

I am very grateful to my roommates, Research Assistants Oğuzhan ARI and Kübra ÇALIŞKAN AKBIYIK, and my colleagues Dr. Tuğba KOÇ, Dr. Naciye Güliz UĞUR, and Dr. Büşra ALMA ÇALLI for their support. I thank Emre UĞUR, chairman of the board of PIKDER, for his support. And I thank Kadir ÇAKIR for his friendship and insightful discussions.

I thank Dr. Ömer Sezai AYKAÇ, Büşra AYKAÇ, and Umut Ali AYKAÇ, who have never failed to show their love and support since the day I set foot in Sakarya and made us feel at home here.

I am indebted to my mother, Müslime (Elmas) AKTAŞ, my father, Halit AKTAŞ, my brother, Cihangir AKTAŞ, Emine AKTAŞ, and Kıvanç Halit AKTAŞ, to whom I can never repay.

I am grateful for the presence and support of my partner, beloved wife, and colleague, Neşe BAZ AKTAŞ. She never hesitated to support me.

Even though she has been in our lives for a very short time, I am thankful for my daughter, my baby Tomris, who makes me wake up more ambitious and motivated every morning.

Bahadır AKTAŞ

24.06.2024

TABLE OF CONTENTS

ABBREVIATIONS

ADABOOST : Adaptive Boosting Classifier
AI : Artificial Intelligence
BCT : Boundaryless Career Theory
CATBOOST : Categorical Boosting Classifier
CCT : Career Construction Theory
CRISP-DM : Cross-industry standard process for data mining
CV : Curriculum Vitae
DSR : Design Science Research
GBC : Gradient Boosting Classifier
IS : Information Systems
IT : Information Technology
KNN : K-Nearest Neighbour
LR : Logistic Regression
ML : Machine Learning
NEET : Not in Employment Education Training
ODCM : Ontology-Driven Conceptual Modeling
ONTOUML : Ontology-based Unified Modeling Language
PSMP : Professional Social Media Platforms
RFC : Random Forest Classifier
SABiO : Systematic Approach for Building Ontologies
SVM : Support Vector Machine
UFO : Unified Foundational Ontology
UML : Unified Modeling Language
XGBOOOST : Extreme Gradient Boosting

TABLES

FIGURES

ABSTRACT

Aktaş, B. (2024). *Ontology driven, artificial intelligence based career planning system for individuals* (Unpublished doctoral thesis). Sakarya University.

With rapid technological advancements and evolving job market requirements, individual career planning has become increasingly challenging. IT professionals, career changing individuals, and NEET (Not in Employment, Education, or Training) individuals often lack the resources for effective career development and lifelong learning. This thesis proposes the development of an AI-based career planning system that leverages data science and machine learning to provide personalized career guidance. The main research problem addressed in this thesis is "How does an artificial intelligence-based career planning system, utilizing ontology, data science, and machine learning techniques, affect the career development and job alignment of information technology professionals, career changers, and NEET individuals?"

The thesis employs a Design Science Research (DSR) methodology, focusing on the creation and evaluation of innovative artifacts to solve practical problems. The research follows the CRISP-DM model for data mining, encompassing phases such as business understanding, data understanding, data preparation, modeling, evaluation, and deployment. The primary data source consists of career profiles from Professional Social Media Platforms (PSMPs). Key components include an ontology-driven conceptual model utilizing Unified Foundational Ontology (UFO) and OntoUML for accurate representation of career-related data, machine learning models to calculate job fit scores and generate skill improvement recommendations, and a functional prototype to validate the system's feasibility and functionality, focusing on IT sector positions.

This thesis provides a significant contribution to the field of career planning by developing an AI-based system that addresses the dynamic needs of the labor market. It emphasizes the importance of personalized career guidance and lifelong learning, particularly for individuals in the IT sector and those seeking career transitions or re-entry into the workforce.

The proposed system aligns with Turkey's strategic goals for developing a qualified workforce and supporting national initiatives in lifelong learning and human resource development. It also sets a foundation for future research and development in AI-based career planning systems.

Keywords: Career Planning, Artificial Intelligence, Ensemble Learning, Ontology Driven Conceptual Modeling, Design Science Research

ÖZET

Aktaş, B. (2024). *Bireyler için ontoloji odaklı, yapay zeka tabanlı kariyer planlama sistemi* (Yayımlanmamış doktora tezi). Sakarya Üniversitesi.

Hızla gelişen teknoloji ve değişen iş piyasası gereksinimleri ile bireysel kariyer planlaması giderek daha zor hale gelmiştir. BT profesyonelleri, kariyer değiştiren bireyler ve NEET (İstihdamda, Eğitimde veya Öğretimde Olmayan) bireyler, etkili kariyer geliştirme ve yaşam boyu öğrenme için genellikle kaynaklardan yoksundur. Bu tez, veri bilimi ve makine öğrenimini kullanarak kişiselleştirilmiş kariyer rehberliği sağlayan bir yapay zeka tabanlı kariyer planlama sistemi geliştirmeyi önermektedir. Bu tezde ele alınan ana araştırma problemi, "Ontoloji, Veri bilimi ve makine öğrenimi tekniklerini kullanarak bir yapay zeka tabanlı kariyer planlama sistemi bilişim teknolojileri profesyonelleri, kariyer değiştiren bireyler ve NEET bireylerin kariyer gelişimlerini ve iş uyumlarını nasıl etkiler?" sorusudur.

Tez, pratik problemleri çözmek için yenilikçi artefaktların oluşturulması ve değerlendirilmesine odaklanan bir Tasarım Bilimi Araştırma (DSR) metodolojisini kullanmaktadır. Araştırma, iş anlayışı, veri anlayışı, veri hazırlığı, modelleme, değerlendirme ve dağıtım gibi aşamaları kapsayan veri madenciliği için CRISP-DM modelini takip etmektedir. Birincil veri kaynağı, Profesyonel Sosyal Medya Platformlarından (PSMP'ler) elde edilen kariyer profillerinden oluşmaktadır. Ana bileşenler arasında kariyerle ilgili verilerin doğru temsili için Birleştirilmiş Temel Ontolojiyi (UFO) ve OntoUML'yi kullanan ontoloji tabanlı kavramsal model, iş uyumu puanlarını hesaplamak ve beceri geliştirme önerileri oluşturmak için makine öğrenimi modelleri ve sistemin uygulanabilirliğini ve işlevselliğini doğrulamak için çalışan bir prototip bulunmaktadır.

Bu tez, dinamik iş piyasası ihtiyaçlarını karşılayan bir yapay zeka tabanlı sistem geliştirerek kariyer planlama alanına önemli bir katkı sağlamaktadır. Özellikle BT sektöründeki bireyler ve kariyer değişimi veya iş gücüne yeniden giriş arayanlar için kişiselleştirilmiş kariyer rehberliğinin ve yaşam boyu öğrenmenin önemini vurgulamaktadır. Önerilen sistem, nitelikli bir iş gücü geliştirmeyi ve ulusal yaşam boyu öğrenme ve insan kaynakları geliştirme girişimlerini desteklemeyi amaçlayan Türkiye'nin stratejik hedefleriyle uyumludur. Ayrıca yapay zeka tabanlı kariyer planlama sistemleri alanında gelecekteki araştırma ve geliştirme için bir temel oluşturmaktadır.

Anahtar Kelimeler: Kariyer Planlama, Yapay Zeka, Ensemble Öğrenme, Ontoloji Tabanlı Kavramsal Modelleme, Tasarım Bilimi Araştırması

INTRODUCTION

With the rapid development of technology and information systems, new professions are emerging, and job definitions and requirements are rapidly changing (Hirschi, 2018; Kuzgun, 2021). In this dynamic and volatile job market, people plan their careers according to the requirements and needs of the company they work for and their own goals, talents, and competencies (Arthur, 1994; Forret & Sullivan, 2002). This approach may facilitate developing and maintaining a skilled workforce that benefits the labor market. However, when it comes to human resources processes in organizations, the focus is on managing careers within the organization rather than individual career development (Erdoğmuş, 2003). For this reason, companies seem to be limited in their ability to provide individual career planning at the macro level and independent of the company. Therefore, the significance of individual career planning, lifelong learning, and the necessity of career guidance becomes apparent.

In a fast-paced work environment, people find it challenging to keep up with rapidly changing job requirements. This may lead to professional isolation and disconnect (Hirschi, 2018; Kuzgun, 2021). To avoid such situations, individuals need guidance, especially when making decisions such as changing jobs, finding a new position, or relocating to a different city. Therefore, it is important for individuals to conduct strategic career planning (Granrose & Portwood, 1987; Hall, 1986) and commit to lifelong learning (Super, 1963) to be successful in their careers and excel through career-changing decisions. However, studies on career development and career planning are mostly focused on the career development and planning of students (Çarkıt, 2019; Duffy & Sedlacek, 2007; İstanbullu Dinçer et al., 2013; Jackson, 2017; Özdemir & Kibar, 2018; Polat et al., 2016; Sevinç & Siyez, 2018; Tomy & Pardede, 2019; Waddell & Bauer, 2005; Willis & Wilkie, 2009). This reveals that career planning and lifelong learning is needed for individuals who are already in the workforce but want to improve or transition between careers and individuals who graduated but are not employed, such as Not in Employment Education or Training individuals (NEETs).

Career development is defined as a process of growth from childhood to retirement (Super, 1963). In this context, career development can be associated with lifelong education. In life-span, life-space theory, phases of professional life are proposed within the developmental framework (Super, 1980). These phases are the growth phase (4-14 years), exploration phase (15-24 years), establishment phase (25-44 years), maintenance phase (45-64 years), and decline phase (65 years and later). Individuals follow a certain path within the scope of primary, elementary, and undergraduate education in the growth and research phases and clearly see the requirements to achieve their goals (Luo, 2016; Van der Horst et al., 2017). However, with the establishment phase, the roadmap for one's future career may become unclear. This is especially true for those who have completed their education and entered the workforce, where they may become subject to the career planning goals of their organizations (Montgomery, 2017; Seger, 2016; Van der Horst et al., 2017). Furthermore, NEET individuals are left completely without a roadmap (Felaco & Parola, 2022; Robertson, 2018).

Research Problem

Even though the literature states that the responsibility for career development lies with the individual (Adekola, 2011; Leibowitz et al., 1991), information systems developed with a focus on career development in organizational career management and human resources planning. This thesis argues for the use of artificial intelligence-based systems to support decisions related to the individual career development and lifelong education plans of individuals during and after the establishment phase. In addition, this study argues for the utilization of professional data, including experience, education, and skills, shared in professionally oriented social media platforms (PSMPs) for the development of machine learning and artificial intelligence models in order to represent the current state of the job market truly. This thesis also suggests modeling PSMP data and the research domain using an ontology-driven conceptual modeling method to determine and validate the suitability of the data to the research domain and the objectives. Consequently, it aims to design and develop an ontology-driven artificial intelligence-based system that will facilitate this process.

The main research problem to be addressed within the scope of the research is determined as "How does an artificial intelligence-based career planning system, utilizing ontology, data science, and machine learning techniques, affect the career development and job alignment of information technology professionals, career changers, and NEET individuals?"

In order to tackle the research problem, the subproblems to be addressed in this research are:

- Is ontology-based conceptual modeling an effective method for ensuring the scalability, adaptability, and interoperability of an artificial intelligence-based career planning system?
- Are data science methods effective in accurately determining job alignment for specific information technology positions based on an individual's current skills, experiences, education, and qualifications?
- How effectively can an artificial intelligence-based system provide personalized, accurate recommendations on the skills and competencies individuals need to acquire or improve to align with targeted information technology positions?
- How effectively does the prototype of an ontology-driven, artificial intelligence-based career planning system demonstrate its feasibility and functionality?"

Research Aim

This thesis aims to design and prototype an ontology-driven AI-based system that can be used for individual career planning. Thus, it will ensure the individual career development of people who do not have access to career counseling and planning services within the scope of lifelong education and contribute to the establishment of a sustainable, qualified workforce. In this context, this study focuses on defining ontological foundations of the domain and developing an artificial intelligence-based information system that will provide position matching with a job fit score and skill recommendations in career planning processes for IT sector positions. This system will be aimed at people looking for an IT position, people in the process of making a career-changing move, and NEETs.

This thesis aims to use ontology-driven conceptual modeling, data science, machine learning, and artificial intelligence techniques to achieve these goals: (1) creating a well-founded ontological model of the PSMP data and the research domain, (2) developing an ML/AI-based model that calculates the job fit of individuals for selected IT positions based on their capabilities, (3) presenting suggestions on which competencies and skills individuals need to acquire or improve (within the scope of lifelong education) to improve their job fit for a specific IT position, (4) creating a working prototype of this system. To ensure this, this thesis will employ design science research (DSR) methodology to provide a basis for research design, and the goals of the research will be represented as artifacts of the DSR process.

Importance of the Research

The artifacts of this thesis will guide individuals at the establishment phase in their career (according to life-span, life-space theory) to progress and excel, enhance their skills to avoid disconnection with the fast-paced IT industry, and make transitions between positions or sectors by helping them equip themselves with the necessary skills and competencies. Additionally, it will assist NEETs who are not currently active in the labor market or education by helping them compare their competencies with the current demands of job positions in the labor market. This aims to help them take the first step towards integrating into the active workforce by giving them a job fit score for a position with an AI-based system and encouraging them for lifelong learning. This will promote the adoption of lifelong learning and education through micro-credentials in the Turkish labor market and contribute to sustainably raising the level of education and welfare of society. By providing the level of fit for a position and tailored suggestions to individuals for achieving their targets, whether it is switching sectors, aiming for a new position, or relocating for work. This approach will provide guidance for people to work in their desired job, sector, or city. This empowerment helps people find the right opportunities and allows them to improve their quality of life.

This thesis will contribute to Turkey's efforts to acquire a sustainable, qualified workforce by providing guidance to individuals (employees or NEETs) through their career development stages. Specifically, this thesis aligns with the initiatives outlined in Turkey's 11th and 12th Development Plan under the section "Developing Qualified Human Resources and Closing the Human Capital Gap". The 11th Development Plan of Turkey, which was the guiding framework (and was the current development plan) during the development of this thesis, emphasized lifelong learning as a key component of human resources development strategies. The plan recognizes that enhancing the quality of human resources will contribute to the competitiveness of the nation. It aims to enable career development and lifelong learning at both early and mature career stages to support training qualified and competent individuals (Cumhurbaşkanlığı Strateji ve Bütçe Başkanlığı, n.d.; *On İkinci Kalkınma Planı*, n.d.).

This strategic focus is intended to develop a well-trained workforce that can meet the evolving demands of the global market. With career development and lifelong learning initiatives implemented in this thesis through an AI-based career planning system, Turkish firms will be equipped with qualified employees, enhancing their competitiveness in global markets. As the Turkish IT industry gains access to a more skilled and sustainable labor pool, it is anticipated that Turkey will secure a larger share of the global market, strengthening its overall economic standing.

Finally, this thesis will provide evidence that an ontology-driven AI-based system can be effectively used in individual career development processes, paving the way for future research. The ontological conceptual models developed as a part of this study are aimed to serve as a foundation to determine and validate the suitability of the data to the research domain and the objectives, guiding future research aimed at expanding upon this work. This will validate the use of information systems in career development and planning processes and inspire and support continued innovation and improvement.

Research Methodology

This thesis argues for the use of an ontology-driven AI-based system to support decisions regarding career development and life learning during and after the settlement phase of the life-span, life-space theory (Super, 1980). In the context of this argument, a design science research (DSR) approach will be used. The reason for adopting the DSR approach is to provide an innovative solution to a problem, to create an artifact, and to observe the feasibility of this solution in real-world circumstances. This thesis will employ the DSR methodology to guide the design and development process to achieve the research objective of developing an artificial intelligence-based individual career planning system. This approach is intended to enhance both the theoretical contributions of the research and its practical implications through several artifacts such as the ontology model, the system design, and ML models.

The methodology used as a basis for this thesis's data science, machine learning (ML) and AI model development processes is Cross-industry standard process for data mining (CRISP-DM). This methodology guides data science projects' start, progression, and outputs. According to this model, data science projects are structured into six main phases: (1) Business Understanding, (2) Data Understanding, (3) Data Preparation, (4) Modeling, (5) Evaluation, and (6) Deployment (Chapman et al., 2000). This study follows these steps in the methodology for the ML/AI model development phase in Chapter 3.

During the system development, it is aimed to use the data collection tool developed as a part of this thesis, as well as the data collected using this tool. The dataset, which will be used for developing machine learning and artificial intelligence models, consists of job experience, education, skills, and competency information that individuals publicly share on PSMPs. The literature (Mashayekhi & Head, 2018) and previous efforts (Aktaş & Akbıyık, 2019) have shown that users of PSMPs view these platforms to share and promote their digital resumes, and they approach them differently from other social media platforms.

Additionally, beyond the support of existing literature and previous efforts, this thesis aims to model the PSMP data to be obtained and the research domain using the ontology-driven conceptual modeling (ODCM) method. ODCM aims to determine and validate the suitability of the data to the research domain and the objectives of this thesis. Within the scope of this thesis, ODCM was conducted using Unified Foundational Ontology (UFO), OntoUML ontological modeling language, and Systematic Approach for Building Ontologies (SABiO) ontology development methodology (all of which are described in Chapter 2). OntoUML model was created using The Visual Paradigm application with open source OntoUML plugin. The resulting model will establish the link between the research domain, the dataset, and the research questions, providing a guiding conceptual model for the development of the AI-based information system.

Research Scope and Limitations

As with every academic study, this thesis also has certain limitations and a defined scope. Firstly, this study's scope was limited to specific positions within the field of information systems (IS) and information technology (IT). This limited scope was chosen because the IT industry effectively represents rapidly changing requirements. Another reason is the high adoption rate of PSMP use among information technology professionals which is the main data source of this study (J. Davis et al., 2020).

Another limitation of the study is that the information technology position data in Turkey, which is to be collected from PSMPs within the scope of this research, is heavily concentrated in a single city and a region, Istanbul. Consequently, it was not possible to conduct a location-based analysis or to use the location as a criterion within the scope of this thesis.

Organization of the Thesis

This study consists of five chapters and is structured as follows. In the the study, the theoretical framework was discussed under two main topics. In the first chapter, individual career planning is examined with a focus on information systems. Subsequently, in the second the ontology approach and the ontology-based conceptual modeling approach have been discussed.

In the third chapter, the methodology was discussed under two main topics. First, the design science research approach was discussed. Secondly, the Crisp-Dm approach was discussed, and the research was detailed following this approach.

In the third and fourth chapters, the findings of the research were shared. This chapter consists of three main parts: the results of the ontological modeling, machine learning and artificial intelligence models, and the prototype.

CHAPTER 1. THEORETICAL BACKGROUND ON CAREER AND CAREER DEVELOPMENT

This chapter examines career and career development concepts, examining their evolution and significance in the modern job market. The chapter begins by defining the career and explores traditional and modern career development approaches, such as boundaryless careers. The chapter then discusses individual career planning and organizational career management, highlighting the interaction between the two. Next, the chapter discusses individual career planning by explaining two main approaches. Additionally, lifelong learning and life-span, life-space theories were discussed within the scope of career planning. Lastly, the chapter discussed the use of artificial intelligence in career development, the role of ontology in explaining competencies and skills, and, therefore, the role of ontology in career development and career planning.

1.1.Career and New Career Approaches

A career is the success and progression that individuals achieve through psychological growth, professional development, and the accumulation of work experiences over time (Cumbler et al., 2018). Arthur (2008) defines the career as a series of an individual's work experiences that develop and evolve progressively. This definition emphasizes the progressive nature and the work-related experience in an individual's professional journey. The perspective that defines a career as a sequence of challenges and opportunities for individual development emphasizes the importance of acquiring skills, knowledge, and competencies in career advancement. In this context, career success is closely related to the acquisition of competencies and skill sets appropriate to one's goals and values (Kuijpers et al., 2006). According to a comprehensive definition, a career is a process linked to work experience and activities throughout one's life, progressing and evolving in line with both organizational and personal objectives (Hall, 1986).

The concept of career, which was primarily positioned in an organizational-centered perspective from the late 1950s to the 1980s, evolved to be recognized as both an individual and organizational concept from the late 1980s to the present day (Ardıç & Özdemir, 2017; Aytaç, 2006; Hall, 1986). In this evolving and expanding context, the concept of career is defined as the cumulative sum of career-related events that individuals encounter over their lifetimes, extending beyond their working lives and professions (Super, 1980). In this regard, a career is defined as the lifelong process of crafting one's own career pattern, integrating work and life roles, and revealing the differences that stem from one's abilities, interests, and values; and the growth that an individual experiences through this proves is defined as career development (Herr et al., 2004; McDaniels & Gysbers, 1991; Yeşilyaprak, 2019, 2022).

With the development of information technologies and the rapid evolution of job roles, the concept of a boundaryless career has emerged. This approach emphasizes a career path that extends beyond a certain employer, company, or sector and that is shaped by the individuals rather than being dictated by the employer (Arthur, 1994; Forret & Sullivan, 2002). The concept of a boundaryless career marks a departure from traditional organizational career structures. It highlights independence from conventional career paths and concentrates on opportunities that extend beyond the confines of a single employer (Sullivan & Baruch, 2009). Boundaryless careers have become increasingly significant in today's context, where the traditional stable employment phenomenon is on the decline. As traditional career boundaries and patterns dissolve, career paths become more flexible and adaptable. Boundaryless career challenges rigid assumptions about roles and career paths that can hinder organizational and personal progress (Arthur, 1994). In fact, it argues that mobility across career borders can foster innovation and development by combining diverse knowledge and experience.

Boundaryless career indicated three main shifts in career perspectives: (1) the shift in individual's interest away from high-paying, high-status jobs and towards satisfying personal goals and emphasizing a work-life balance, (2) the shift in individual's interest away from acquiring organization-specific skills and towards transferable skills, (3) the shift in individual's loyalty away from the organization and towards dedication to one's profession (Forret & Sullivan, 2002).

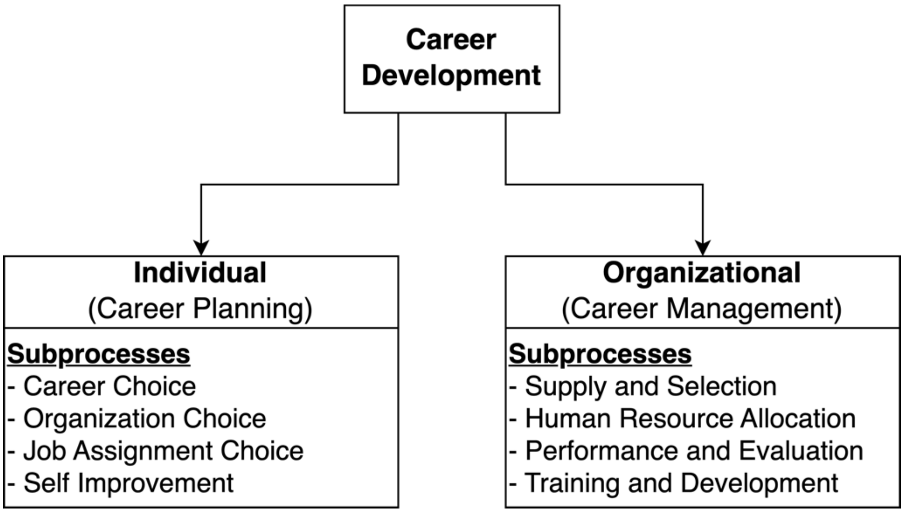
The boundaryless career challenged traditional notions by drawing validation from sources outside the present employer, emphasizing the importance of external and diverse experiences and extra-organizational networks and information (Arthur, 1994). This challenges the idea of career progression solely within the confines of a single organization. By breaking traditional organizational career boundaries, such as hierarchical reporting and advancement principles, boundaryless careers promote flexibility and adaptability in career paths, thus challenging linear and rigid career orıgression in the confines of an organization.

1.2.Career Development and Individual Career Planning

Career development is a process that benefits both individuals and organizations (Herr & Shahnasarian, 2001; Leibowitz et al., 1991; Van Dijk, 2004). It arises from the interaction between individual career planning and organizational career management (Hall, 1986). Career development is explored under two primary categories: individual career planning and organizational career management (Figure 1) (Erdoğmuş, 2003).

Figure 1

Career Development



Source: Erdoğmuş (2003)

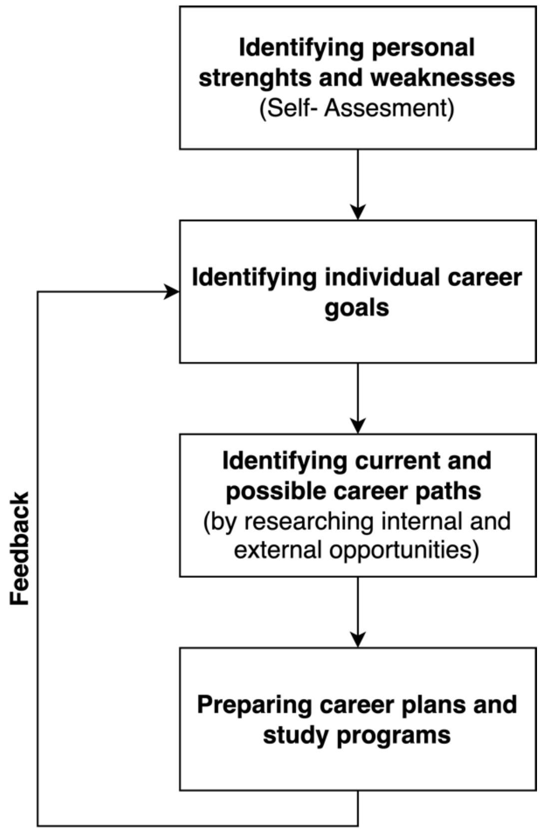
Career planning is the initial step in the career development process (Granrose & Portwood, 1987; Hall, 1986). Career planning is the process by which individuals (1) become aware of opportunities, options, and constraints relevant to achieving a specific goal, (2) establish career-related objectives and goals, and (3) organize and schedule work, education, and other related developmental activities (Hall, 1986; Leibowitz et al., 1991).

From an individual career perspective, career planning is a process that is centered on the individual rather than the position (Özgen & Yalçın, 2017; Quigley & Tymon, 2006). It involves individuals strategizing to achieve their career goals, taking into account their skills, interests, and values (Vergiliel Tüz, 2003). The definition and assessment of a career are highly individualized, reflecting how individuals value their career paths

(Smith-Ruig, 2009). The individual career planning process consists of five steps. According to Anafarta (2001), these steps are: (1) identifying personal strengths and weaknesses through a self-assessment, (2) identifying individual career goals, (3) identifying current and possible career paths by researching internal and external opportunities, (4) preparing career plans and study programs, (5) creating feedback where the individual evaluates their achievements to date (Figure 2).

Figure 2

Individual Career Planning Process



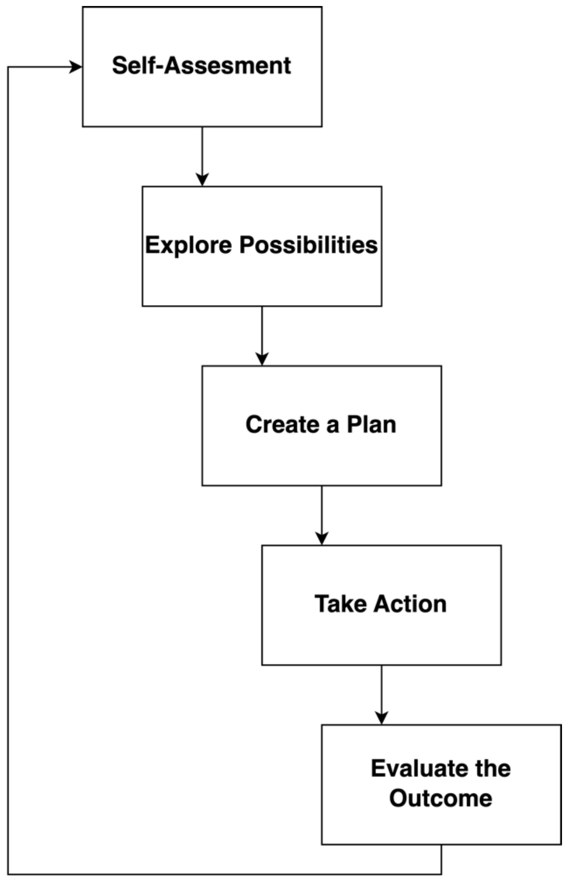
Source: Anafarta (2001)

According to Jaffe and Scott (1991), the five stages of career planning are (1) self-assessment, (2) exploring possibilities, (3) creating a plan, (4) taking action and (5) evaluating the outcome (Figure 3). In the first phase, self-assessment is important as it allows individuals to evaluate their skills, values, and interests. This helps people get a better understanding of themselves and make informed decisions about their careers. In the second phase, while exploring possibilities, individuals investigate different career options, different industries, and different job positions to identify potential opportunities that align with their knowledge and characteristics. In the third phase, creating a plan, individuals set specific career goals and timelines to achieve their goals, and they determine the steps necessary to achieve them. The fourth phase involves taking action to realize the plans that one has created. This process includes applying for jobs, networking, gaining relevant experience, and continuously learning and growing. Finally, evaluating the outcomes involves reflecting on achievements and challenges, updating career development plans, and repeating the individual career development cycle through self-assessment (Hansen et al., 2016; Jaffe & Scott, 1991).

Both Anafarta's (2001) and Jaffe and Scott's (1991) five-step approaches provide comprehensive frameworks for individual career planning. This theoretical background is utilized in this study to inform the development of the AI-based career planning system, ensuring that the system supports users through each critical phase of their career development process.

Figure 3

Five Stages in Career Planning



Source: Jaffe & Scott (1991)

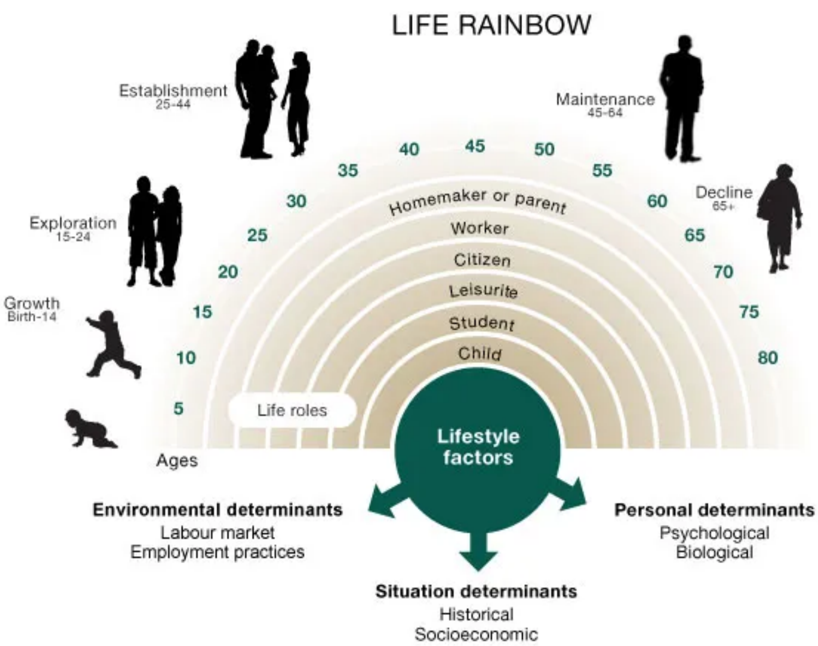
1.3.Lifelong Learning

Individuals navigating boundaryless careers often engage in continuous learning and skill development to adapt to diverse work environments and seize opportunities beyond traditional organizational boundaries. Positioned beyond organizational boundaries and not accepting traditional career paths, a boundaryless career requires a proactive approach to learning and skill acquisition. This enables individuals to effectively position themselves in a dynamic and ever-changing professional environment (Sullivan, 2001). Lifelong learning plays a crucial role in enhancing and acquiring competencies that are also essential for success in a boundaryless career (Guan et al., 2019).

Lifelong learning also plays a pivotal role in individual career development, particularly in the acquisition of new skills, knowledge, and competencies. This continual engagement aims to open up diverse career opportunities (Enache et al., 2013). Lifelong learning serves as a cornerstone, guiding the healthy sustainability and development of individuals' careers and helping individuals remain competitive in the labor market (Forrier et al., 2009; Volmer & Spurk, 2011).

Figure 4

Life-Span, Life-Space Theory's Life Career Rainbow



Source: Super (1980)

Career development is defined as a development process that spans from childhood to retirement (Super, 1963). Within this context, career development can be associated with lifelong education. In the life-span, life-space theory, stages of life are identified as a life career rainbow (Figure 4). These phases (life-span) include the growth phase from ages 4 to 14, the exploration phase from ages 14 to 24, the establishment phase from ages 25 to 44, the maintenance phase from ages 45 to 64, and the decline phase from ages 65 and beyond (Super, 1980).

The growth phase is characterized by the development of the concept of self and the general world of work. The exploration phase is a period where temporary choices are made, and skill development happens. During this phase, individuals experiment with different courses, hobbies, and even jobs. The establishment phase is primarily focused on skill development and work experience. In the maintenance phase, individuals are in the process of making adjustments to improve their career positions. The decline phase is defined as the period leading up to retirement, during which an individual's productivity and output generally decrease (Super, 1980). Individuals also take on different roles in this process (life-space). According to Super's theory, these roles are mainly defined as parent or homemaker, worker, citizen, leisurite, student, and child (Super, 1980). This study will mainly focus on the life-span part of Super's theory.

The life-span, life-space theory seeks to explain the career development process of individuals through different roles individuals take throughout their lives (Scholl & Cascone, 2010). It provides a framework for career planning and counseling by focusing on how one's life progresses through various career development processes (Perrone et al., 2006). Super's life-span, life-space theory conceptualizes a career as a sequence of positions held across one's life, encompassing before employment, during their working years, and after retirement (Perrone et al., 2006). Super's theory has influenced the understanding of career decisions, transforming career management from a static process focused on a specific path to a dynamic process that necessitates flexible strategies (Sterner, 2012). Lifelong learning plays a crucial role in individual career development by continuously enhancing skills and knowledge to adapt to the dynamic job market requirements. Career counseling in individual career planning provides the necessary guidance and support to navigate this ongoing learning process, helping individuals to plan and achieve their career goals strategically.

This study focuses on individuals who are in the establishment phase of their professional careers, typically ranging from ages 25 to 44 according to the life-span, life-space theory.

1.4.Career Counseling and Individual Career Planning

Career counseling provides support and guidance to help individuals make informed career decisions, enhances adaptability, and enables them to navigate the complex and volatile labor market. Career counseling is the process of identifying one’s goals and starting to take action to achieve them (Osborn & Baggerly, 2004). Career counseling aims to enable self-understanding and informed career decisions (Osborn & Baggerly, 2004). Career counseling encompasses career development, positioning in the labor market, and strategic career design, and it goes beyond only vocational guidance (Savickas, 2014). It aims to support career planning by helping individuals overcome potential obstacles to becoming skilled employees, and in the process of doing so, it seeks to identify one’s values and mindset, which are essential for effective career development (Racene et al., 2019).

Individual career counseling has proven to be beneficial and effective in guiding people’s career journeys (Lau et al., 2020). It aims to help individuals gain a deeper understanding of themselves, recognize available opportunities, and increase their readiness for work, thus enabling individuals to make well-informed decisions based on this insight (Behrendt et al., 2019). It also plays a significant role in boosting individuals’ employability by enabling them to understand their own skills and competencies and encouraging them to focus on development (Behrendt et al., 2019). The primary goals of career counseling include (1) enhancing self-awareness, (2) increasing awareness of opportunities, (3) making career decisions, (4) finding a new job, (5) improving work-life balance, and (6) enabling better work relationships (Brott, 2001; Verbruggen et al., 2016).

One specific approach to career counseling is the constructivist approach. It emphasizes acquiring transferable skills and centers on the idea that individuals actively shape their career paths by developing and applying these versatile transferable skills (Miller, 2004). This approach recognizes the importance of developing skills that are transferable across various career fields. Constructivist approaches in career guidance and counseling focus on brief, positive strategies that support individuals in overcoming career-related challenges and advancing toward their goals. This model allows career counseling practitioners to apply a solution-focused strategy that enables their clients to navigate career transitions and make informed decisions. It connects career counseling with the individual's goals, making the career development process logical and practical while encouraging relevant (and progressive) steps to be taken. (Miller, 2004).

The constructivist approach offers a three-phase framework for the counseling process: (1) problem clarification and goal setting, (2) solution building, and (3) constructing meaningful homework. The first phase begins by understanding the individual's problem. From there, it proceeds to the goal setting in line with the individuals' goals. It shifts the focus of the discussion from the problem. The focus of the discussion shifts from problem discussion to solution exploration through goal-setting questions that uncover desired outcomes from the process. The second phase focuses on drawing on past experiences to strategize a solution and evaluate options. The third phase involves individuals taking actions in line with their goals. During this phase, the solution plans developed in the second phase are realized, and the progress is monitored (Miller, 2004).

This study argues that the three-phase framework of counseling can be enhanced and improved through the use of information systems. In the first phase, during the problem identification and goal setting, information systems can analyze an individual's career data to determine how they are poisoned in the current job market. In the second phase, the information system can assist in planning solutions to advance the individual's career. For the third phase, information systems can be used to plan future actions that should be taken in their career development process, enabling them to effectively plan for the future.

Career counseling in individual career planning offers personalized guidance and support to help individuals make informed career decisions and achieve their professional goals. Integrating artificial intelligence into career development enhances this process by providing data-driven insights and recommendations, thereby making career planning more effective and accessible.

The discussion of career counseling and the integration of information systems show the need for the exploration of human capabilities, competencies, and career data. The subsequent discussion focuses on these, as well as the necessity of a comprehensive ontology to represent and utilize career-related information accurately. The ontological approach is crucial for enhancing the precision and effectiveness of career counseling and individual career planning. By establishing a robust framework to classify and understand various aspects of human capabilities, competency, and career-related data, career counseling, and individual career planning services can be better aligned with individual needs.

1.5.Ontological Foundations of Competency

Competency is a combination of skills, knowledge, and behavior necessary to perform a job effectively. Competencies are dynamic attributes that are associated with an individual's or organization's capability to perform specific tasks and activities. They evolve over time as the individuals gain experience and knowledge (Bergenhenegouwen, 1996). An individual's competencies continuously change in conjunction with their professional development and through ongoing learning processes, which are crucial in career development. Emphasizing competencies fosters a more comprehensive integration of formal education, vocational training, and professional development, aligning seamlessly with lifelong learning strategies (Azevedo et al., 2015; Le Deist & Winterton, 2005; Miranda et al., 2017). Therefore, competencies, which continuously evolve through an individual's experiences, form the foundation of broader human capabilities and play a significant role in both personal and professional development.

The career data of individuals, which includes the positions they have held, their skills, educational background, professional competencies, and career-related achievements, is crucial for career planning and development. This information helps both individuals and organizations make informed decisions regarding career trajectories and professional growth. Effective analysis of career data can offer a strategic approach to human resources management and career development processes, ensuring alignment between individual aspirations and organizational needs.

Competencies and skills are crucial elements in both the theory and practice of career development. Thus, accurately modeling these aspects is essential for effective career planning and management (Dneprovskaya et al., 2022; Kuijpers et al., 2006; Schmidt & Kunzmann, n.d.). In this context, the ontology-driven approach provides a structural framework to capture the complexity and dynamics of human competence in the labor market. Modeling career data, which encompasses human capabilities and competencies, allows for the systematic capture of various attributes, such as skills, experience, and knowledge, that individuals manifest in their resumes and actively use in their professional roles (Calhau et al., 2021; Calhau & Almeida, 2022; Fazel-Zarandi & Fox, 2012; Miranda et al., 2017). Within this framework, Human Capital Theory underscores the importance of skills and knowledge as pivotal economic resources, emphasizing the necessity of modeling to optimize career outcomes and facilitate organizational growth (Aliaga, 2001; Nafukho et al., 2004). In this context, ontologies offer a robust framework for categorizing and defining competencies, capabilities, and other related career data. This representation facilitates understanding the relationships and dependencies.

The benefits of using ontology to model a domain are manifold. Firstly, ontology-based models facilitate interoperability and standardization (Corcho et al., 2006; Falbo et al., 2002). In the context of this research, employing an ontology-based approach enables the uniform interpretation and comparison of competencies across diverse career positions. Through this approach, the possibility for advanced analytical capabilities is provided, enabling sophisticated data analysis and reasoning. This methodology supports the potential to uncover viable career paths and development opportunities for individuals. It aligns with Career Construction Theory (CCT), which emphasizes the necessity for adaptability and resilience in career development (Savickas et al., 2009; D. Wang & Li, 2024). In addition, the dynamic nature of Ontology-based structures allows for development, updating and modification to keep pace with emerging trends. This adaptability reflects the principles of the Boundaryless Career Theory (BCT), ensuring that career advice remains relevant over time (Arthur, 1994; Forret & Sullivan, 2002). BCT emphasizes a career concept in which individuals shape their own career paths, unrestricted by a single employer or sector.

This shift towards a more individualized and adaptable career approach aligns with the integration of advanced technologies. The adoption of these technologies, such as artificial intelligence, aims to enable individuals to navigate diverse career opportunities more effectively.

1.6.Artificial Intelligence in Career Development

In the comprehensive systematic literature review study, Pandya and Wang (2024) examined artificial intelligence in human resource development. Analyzing 101 published articles, they outlined the multifaceted effects of artificial intelligence on career development under seven main themes. The paper explores the implications of AI in career development from a positive and negative angle. On the positive side, AI applications empower individuals to develop themselves, increase efficiency, and reduce the cost of this process. This supports employees in their career development plans while facilitating the acquisition of skills, knowledge, and abilities. On the negative side, there are concerns about AI replacing human jobs, leading to workforce displacement. The study reveals that integrating AI into the career development process is well-received at the individual level.

Current developments in AI that transform traditional approaches can offer a scalable approach that appeals to a wider audience and improves system accessibility (Drewery et al., 2022; Pandya & Wang, 2024). The integration of AI into companies' career development process can enable upskilling and reskilling to meet the labor market's needs. However, studies reveals that these might increase employee turnover, as individuals can explore career paths freely, and they are more likely to evaluate and pursue new opportunities (Petkov, 2021; Presbitero & Teng-Calleja, 2023; Westman et al., 2021).

AI transformed job matching and recruitment by leveraging data analysis capabilities to streamline the process of identifying ideal candidates for positions. This enhanced efficiency and reduced the time-to-hire. In addition to benefiting employers, AI can also provide personalized recommendations to job seekers, identifying positions that match their qualifications and align with their unique career goals. This gives individuals self-confidence as they navigate the job market (Drewery et al., 2022; Wilson et al., 2022). The integration of AI and lifelong learning plays a significant role in people's career mobility and employability by enabling individuals to adapt to an ever-changing work environment. Skill gap analyses and recommendations for people's needs and preferences can increase people's motivation and engagement toward lifelong learning (Gouda, 2022; Lee et al., 2018; Wirtz et al., 2018). In AI-centered environments, there is a difficult balance between the collection of data for development and the privacy of individuals. It is important for the ethical use of data that AI systems do not make biased recommendations and that data is anonymized to ensure data security (Hagendorff, 2020; Kantar & Bynum, 2021).

Westman et al. (2021) state that studies on artificial intelligence systems developed to provide career guidance for individuals during their higher education and early career processes are limited. By investigating the use of artificial intelligence in the career guidance process in higher education institutions, this study highlights the potential of these systems. However, the study indicates that for this process to work effectively, students, staff, and educational institutions need to be well-informed about AI-based education.

Gedrimiene et al. (2024) investigated the advantages and challenges of AI-supported learning analytics applications to aid career decisions from the user's perspective. The study identified five key benefits of AI-based learning analytics tools. These include assessing career information, researching and analyzing it, differentiating between various career paths, providing decision support, and helping individuals conduct self-assessments. In addition, the study reveals that AI-based systems positively benefit individuals' career changes. These insights offer valuable guidance, particularly for the development and implementation of learning analytics applications.

There are recent studies in the literature on career development and supporting career decisions with information systems and artificial intelligence applications.

Tomy and Pardede (2019) developed the Map My Career to enhance university students' satisfaction by offering a career-oriented educational experience. The application aims to boost graduate employability by helping students develop core competencies. It links university courses with relevant skills and career prospects through text mining and data analytics methods. The study aims to support students' academic preparation, skill development, and workload management to increase their educational satisfaction. The study seeks to help students select courses aligned with their career interests by facilitating skill-job matching.

Ghosh et al. (2020) proposed an interpretable monotonic nonlinear state space model for career path modeling, aiming to predict company affiliation, job title, and talent using large datasets. This study evaluated only the jobs themselves in the career planning pathway and suggested that future research could also consider the duration of employment as a criterion for assessment.

In addition, there are many studies in the career development, planning and counseling literature that examine students' career processes (Çarkıt, 2019; Duffy & Sedlacek, 2007; İstanbullu Dinçer et al., 2013; Jackson, 2017; Özdemir & Kibar, 2018; Polat et al., 2016; Sevinç & Siyez, 2018; Tomy & Pardede, 2019; Waddell & Bauer, 2005; Willis & Wilkie, 2009). These studies often investigate the challenges students face when choosing a career path, the factors influencing their decisions, and the role of effective career planning guiding them toward fulfilling career trajectories.

However, despite the recent AI applications for career development and human resources development, there exists no ontology-driven approach that incorporates ontological rigor to provide the necessary adaptability, interoperability, and depth. Such an approach is essential for creating comprehensive, flexible, and ontologically well-founded career planning systems that can effectively integrate diverse data sources and adapt to the evolving needs of individuals and the job market.

Ensuring ontological rigor and a robust semantic infrastructure is crucial for developing successful AI systems, particularly in the realm of career development. In AI systems, ontology provides the foundational structure for data integration, interoperability, and semantic clarity. In AI-based career development and planning systems, diverse data sets can be effectively combined, enabling seamless management and utilization. The use of an ontological approach may allow these systems to capture the complexity and nuances of career data, ensuring consistent interpretation and meaningful retention across various systems and contexts (Gemino & Wand, 2005; Guizzardi & Proper, 2021; Olivé, 2007). Furthermore, ontological approaches allow AI systems to incorporate novel data sources effectively. AI applications grounded in ontological rigor ensure datasets' relevance and long-term usability, and they can easily adapt to changes within these datasets (Altan, 2022).

Ontology also enhances the scalability and sustainability of AI systems, allowing these applications to adapt to rapidly evolving industry demands, such as the frequent changes in career position requirements. This ability to swiftly adapt to changes is crucial for maintaining the relevance and effectiveness of AI-driven career development applications, which aim to represent a rapidly evolving job market accurately (C. Davis et al., 2020; Nicholson et al., 2023).

In summary, ontological rigor is essential in AI development, particularly when integrating novel data sources. Establishing solid ontological foundations is crucial for creating a comprehensive, flexible, and robust AI system capable of meeting evolving market needs.

CHAPTER 2. THEORETICAL BACKGROUND ON FOUNDATIONAL ONTOLOGY

This chapter examines the concept of ontology, Unified Foundational Ontology, Unified Foundational Ontology stereotypes, and OntoUML modeling language. The chapter starts with identifying the need for ontology through the example of false agreement. Then, the chapter explains ontology and Unified Foundational Ontology (UFO). The chapter then discusses the terms and stereotypes of UFO, which will be essential for the development of the ontology-driven conceptual model. Additionally, OntoUML, an ontologically well-founded modeling language that extends UML (Unified Modeling Language) with ontological distinctions, is explained.

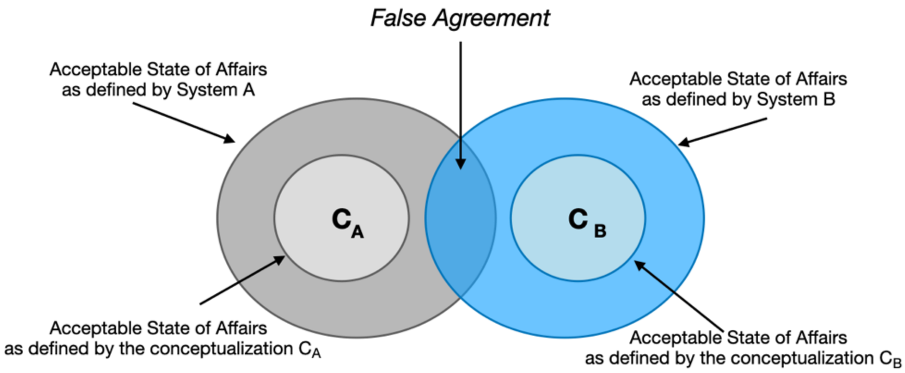
2.1. The Concept of False Agreement

An ill-defined domain and ambiguous definitions of classes can lead to uncertainties. The inability to model complex business problems and challenges encountered during the conceptualization may pose significant issues in the stages of system development and integration. While it may be presumed that the model is correct, the constraints of the method or the language can hinder the accurate and comprehensive definition of the problem. For instance, conventional methods of modeling and conceptualization often face challenges in effectively dealing with complex business problems. In order to explain the possible ambiguity further, it proves beneficial to refer to the construct 'False Agreement' that was introduced by (Guarino, 1998).

During system conceptualization and modeling, there is a risk of creating an insufficiently constrained model due to the lacking features, such as ontology and semantics, within the modeling language. Consequently, this can lead to overgeneralized inferences that do not correspond to the realities of the domain. This phenomenon, known as 'False Agreement,' creates the (false) assumption of agreement between two systems which, in reality, neither coincide nor concur (Figure 5). False Agreement may lead to significantly different concepts being incorrectly perceived as identical during the system integration (Guarino, 1998; Guizzardi, 2005).

Figure 5

False Agreement



Sources: Guarino (1998) and Guizzardi (2005)

Similarly, establishing more rigid constraints than that are required to conceptualize the systems or domains may lead to conceptual boundaries that are more constrained than those of the actual physical domain. This results in perceived, yet false, disagreement between the systems on certain concepts due to an over-restrictive approach despite the systems being aligned (Guarino, 1998). It is argued that false agreement in the IS field of study and practice can be fixed by using Ontology-Driven Conceptual Modeling (ODCM) languages to model wicked business problems.

2.2. Ontology

Ontology can be described in a philosophical or a computer science sense. Ontology, in a philosophical sense, is a theory about the kind of entities and their ties that are assumed to exist by a given description of reality (Guizzardi, 2005). It refers to the study of exploration of the nature of existence in a general sense. On the other hand, as a computer science term, ontology is an area devoted to developing domain-independent toolboxes with tools for supporting ontological analysis (Masolo et al., 2003; Smith & Welty, 2001). Within the latter context, ontology refers to the examination of the entities and their relationships that exist in a specific physical domain or a universe of discourse. Ontology provides a solid theoretical foundation for conceptual modeling. It ensures that conceptual models can accurately represent the complexity and nuances of real-world entities and their relationships.

2.3. Conceptual Modeling

Concepts are abstract notions that are expressed in a way that can be understood by people who understand a certain lexical language. Expressing a concept in a lexical language hinders deeper understanding and communication. These concepts must be expressed and communicated as artifacts to be understood by all stakeholders and consequently analyzed and debated upon (Guizzardi, 2005). Conceptual modeling is referred to as a collection of specification statements relevant to a problem (Lindland et al., 1994), and they are used by information systems professionals to express the aspects of a domain (Fettke, 2009).

Conceptual modeling aims to build a formal representation of a modeling domain (Fettke, 2009). It ensures that the models accurately represent real-world entities, often graphically (Fettke, 2009), without depending on a specific design, a specific systematical or technological choice, or a lexical language (Mylopoulos, 1998). Conceptual modeling represents an integral and foundational field for information systems and computer science (Fettke, 2009; Lindland et al., 1994; Olivé, 2007). It influences various computer science areas, including, but not limited to, database design, software engineering, information integration, semantic interoperability, and natural language processing (Guizzardi, 2005). Even though it is widely used in the field of computer science, it is often discussed that conceptual modeling requires improvements in areas such as interpretability and complexity management (Falbo et al., 2002; Fonseca et al., 2022; Guizzardi, 2005; Guizzardi et al., 2015; Mylopoulos, 1998). The ODCM approach has been developed to address this gap, integrating conceptual modeling with ontology.

2.4. Unified Foundational Ontology (UFO)

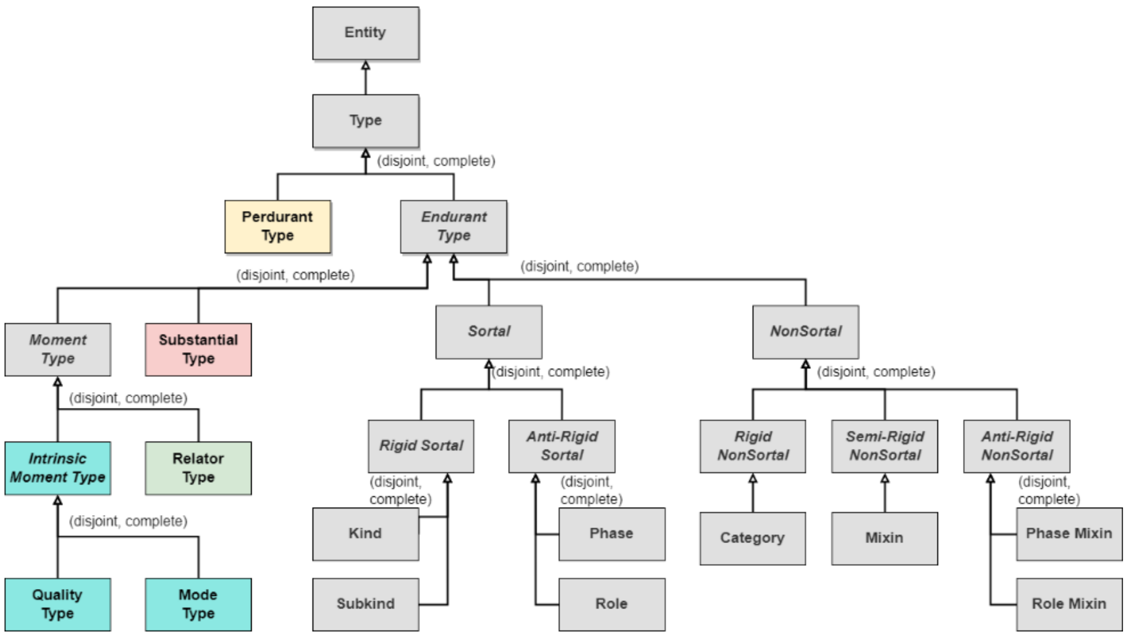
Ontology, in a computer science sense, plays an important role in ODCM. ODCM is a scientific and systematic approach that is used in creating, interpreting, and analyzing conceptual models (Guizzardi, 2005; Guizzardi et al., 2015; Mylopoulos, 1998). This approach is based on principles of foundational ontologies, in the case of this study, specifically, UFO. UFO represents the underlying structures that represent the categories of things.

UFO enables researchers, practitioners, modelers, and analysts to accurately represent the physical domain. UFO uses stereotypes to present a detailed representation of real life and the physical domain. UFO categorizes entities under endurant types and perdurant types (Figure 6).

Endurant types maintain their identity and existence over time while possessing different properties and relationships. On the contrary, perdurant types are referred to as events or processes that occur within a specific time period. Their existence is temporary within that time period (Guizzardi, 2005; Guizzardi et al., 2015). Both endurant and perdurant types hold significant importance within the field of conceptual modeling. However, the emphasis of this particular study will be focused on endurant types and their further classification.

Figure 6

UFO Stereotypes



Sources: Guizzardi (2005) and Guizzardi et al. (2015)

Endurant types are divided into sortal, non-sortal, and also moment types. Sortals are endurant types that carry specific identity and classification principles. They define 'what it means for an entity to remain the same over time?', and they can answer the question of 'what is it that makes this object what it is?'. Non-sortals group and categorize entities based on their common characteristics and attributes without providing a specific identity. However, they don't determine what makes something the same type over time, and they don't necessarily answer the previously mentioned question that sortals do answer. For example, the non-sortal type "Red Things" includes any object that is red, such as a red apple, a red car, or a red book. Lastly, moment types provide a structure for representing characteristics, traits, and circumstances that rely on other entities. The existence of moment types and substantial types are both dependent on other entities (Fonseca et al., 2022; Guizzardi, 2005).

Sortals and non-sortals are also divided into different categories in terms of their rigidity. An entity that is defined as rigid always carries its identity and does not exhibit a change in its inherent identification. No matter what changes, a rigid entity will never lose or change its identity. For example, a human will always be a human even if time passes or it is affected by the outside world. Anti-rigid type, on the other hand, is the opposite. An anti-rigid entity does not necessarily carry its identity, allowing for the possibility of change. Anti-rigid entities carry their identity only under certain conditions or during specified time intervals. As an example of anti-rigid, a person's status as a student can be considered. A person doesn't start life as a student and doesn't remain a student indefinitely. This status relies on particular relationships and conditions that can change or disappear over time. In such a case, that individual would no longer be identified as a student (Fonseca et al., 2022; Guizzardi, 2005).

In order for us to understand and work with UFO, it is necessary to go over some of the UFO stereotypes that will be discussed within the scope of the study. In this section, stereotypes that are considered an endurant type will be explained (See Table 1). This insight will provide a better understanding of the concept.

2.4.1. Sortals

Sortals are a category of endurant types with distinct principles for identity and classification. They establish the concept of entity's persistence over time and addresses the fundamental question of what defined the essence of an object. Sortal's ability of answering this question enables us to be precise and eliminate uncertainties when modelling. They play a vital role in ODCM by providing a structured, systematic and clear way to represent entities. This leads to more accurate, coherent and semantically rich models. In the context of UFO and withing the scope of this study, the sortal stereotypes identified are kind, subkind, phase and role.

Kind is a rigid sortal type that exists independently and doesn't change its identity through time or across different contexts. Identifying kind helps to understand the essential and unchanging elements in a domain, providing a stable foundation for the model. In a conceptual specification, every entity (or object) must be either directly or indirectly an instance of a Kind. In addition, an object cannot simultaneously be an instance of more than one kind.

Kinds can be further detailed as **subkinds**, which are specialized rigid subtypes that embody the attributes of the kinds. For instance, a person can be considered a kind, while its subkinds could include biological male and biological female.

Role is an anti-rigid sortal type of thing that is context-dependent and typically temporary. Roles are inherently dynamic and depend on certain conditions or relationships to exist. Roles allow us to capture the dynamic and context-dependent aspects of entities, making models flexible and responsive to different situations. For example, being a student can be defined as a role. A person, which is an instance of a kind, can be defined in the role of a student as long as they are enrolled in an educational institution. If the required relationships and conditions are no longer valid, the person won't possess the role of a student any longer. This arises from the anti-rigid nature of the role.

Phase is another type of anti-rigid sortal type. Phases represent the natural state of entities at various stages of their lifecycle. Unlike roles, phases don't rely on the existence of a relationship. For example, a butterfly has a certain lifecycle that progresses through its life. In this case, the phases of the butterfly entity can be defined as egg, larva, pupa, and adult butterfly (Guizzardi, 2005).

2.4.2. Non-Sortals

Non-sortals classify entities by their shared characteristics and attributes without assigning individual identity. Unlike sortals, they do not incorporate the concept of an entity's enduring existence over time. Non-sortals offer flexibility and adaptability in ODCM. It is particularly useful when dealing with entities that don't follow strict rules about identity and it is advantageous to group entities on their attributes and roles. Non-Sortals are also useful for adapting a wide range of needs, especially when the rigid rules that sortals impose don't fit the situation. For instance, consider the non-sortal type "Furniture." This category encompasses items such as chairs, tables, and sofas, which share common characteristics as objects used to furnish spaces. However, "Furniture" does not provide a specific identity or address what maintains the continuity of a particular chair's identity over time. Non-Sortals that are going to be covered in this stud are mixin, role mixin, phase mixin and category.

Mixins are non-sortal types that enable more flexible and nuanced conceptual modeling by providing the ability to represent similarities and associations between different entities without oversimplifying the relations. Mixins allow for modeling entities that differ in type but possess shared properties or characteristics.

Role mixin is a special type that represents a role that can be played by individuals of different kinds. It's a way to capture roles that are shared across several kinds, without having to specify them separately for each kind. Role mixin helps in generalizing common roles across different kinds, reducing redundancy and improving the conceptual clarity of the models. For instance, a customer can be modeled as a roleMixin if there exists an individual customer (kind of Person) and a corporate customer (kind of Organization).

Phase mixin is used to define a phase that belongs to different kinds with certain characteristics. To illustrate, the functional state of a vehicle, such as broken or operational, can be considered as a phase mixin. This definition of a phase mixin can apply to different kinds of entities, such as a mobile phone or a car, but the functional state of being broken or operational is a shared attribute between them (Guizzardi, 2005).

Category is used to define and group a set of distinct entities that, despite their differences, share a common meta-property or characteristic. The definition of category stereotype applies across different kinds of entities, but the shared meta-property unites them under a common conceptual umbrella (Guizzardi, 2005). Categories are essential in complex modeling cases where entities are not strictly confined to a single, well-defined class.

2.4.3. Moments

Moment types offer a framework for modeling characteristics and conditions that are dependent on other entities. Similar to substantial entities, the existence of moment types is tied to other entities. This adds richness and nuance to conceptual modeling and offers a way to model unique relationships and characteristics of specific domains. The moment-type stereotypes that will be covered within the scope of this study are relators, modes, and qualities.

Relator is a type that mediates a relationship between other types, often roles. It typically represents a fact or situation that links things together. Relators help encapsulate relationships and link entities together, giving us a more nuanced understanding of how entities interact within the domain. For instance, the enrollment of an student in an educational institution can be modeled as a relator.

Mode Type is used to model an entity's intrinsic or dispositional abilities and competences. These abilities and competences may or may not have yet been realized, and they depend on various conditions and contexts. For instance, a person's ability to speak a foreign language can be modeled as a mode. This is an intrinsic competency that the individual possesses and can be evident when speaking that particular language.

Quality Type represents an object's intrinsic characteristics with a specific measurement. Quality type is generally paired with a measurement unit. An example of a quality would be the height of a person (Guizzardi, 2005).

In summary, the exploration of UFO stereotypes reveals their foundational significance in ontological modeling. As these findings are integrated with the objective of revealing the potential of ODCM in the field of IS, their relevance becomes apparent. UFO stereotypes with their emphasis on precise categorization and identity standards, align with ODCM principles. By leveraging the use of these stereotypes, it is possible to be positioned well to tackle the unique challenges encountered in IS, enabling more agile and effective system development, integration, and knowledge management practices.

The motivation for employing an ontology-based approach in this study is multifaceted. Traditional methods of modeling and conceptualization often struggle to communicate and manage complex business problems effectively. Ontology-Driven Conceptual Modeling (ODCM), however, offers advantages beyond conventional techniques by emphasizing the semantics of the domain being modeled. The use of ODCM ensures semantic integrity and precision, which prevents ambiguity and misunderstanding among stakeholders, particularly in large and distributed systems (Corcho et al., 2006; Falbo et al., 2002). Grounded in Unified Foundational Ontology (UFO), ODCM models are reusable and interoperable, enhancing their applicability and reliability across various domains (Unschold & Gruninger, 1996). This approach allows for better validation and verification, improving the accuracy and comprehensiveness of the models and systems developed.

Table 1

UFO Terms and Stereotypes

UFO Terms & Stereotypes	Brief Definition
Endurant Types	Maintains their identity and existence over time while possessing different properties and relationships
Perdurant Types	Events or processes that occur within a specific time period
Sortal	Endurant types that carry specific identity and classification principles. Defines: what it means for an entity to remain the same over time?
Non-Sortal	Groups and categorizes entities based on their common characteristics and attributes without providing a specific identity.
Moment Type	It offers a framework for modeling characteristics and conditions that are dependent on other entities.
Rigidity	It always carries it identity and does net exhibit a change in its inherent identification
Anti-Rigidity	It doesn't necessarily carry its identity, allowing for possibility of change. They may carry their identity only under certain conditions and for specific time intervals.
Kind (Rigid, Sortal)	It exists independently and doesn't change its identity through time or across different contexts.
Role (Anti-rigid, Sortal)	It is context-dependent and typically temporary. Their existence is dependent on certain conditions or relationships to exist.
Phase (Anti-rigid, Sortal)	It represents the natural stage of entities at various stages of their lifecycle.
Role Mixin (Anti-Rigid, NonSortal)	It is a role that can be played by individuals of different kinds.
Phase Mixin (Anti-Rigid, NonSortal)	It is a phase that belongs to different kinds with certain characteristics.
Mixin (Semi-Rigid, NonSortal)	It represents similarity and associations
Category (Rigid, Sortal)	It is a used to define a set of entities that share a common meta-property but belong to different kinds.
Relator	It helps encapsulate relationships and links the entities together
Mode Type	It represents intrinsic ability
Quality Type	It represents intrinsic characteristic

Source: Guizzardi (2005)

Traditional approaches fail to address the dynamic and complex nature of individuals' competencies and skills. These methods lack the ability to provide a comprehensive and systematic representation of career data, which includes skills, experiences, and qualifications (Altan, 2022). Such shortcomings lead to a limited and incomplete understanding of an individual's capabilities and limit the effectiveness of career planning and development strategies (Miranda et al., 2017). Ontology, with its ability to capture and represent complex relationships and structured frameworks (Corcho et al., 2006; Falbo et al., 2002), offers a robust solution to this problem.

This study aims to overcome the limitations of traditional approaches by employing an ontology-based framework. Ontology provides standardization and ensures a common interpretation of competencies across different positions, facilitating comparisons and skill transfer (Paquette et al., 2021). This approach aligns with the principles of lifelong learning and the boundaryless career approaches, supporting continuous improvement, which is critical for career development.

The current literature highlights the lack of ontological rigor as a significant gap in AI-based systems, particularly in AI-based career development systems (Boudi et al., 2020). In Chapter 1, the applications of AI in career development were discussed, revealing that many studies lack ontological foundations (See Chapter 1). This may lead to inconsistencies and limited adaptability. This study aims to address this research gap by developing a comprehensive, flexible, and robust framework that adapts to the ever-changing, dynamic labor market requirements. This will be achieved by grounding the AI-based career planning system in ontological principles.

2.5. OntoUML

Foundational Ontology, UFO in this case, forms the basis for OntoUML, providing a set of conceptual categories that are ontologically well-founded and thus can more accurately represent the domain (See Table 1). OntoUML is an ontologically well-founded modeling language that extends UML with the ontological distinctions from UFO (Guizzardi et al., 2015). OntoUML addresses some of the limitations of traditional modeling approaches by making ontological distinctions. Foundational Ontology and OntoUML enable a more precise representation of all domains. They address some of the limitations of the conventional modeling approaches by making ontological distinctions explicit (Guizzardi, 2005).

CHAPTER 3. RESEARCH METHODOLOGY

This chapter outlines the research methodology employed in this study, focusing on the systematic approach used to investigate and develop the AI-based career planning system. The chapter begins with an explanation of the Design Science Research (DSR) methodology, which guides the research design and the creation of artifacts for the research. The chapter then describes the CRISP-DM framework used for the data science process, encompassing steps such as business understanding, data understanding, data preparation, modeling, evaluation, and deployment. DSR and CRISP-DM create the main structure for this chapter. Additionally, the chapter discussed the data sources, including professional social media platforms (PSMPs) data, and the methods of data collection, data preparation, feature engineering, and finally, AI/ML model development. This process is closely integrated with the ontology-driven approach, whose theoretical background and methodology were discussed in Chapter 2, and the outputs of these combined processes are presented in Chapter 4.

3.1.Design Science Research

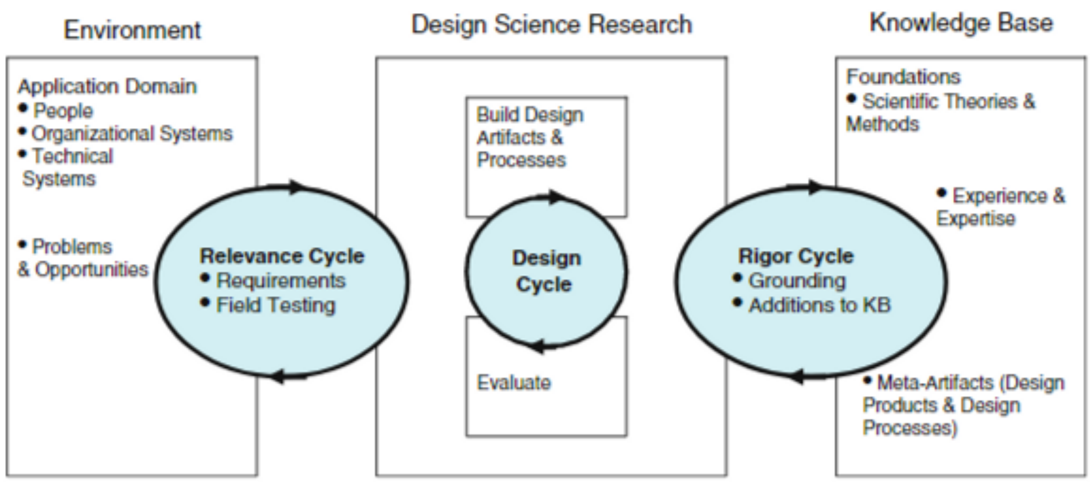
Information systems is a unique research field that integrates ideas, concepts, and methods from various areas. The constantly evolving hardware, software, and user interfaces in information systems pose specific design challenges. The design science research approach addresses these challenges and shares the innovative artifacts it discovers with the scientific community. Nunamaker et al. (1991) and Simon (1996) brought Design Science Research and Artifact Development to the attention of the academic community. Later, Hevner's article published in MISQ in 2004 clearly explained DSR's place within the MIS discipline (Hevner et al., 2004). Hevner et al. (2004) introduced the Design Science Research paradigm in their influential paper, highlighting its relevance and application in Information Systems research projects. They also defined artifacts within this paradigm, categorizing them as constructs, models, methods, or instantiations.

Over the years, more methods have been published outlining specific ways to implement design science research (Hevner, 2007; Vaishnavi & Kuechler, 2007). Design science offers an effective way of tackling the relevancy gap that has been a problem for the Management Information Systems research area (Hevner & Chatterjee, 2010). As a result of the field studies and iterations that occur in DSR, the relevance of the study can be expressed more clearly, and the relevancy gap can be identified.

Design Science emphasizes problem-solving and aims to introduce evaluated artifacts. It is a paradigm in which a designer responds to questions relevant to human problems by creating innovative artifacts to answer questions that contribute to the body of scientific evidence (Hevner & Chatterjee, 2010). Artifacts are the outputs produced within the scope of DSR projects, and they aim to solve a problem or improve existing solutions. A designed artifact can serve as a tool and a basis for understanding the problem (Hevner & Chatterjee, 2010). Artifacts can be constructs, models, methods, or instantiations. Constructs consist of vocabulary and symbols required to express better DSR problems, processes, and solutions (Boland Jr., 2022; Hevner et al., 2004). Effectively expressing the problem is crucial to finding a solution that works (Weber 2003). Models include abstractions and representations of ideas. These abstractions and representations enable a deeper understanding of the design problem and its solution. Models help to understand the problem and solution by representing their relationship. Instantiations show that constructs, models, and methods can be implemented in a running system. It allows us to see the viability of a solution, its suitability for its planned purpose, and the relationship of the artifact with the real world and its users. DSR studies in the IS field often deal with information systems design. Therefore, instantiations emerging in this field can be prototypes of software tools aimed at improving IS processes (Hevner et al., 2004).

Figure 7

Design Science Research Cycles



Source: Hevner (2007)

Hevner et al. (2004) defined the Design Science Research paradigm in their 2004 MISQ paper and discussed its use in IS research projects; however, they did not share how DSR should be done in this study. Later, they explained the three cycles of design science research in his 2007 study (Hevner, 2007). These three cycles are Relevance, Rigor, and Design (Figure 7).

The Relevance Cycle establishes the relationship between the design science activities and the research project's contextual environment. It initiates the design science research process. The application domain provides research requirements as input and defines acceptance criteria for final evaluation (Hevner, 2007).

The design cycle is at the center of DSR projects. At this stage, an artifact is generated and evaluated, and feedback is given iteratively to improve the design further. This cycle consists of comparing and evaluating alternatives until they meet the requirements and repeating the process until a satisfactory design emerges. In the design cycle, requirements are inputs retrieved from the relevance cycle. Design theories and evaluation theories are inputs retrieved from the rigor cycle. The design cycle is interconnected with the other cycles in these points (Hevner, 2007).

The rigor cycle enables studies to benefit from past knowledge. In this way, current knowledge can be referenced throughout the study, thus highlighting the benefits of the research and its difference from routine professional design application. Rigorous advancements and improvements in design distinguish a DSR project from routine professional design practice. Therefore, it is important to determine the appropriate source of rigor for the research project (Hevner, 2007).

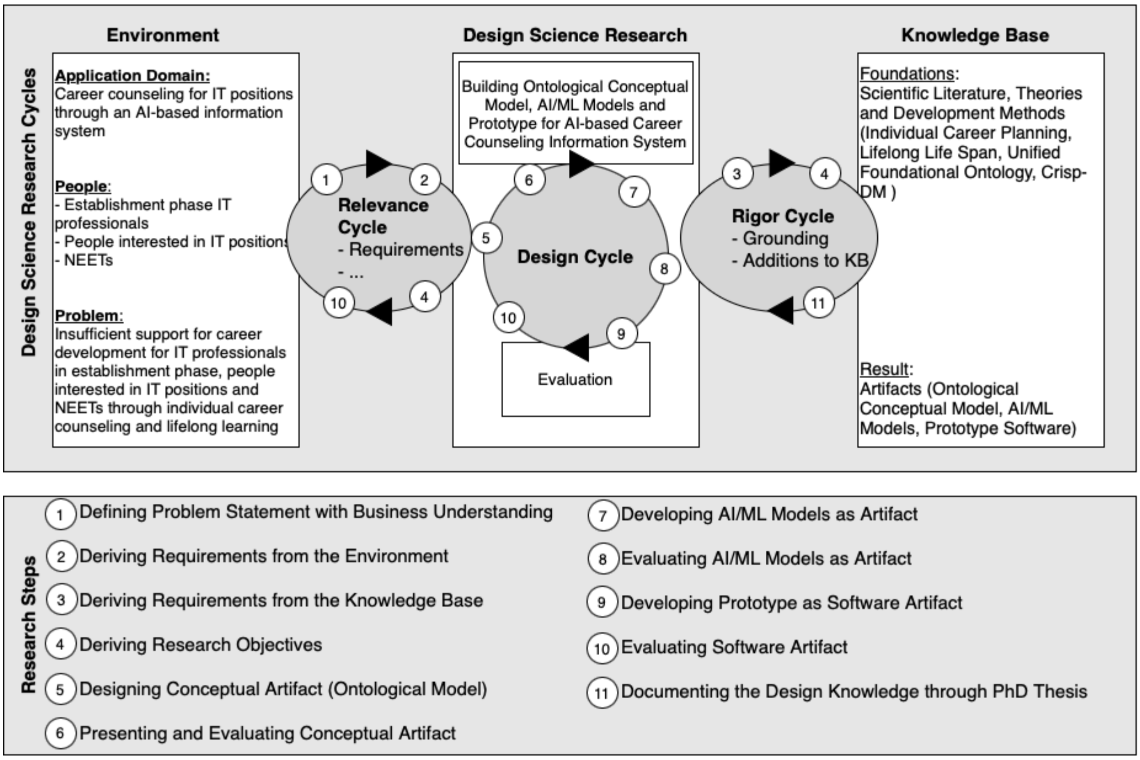
3.1.1. Research Design

To address the research goal by developing an innovative AI-based career planning system for IT professionals, individuals interested in IT positions, and NEETs, a DSR method is employed. Using this approach, the aim is to (1) provide a relevant solution for the predominant problem setting as described in the introductory section by applying a scientific approach and (2) derive generalized design implications for the Information Systems (IS) research discipline.

The three-cycle information systems research framework by Hevner (2007) and Hevner et al. (2004) was used as a basis, and the research steps were applied according to the relevance, rigor, and design cycles displayed in Figure 8.

Figure 8

Research Design



In the first stage of the research, the research problems were clearly defined. Using a rigorous approach based on theories of individual career planning (Jaffe & Scott, 1991), career development (Erdoğmuş, 2003; Granrose & Portwood, 1987; Hall, 1986; Super, 1963), and life-span life-space (Super, 1980), the research problems were identified as stated in the introduction section (Step 1). The expert interviews conducted allowed human resources professionals and scholars to derive relevance from the environment, which was utilized as input for the design cycle (Step 2). These interviews and their outcomes are described and reported in the sections on understanding data and sample selection.

Scientific literature and relevant theories - individual career planning theories (Jaffe & Scott, 1991), early career behavior theories (Hall, 1986; Super, 1963), life-span, life-space theory, and lifelong education approach (Super, 1980) – along with the researchers' contributions (Aktaş et al., 2022; Aktaş & Akbıyık, 2019, 2023), enabled the study to proceed rigorously by drawing extensively from the knowledge base (Step 3). Then, by leveraging both the knowledge base (literature) and the environment, the research objectives were determined (Step 4). These aspects are explained in the introduction, understanding data, and sample selection sections of the thesis.

Within the design cycle, the development of the ontological conceptual model (Step 5 and 6) was conducted following the Systematic Approach for Building Ontologies (SABiO) ontology development methodology (Fablo, 2014). This approach utilized Unified Foundational Ontologies (Guizzardi, 2005; Guizzardi et al., 2015), and employed OntoUML modeling language (Fonseca et al., 2021). The relevance and the potential contributions of UFO and OntoUML to the field of Management Information Systems were explored in a previous study by the researcher (Aktaş & Akbıyık, 2023). The processes of AI/ML model development (Steps 7 and 8), which are the components of the design cycle, were carried out within the framework of the CRISP-DM approach (Chapman et al., 2000). The subsequent prototyping (Step 9) took place using the AI/ML models developed in the previous step. The development processes are detailed under the relevant subsections in the methodology section of the study. Evaluation of the artifacts is shared in the findings section, utilizing machine learning success criteria and the software usability scale (Steps 6, 8 and 10).

The contribution of the design science research study to the knowledge base has been documented through academic publications related to the subject conducted by the researcher during the thesis (Aktaş et al., 2022; Aktaş & Akbıyık, 2019, 2023), as well as through the PhD thesis itself (Step 11).

3.2.Relevance Cycle

3.2.1. Defining Problem Statement (Relevance Cycle, Step 1)

The rapid change and advancement in technology continually alter the demands and requirements of the labor market, making individual career planning challenging. This ongoing change causes mismatches among employees, and it becomes increasingly difficult to integrate NEET individuals into the labor market (Hirschi, 2018; Kuzgun, 2021). In this dynamic and volatile job market, people should make individual career planning not only in line with the needs and expectations of their employers but also in line with their own goals, talents, and competencies (Arthur, 1994; Forret & Sullivan, 2002).

The literature indicates that individuals often require guidance during various stages of individual career planning, such as transitioning to different roles, advancing within their current positions, switching to a different field, or seeking employment in a new city (Dini et al., 2022). In this context, lifelong education is emphasized as crucial (Super, 1980). However, research on career development and career planning often concentrates on the career development processes of students or relies on data sources from student populations and this limits the research findings' applicability to broader audiences at different stages of their career and lives (Çarkıt, 2019; İstanbullu Dinçer et al., 2013; Jackson, 2017; Özdemir & Kibar, 2018; Polat et al., 2016; Sevinç & Siyez, 2018; Tomy & Pardede, 2019).

Despite the dominant view in the literature that individuals are primarily responsible for their career development (Adekola, 2011; Leibowitz et al., 1991), this thesis suggests that artificial intelligence-based systems can significantly enhance decision-making in individual career progression and lifelong learning. It argues for utilizing professional data from social media platforms dedicated to professional networking to train AI and machine learning models that accurately reflect job market conditions. Furthermore, this study proposes using ontology-driven conceptual modeling to map and validate the relevance of this data to the research aims and the field. Ultimately, the aim is to develop an ontology-driven, artificial intelligence-based system to facilitate this process.

To achieve this aim, the thesis addresses the following research problem: "How does an artificial intelligence-based career planning system, utilizing ontology, data science, and machine learning techniques, affect the career development and job alignment of information technology professionals, career changers, and NEET individuals?"

To address this research problem, the study also explores several sub-problems. These sub-problems are "Is ontology-based conceptual modeling an effective method for ensuring the scalability, adaptability, and interoperability of an artificial intelligence-based career planning system?", "Are data science methods effective in accurately determining job alignment for specific information technology positions based on an individual's current skills, experiences, education, and qualifications?", "How effectively can an artificial intelligence-based system provide personalized, accurate recommendations on the skills and competencies individuals need to acquire or improve to align with targeted information technology positions?" and "How effectively does the prototype of an ontology-driven, artificial intelligence-based career planning system demonstrate its feasibility and functionality?"

3.2.2. Deriving Requirements from the Environment (Relevance Cycle, Step 2)

The human resource and career management literature often points out that job postings do not always align with the actual requirements of a job position (Barber & Roehling, 1993; Felstead et al., 2015). Despite this, many studies in this field use job postings as a data source to describe and identify the requirements for a job position (Karakatsanis et al., 2017; Khaouja et al., 2021; Zhong et al., 2023). This study argues that using career information from individuals on PSMPs can accurately identify job and position descriptions and requirements.

3.2.2.1. Data Source Selection

PSMPs are social networking platforms that focus on professional networking and career development (Caers & Castelyns, 2011). Users of PSMPs often utilize these networks to showcase their professional credentials, primarily for self-promotion. They aim to tap the platform's resources for potential employment opportunities and the acquisition of professional information (Van Dijck, 2013). In addition, active or passive participation in PSMPs and disclosure of personal profiles have significant positive effects on the perceived social connectedness of PSMP users (Aktaş & Akbıyık, 2019; Mashayekhi & Head, 2018). Studies also show that users' profiles in PSMPs are akin to online resumes and can help users extend their professional network by establishing a common ground for professional self-promotion (Aktaş & Akbıyık, 2019; Mashayekhi & Head, 2018, 2022).

In this phase of the thesis, preliminary data collection was conducted from PSMPs. This initial step involved reviewing a dataset consisting of 25 career profiles of IT professionals to assess its suitability for the objectives of the thesis. This evaluation was conducted using expert opinions gathered from human resources business experts and academic advisors. During these expert opinion meetings, the preliminary dataset was examined, and its relevance to the problem being addressed in the study was discussed. These meetings were conducted as one-on-one sessions with three human resources professionals who are members of the Business Professionals Association (PIKDER), an organization that operates in cooperation with Sakarya University School of Business, and three academic advisors who specialize in human resources. The three academic advisors all hold PhDs in human resources management. The three business professionals are human resources managers of three large companies in the Marmara region of Turkey, contacted via PIKDER. The meetings were held either remotely or in person, depending on the convenience of the participant. During the interviews conducted with the open-ended question “How can data from professional social media profiles be leveraged to improve career planning?” to gather expert opinions, human resources professionals and academic advisors agreed that career information shared on personal profiles on professional-oriented social media platforms—such as work experience, education, certifications, skills, and competencies—provides a comprehensive view of an individual's capabilities. There was a consensus that a well-structured and regularly updated user profile can indeed serve as a digital resume. Based on this agreement, the participants concluded that these profiles are a suitable data source for addressing the research problem in the study.

Additionally, human resources business experts have indicated that professional social media profiles are already being reviewed during recruitment processes and candidate screenings, with the consent of the individuals. This practice is also supported by findings in the literature, confirming that recruiters often utilize these profiles to assess candidates' qualifications and fit for job positions. Research indicates that the evaluation of social media profiles is increasingly being integrated into the hiring procedures of organizations (Greenwood, 2009; Kluemper & Rosen, 2009). A study shows that social media profiles, especially PSMS profiles, are used by employers for screening to distinguish the candidates' professional attributes and their job fit (Hoek et al., 2016). Consequently, the study recommends that applicants should proactively manage their profiles on Professional Social Media Sites, such as LinkedIn, to clearly convey their skills and fit for potential job opportunities to employers (Hoek et al., 2016).

3.2.2.2.Sample Selection

Following the data source selection, a second round of expert consultation was conducted concerning the sample. This phase aimed to determine the sampling criteria for the dataset. This targeted selection is crucial for having a well-suited dataset for training the artificial intelligence-based individual career planning system, ensuring it effectively meets its design objectives. Sample evaluation, in line with the thesis objectives, was discussed with three human resources business professionals who are members of PİKDER and three academic advisors, all holding PhDs in human resources management. The three business professionals are human resources managers of three large companies in the Marmara region of Turkey, who were contacted via PIKDER. This consultation was guided by the open-ended question, "What criteria do you believe are essential for selecting companies that exemplify strong human resources processes and corporate culture for inclusion in research studies?".

The compilation of the general recommendations made in the interviews with the Human Resources experts is as follows:

> "The presence of well-defined human resources processes in the companies planned to be included in the research indicates HR institutionalization. This suggests that these companies likely have accurate and precise job descriptions, making them suitable candidates for the study.
>
> "…It's crucial for a company to have a strategic plan and to make long-term investments in HR processes."

Experts also shared recommendations on companies' workspaces and suggested sectors. Compilation of the recommendations on companies' workspace are as follows:

> "The field of work and sector of a company is crucial in terms of accurate job descriptions. "
>
> "IT companies, pharmaceutical firms, and banks often have advantages over other businesses regarding institutionalization, openness to investment, and well-defined HR processes. These characteristics can be highly valuable for your work, providing a solid foundation for effectively implementing and evaluating HR strategies."

Lastly, experts also discussed the company selection criteria for the dataset. Recommendations on company selection criteria and possible candidate companies are as follows:

> "Companies that have received awards and certifications focused on employee satisfaction and corporate culture demonstrate their commitment to long-term investments in HR processes. Such recognition indicates that these companies

> prioritize the well-being and development of their workforce. These organizations can generally be considered corporate entities with accurate job descriptions."
>
> "Businesses undergoing the certification process can be regarded as having reached a certain level of corporate maturity. These are companies that have been through thorough audit and evaluation processes. More importantly, the willingness to participate in such evaluations significantly indicates a business's commitment to maintaining and improving its corporate standards and practices."
>
> "Great Place to Work (GPTW), Peryön (Human Management Association of Turkey) Human Value Awards, Kariyer.net Respect for People Awards, European Foundation For Quality Management (EFQM) Excellence Awards are the leading awards and certifications that indicate employee satisfaction and corporate culture in Turkey."

As a result of the interviews, businesses that participated in employee satisfaction certification programs and received awards were determined as the priority group. Based on these findings and considering the limitations of the thesis study, IT employees from companies that have received the international 'Great Place to Work' employee satisfaction award in the last three years (2020, 2021, 2022) were included in the study. This selection criteria ensures the inclusion of companies with recognized standards of employee satisfaction and corporate culture, providing a robust foundation for the research.

3.3.Rigor Cycle

3.3.1. Deriving Requirements from the Knowledge Base (Rigor Cycle, Step 3)

In the decision-making phase concerning the selection of theories, approaches, and methodologies to be utilized in the design stage of the system, the knowledge base was thoroughly reviewed to establish the necessary literature foundation for the system's development. Scientific literature and relevant theories - individual career planning theories (Jaffe & Scott, 1991), early career behavior theories (Hall, 1986; Super, 1963), life-span, life-space theory, and lifelong education approach (Super, 1980) – along with the researchers' contributions (Aktaş et al., 2022; Aktaş & Akbıyık, 2019, 2023), enabled the study to proceed rigorously by drawing extensively from the knowledge base. These theories and approaches are examined in detail in **Chapter 1**.

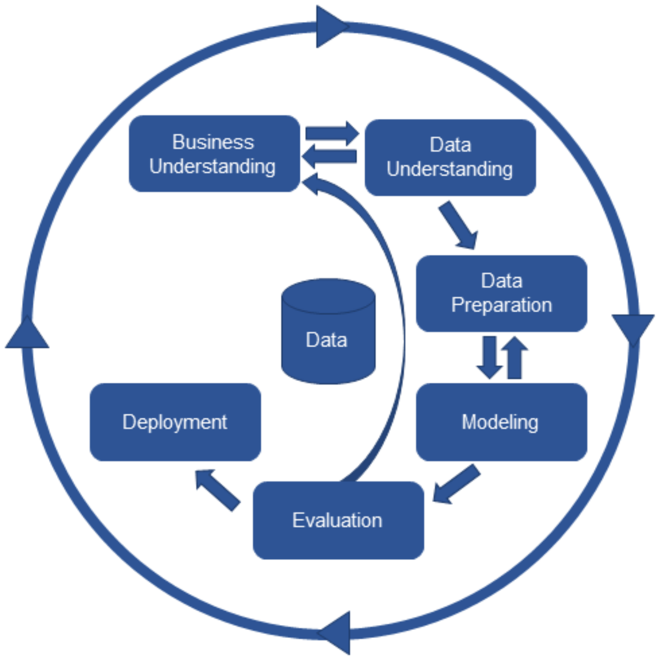
Within the scope of the research, the Unified Foundational Ontology was used for the theoretical foundation of the ontology-driven conceptual model, and the OntoUML modeling language was utilized for its implementation. UFO and ontological stereotypes are examined in detail in **Chapter 2**. The development process of the ODCM model was also conducted using the Systematic Approach for Building Ontologies (SABiO) methodology, which is employed to support ontology development to enhance software development processes. The SABiO methodology comprises five stages: Purpose Identification and Requirement Elicitation, Ontology Capture and Formalization, Design and Implementation, and Testing. (Fablo, 2014). This approach is explained in detail under the Design Cycle Step 5.

The selection of methodologies for system development was informed by the knowledge base, and data science processes will be implemented using the CRISP-DM methodology. The CRISP-DM methodology was selected for this study because it provides a comprehensive, structured approach to data science that ensures all necessary aspects of the project are considered. Its iterative nature allows for continuous refinement and improvement of the project at each phase, ensuring a thorough understanding of both the business problem and the data (Saltz, 2021). The model's clear and systematic process helps maintain focus and direction throughout the project, ultimately leading to more robust and effective solutions (Martinez-Plumed et al., 2019). Additionally, CRISP-DM's widespread acceptance as a de facto standard and proven track record in various industries (Martinez-Plumed et al., 2019; Schröer et al., 2021) add to its reliability and effectiveness for developing an AI-based career planning system. According to this model, data science projects are structured into six main phases: (1) Business Understanding, (2) Data Understanding, (3) Data Preparation, (4) Modeling, (5) Evaluation, and (6) Deployment (Chapman et al., 2000). Directional arrows between the phases illustrate the relationships and connections among them. The large loop surrounding the model highlights the iterative nature of the data mining process. In addition, data is at the center of the figure, emphasizing its pivotal role in data science projects.

The business understanding phase is the initial step in a data science project. This phase involves understanding the objectives and requirements of the data science project from a business perspective. In this step, the problem to be solved is identified and defined. Preliminary plans for addressing the data science challenges are made, and success criteria are established to evaluate the solution's effectiveness. This phase is crucial for aligning the project's objectives with the potential data-driven solutions and outcomes

Figure 9

CRISP-DM Phases



Source: Chapman et al. (2000)

In the data understanding phase, data suitable for addressing the problem is identified, and preliminary data collection is conducted. The focus here is on understanding the data, identifying any quality issues, and making preliminary discoveries. In data science projects, there is a two-way relationship between this phase and the business understanding phase due to the necessity of selecting a data set that is suitable for the problem to be solved and thoroughly understanding the data. This relationship allows for the revision of the business analysis based on challenges or opportunities experienced during data collection. It also ensures that the data gathered is appropriate for the business problem.

Preparing data phase covers the transformation of raw data into a dataset ready for modeling. It encompasses several steps, including data cleaning, data transformation, data standardization, noise removal, and feature engineering. The specific stages of data preparation can vary depending on the dataset and the machine learning algorithms used.

The modeling phase includes choosing and implementing various analytical methods or machine learning algorithms suited to the research question and data. The chosen method is then refined and adjusted to address the research problem effectively. It's important to note that certain models and analytical techniques may demand particular types of data structures. Therefore, there is an interdependent relationship between the modeling and data preparation phases to ensure data is properly formatted to support selected technique.

At the evaluation phase, the effectiveness of the analysis techniques or machine learning algorithms used in the research is assessed against the success criteria established during the problem identification phase. This assessment ensures whether the model or algorithm has been correctly developed to solve the research problem. If the results are satisfactory, the project may progress into the deployment stage. Alternatively, the process might even revert to the business understanding phase to refine the approach and repeat the research cycle for better results.

In the Deployment phase, various analysis techniques or machine learning algorithms are utilized to create a solution for the end user. Although specifics may vary across projects, the deployment phase generally includes (1) developing models that successfully meet the evaluation criteria, (2) designing software to generate reports for the end user, or (3) developing a program or a website designed to meet the specific needs of the end user.

3.3.2. Deriving Research Objectives (Relevance and Rigor Cycles, Step 4)

This thesis aims to design and prototype an ontology-driven AI system for individual career planning. It seeks to facilitate career development for those lacking access to traditional career planning, thereby fostering a sustainable, skilled workforce. The study concentrates on constructing ontological foundations essential for matching individuals with suitable IT positions through job fit scores and skill recommendations. This system specifically serves IT, job seekers, individuals considering career transitions and NEETs.

The study aims to integrate ontology-driven conceptual modeling, data science, machine learning, and artificial intelligence techniques to develop a robust ontological model of PSMP data and the research domain. It also aims to create an ML/AI model that calculates the job fit of individuals for selected IT positions based on their capabilities and provides recommendations for enhancing job fit through skill and competency development within lifelong education frameworks.

3.4.Design Cycle

This thesis aims to develop an ontology-driven, artificial intelligence-based system that will provide individual career planning using data science and machine learning technologies. The design cycle of this thesis encompasses the creation of three artifacts: an ontology-driven conceptual model, AI/ML models, and the prototype. The following sections will explain the methodology used to develop and evaluate the aforementioned artifacts.

3.4.1. Designing Ontology-Driven Conceptual Model (Design Cycle, Step 5)

ODCM employs a comprehensive language rooted in ontology and semantics, facilitating communication among stakeholders such as business analysts, system developers, and end-users. This rich semantic foundation and modularity make ODCM and OntoUML suitable for modeling complex domains, enabling precise definition and mapping of entities and relationships. Research by Verdonck et al., (2019) indicates that ODCM techniques excel in modeling large and complex domains, though no significant difference is observed for simpler domains compared to traditional methods. The ontologically well-founded nature of OntoUML enhances reusability, extendability, and interoperability, fostering a collaborative environment essential for system extension and integration (Gemino & Wand, 2005). Ultimately, ODCM complements traditional system design methods, providing a more explainable and interpretable structure that aligns with the depth and complexity of real-world business problems. Consequently, this study adopts ODCM to leverage these benefits in addressing the complexities inherent in MIS.

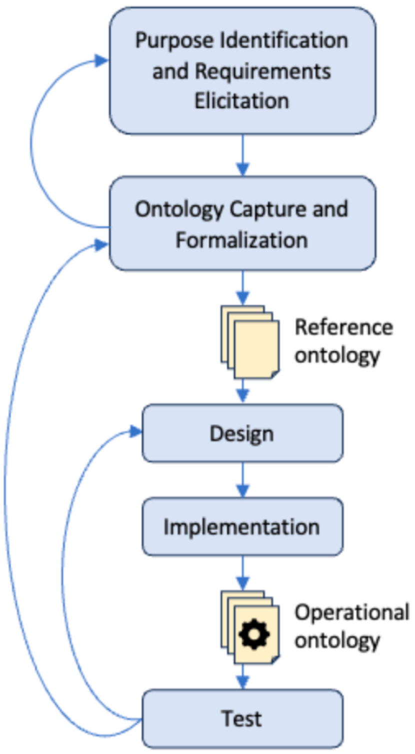
An ontological model created using the SABiO methodology following the competency questions will serve as a common conceptual model and a domain model for application development. SABiO is an ontology development methodology that has been inspired by software engineering practices (Fablo, 2014). SABiO is structured in five phases (Figure 10). The main phases of SABiO are as follows.

Purpose identification and requirement elicitation: In this phase of ontology development, the purpose of the ontology is clearly defined by engaging with relevant stakeholders. The ontology is designed to facilitate communication among these stakeholders by establishing a common vocabulary with precise semantics. Additionally, it serves to provide semantic annotations, which help disambiguate information from diverse data sources and assist in explaining the results of machine learning analyses.

Moreover, both functional and non-functional requirements of the ontology are specified during this stage. To capture the functional requirements, specific competency questions were formulated. These questions are designed to clearly delineate the scope of the ontology based on the knowledge inquiries it is intended to address, thereby defining the range of problems the ontology aims to solve. Non-functional requirements encompass various quality and project-related criteria, such as the need to align with existing models or ontologies and the use of specific languages and tools that may influence the development of the ontology. If the ontology becomes too large and complex, covering a wide range of aspects, it may be necessary to decompose it into sub-ontologies. This approach helps manage complexity and maintain clarity in the ontology's structure and functionality. The outcome of this phase should be a comprehensive document that describes the ontology's purpose, requirements, and structure.

Figure 10

Overview of the SABiO Ontology Development Methodology



Source: Fablo (2014)

Ontology Capture and Formalization: In this phase, a reference ontology is created by defining a conceptual model and adding axioms to achieve the necessary level of formalization. Utilizing a foundational ontology, such as UFO (the Unified Foundational Ontology (Guizzardi et al., 2022), can be particularly beneficial during this phase. The UFO provides foundational concepts that aid in the development of consistent ontologies. For the creation of the reference ontology, OntoUML (*OntoUML*, n.d.), a UML profile specifically designed for defining ontologies based on UFO principles, is employed. OntoUML is implemented as a plugin In the Visual Paradigm UML tool (Fonseca et al., 2021).

***Design and Implementation*:** During these phases, the reference ontology is transformed into an operational ontology by taking into account the non-functional requirements and the specific environments in which the ontology is expected to operate.

***Test*:** In this phase, the operational ontology is tested to ensure it meets both its functional and non-functional requirements. The competency questions defined earlier are used as a benchmark to assess the onology's effectiveness. This testing is typically conducted within an ontology development tool. This procedure allows for the verification of the ontology by checking if it can provide appropriate answers to the questions and whether the underlying rules of the ontology remain consistent across different scenarios.

Competence, skills, and certifications are key elements in human resource management, educational planning, and personal career development. However, understanding the intricate relationships between these elements can be challenging. The project aims to use the OntoUML to develop an ontological model that represents and relates individual competencies, skills, and certifications. This study aims to map the relationships between a person's job experience, education, skills, competencies, and certifications. This will be done along the following competency questions, following the SABiO ontology development methodology: (1) What are the structural components of individual competencies, and how do they relate to a job position and industry-recognized certifications? (2) How can we ontologically model the relationship between a person and their educational information? (3) How can we ontologically model the relationship between a person and their job experience?

The competency questions and the SABiO methodology will guide the ontology-driven conceptual model development. The result of the ontological modeling process is shared in **Chapter 4**.

3.4.2. Evaluation of the ODCM (Design Cycle, Step 6)

The evaluation of the ontology-driven conceptual model (ODCM) was a multi-phase process designed to ensure the model's robustness, relevance, and alignment with the research objectives. The ODCM was evaluated using several key steps. The first step involved verifying the ontological integrity of the ODCM using the foundational principles of the Unified Foundational Ontology (UFO). The UFO provides a set of conceptual categories that ensure the consistency and coherence of the ontology. This step was critical in establishing a robust ontological foundation for the ODCM. The next phase involved validating the OntoUML models used in the ODCM. OntoUML, an ontologically well-founded modeling language, extends UML with the distinctions from UFO. This validation process ensured that the OntoUML models accurately represented the domain and maintained semantic integrity. To ensure the accuracy of the ontological model created in this research, the consistency and syntax of the model will be verified using the Visual Paradigm application. Also, a literature-based validation is conducted, involving a comparison of the model with the results of existing academic studies. This validation process is designed to confirm the reliability and scholarly alignment of the developed model. This process was shared in detail in **Chapter 4**.

3.4.3. Development of the Artificial Intelligence Models (Design Cycle, Step 7)

This section begins by outlining the data collection and preparation processes essential for developing machine learning and artificial intelligence models for career planning. Utilizing web scraping techniques, the study systematically collected career-related data from professionally oriented social media platforms. The development of a specialized web scraping tool facilitated the efficient extraction of this data. This dataset, crucial for training the ML/AI models, was processed and standardized to ensure its suitability for analysis. Following data preparation, the section details the specific algorithms used in model development, including a discussion of various machine learning techniques and the application of the ensemble learning method to enhance the predictive accuracy and robustness of the models.

3.4.3.1.Data Collection

The dataset was collected using a web scraping method using Python programming language and employing several software libraries such as BeautifulSoup, Pandas, and

Selenium. BeautifulSoup was utilized to parse HTML and XML documents, allowing for efficient data extraction from web pages. Pandas was used to store and process large datasets, enhancing data handling capabilities. Selenium was used to automate web page access, enabling systematic data collection.

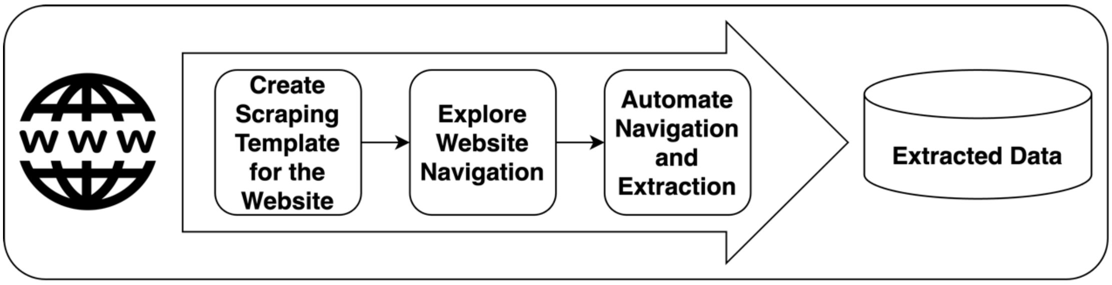
Python is an object-oriented, interpreted, high-level programming language created in 1991 by Guido van Possum. It is widely used by technology companies, including Wikipedia, Google, Facebook, Amazon, Spotify, and Reddit. Useful software libraries such as NumPy, SciPy, Matplotlib, and Pandas have led to the rapid popularization of Python in data analysis and data science. Additionally, software libraries such as BeautifulSoup4, Selenium, and Scrappy make Python a preferred language for web-scrapping applications. Python also has a wide range of open-source libraries that support various functionalities, including automation, graphical interfaces, machine learning, and data processing (Welcome to Python.Org, n.d.).

Web scraping is a method for extracting data from online sources. This method involves obtaining information either by directly visiting websites or other online sources or by accessing the HTML code of the source. Although it is possible to scrape data manually, automation software is often used to gather more data in less time. Typically, the term web scraping refers to these automated processes, which streamline the collection of large amounts of web-based data (Krunal A, 2014; Lawson, 1989; Milev, 2017; Polidoro et al., 2015; Ren & Ren, 2018; S.C.M. de S Sirisuriya, 2015; Slamet et al., 2018). In this study, the web scraping process suggested by (Krunal A, 2014) was followed (Figure 11). According to the web scraping approach suggested by Krunal, the data acquisition process consists of (1) accessing the HTML code of the website, (2) identifying the HTML structure, (3) using this structure to automate the navigation and acquisition of data, and (3) finally, saving the acquired data.

In order to realize the system that is aimed to be developed within the scope of the thesis studies, first of all, necessary data should be collected for the training of artificial intelligence. Data collection tools need to be developed to start the data collection process. The data required within the scope of the thesis is the career information shared in the PSMP. Since there is no ready tool or openly available Application Programming Interface (API) to obtain this type of data, it was decided to develop data collection tools using the Python software language.

Figure 11

Web Scraping Process



Source: Krunal (2014)

A two-step process is planned to reach the data set intended to be obtained within the scope of the study. The first step aims to reach the companies planned to be included in the research and to access the list of PSMP profiles of the individuals associated with and working in the company. In the second step, publicly shared professional information of the profiles determined in the first step will be accessed and recorded.

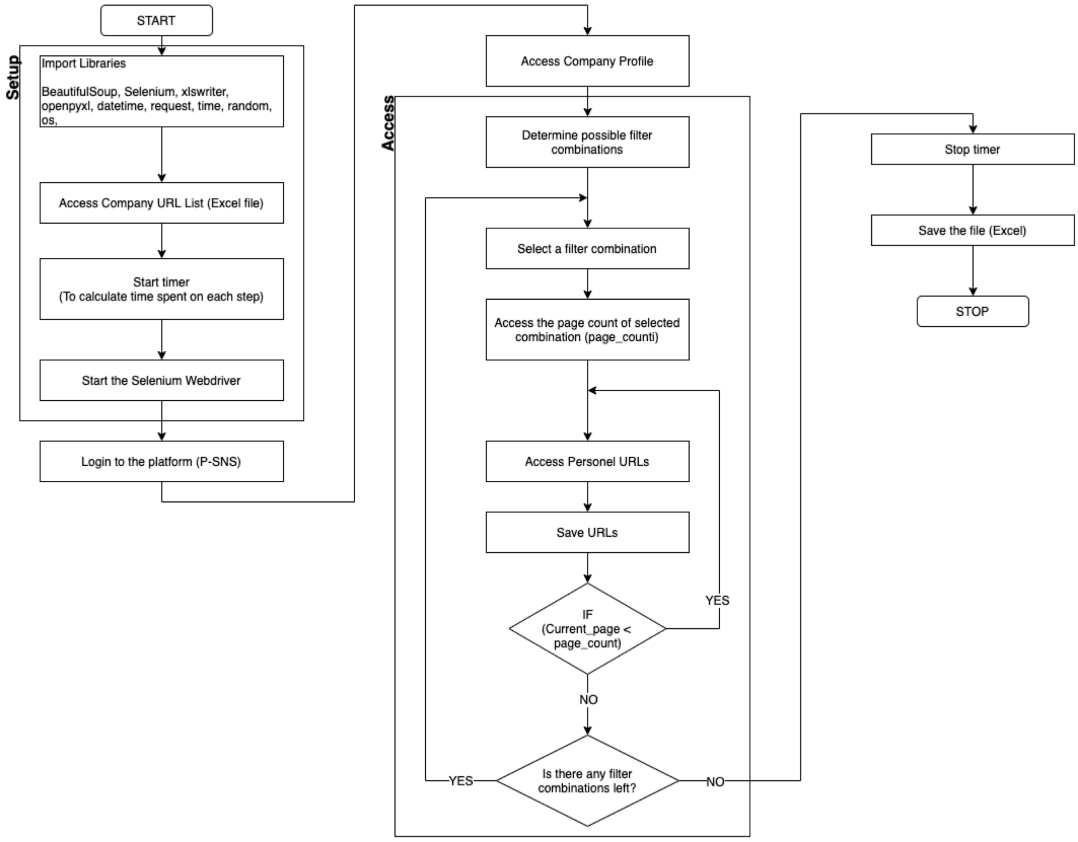
The two-step web scraping structure was designed based on several algorithms. The purpose of the first step is to facilitate the initial stage of the data collection process of the thesis study, which is to collect and record the URL addresses of the PSMP profiles (Figure 12). These profile addresses belong to active, publicly available PSMP users who are associated with the predefined pool of companies.

Python software language was used in the development of this software. In the software development process, Selenium library was used for automation, BeautifulSoup4 library was used for HTML parsing and scraping processes to provide access to only the desired data from the data, openpyxl library was used to access the Excel document containing the company list to be accessed, and xlsxwriter library was used to save the obtained data permanently. All source code is original and written by the PhD candidate.

The flowchart in Figure 12 illustrates the automated process of extracting personnel URLs from company profiles. It begins with setting up by importing necessary libraries, accessing the company URL list, starting a timer, and initializing the Selenium WebDriver to log into the platform. The process then involves accessing company profiles, determining and applying filter combinations to retrieve personnel URLs, and saving these URLs. This loop continues for all filter combinations until all possible URLs are extracted. Finally, the timer is stopped, and the results are saved in an Excel file. This method ensures efficient and systematic data extraction for further analysis.

Figure 12

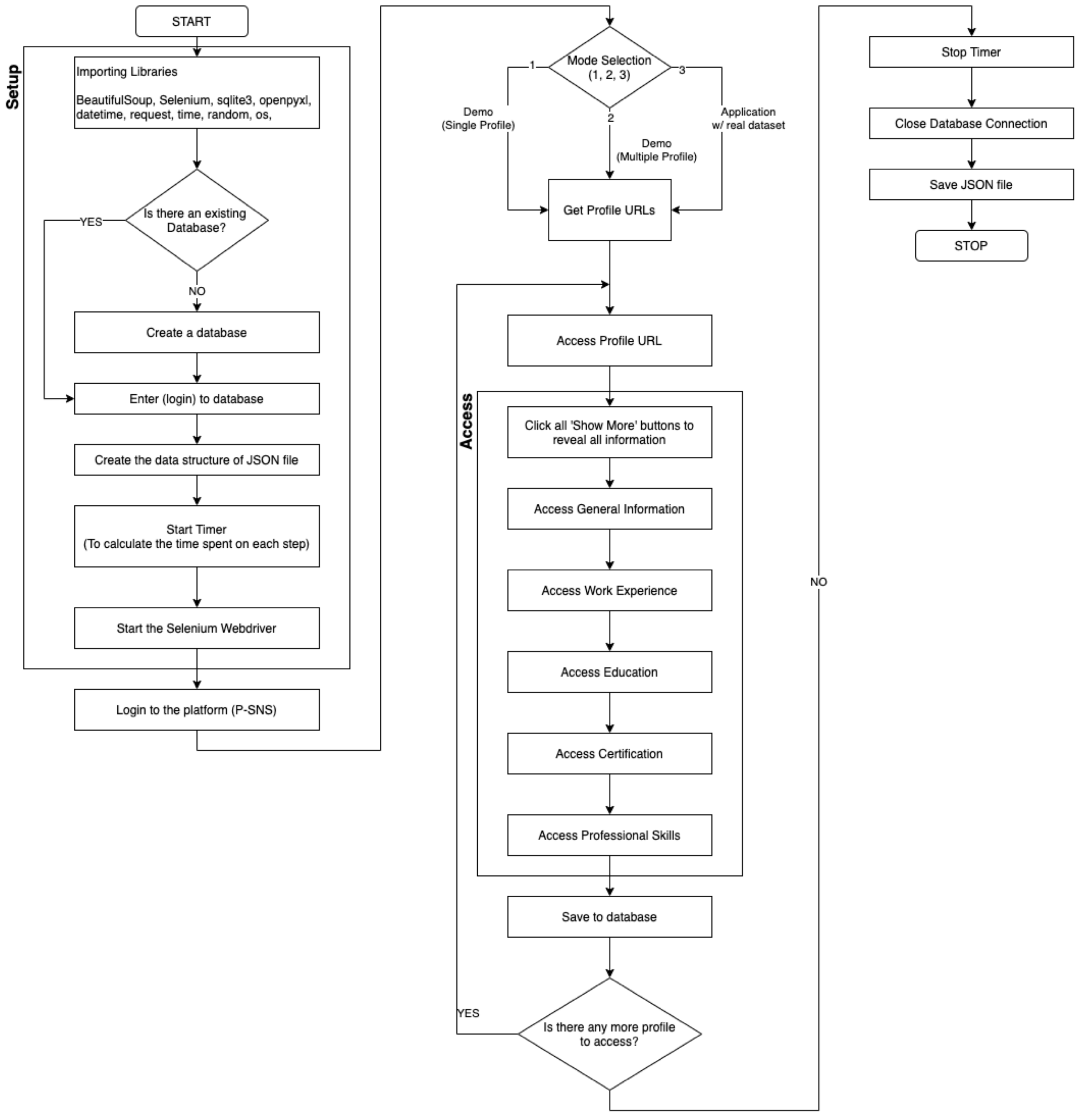
Automated Personnel URL Extraction Process Flowchart



The purpose of flowchart shared in Figure 13, which was designed to facilitate the second stage of the study's data collection process, is to collect and record the professional information that users share publicly on their social media accounts (Figure 13). The flowchart details the automated extraction process of profile information. Once necessary libraries are imported, a database is created or accessed. Then, the mode selection (Demo for debugging or production-ready with real data) is performed to retrieve profile URLs. The system reveals all information for each profile URL by clicking the 'Show More' buttons, accessing general information, work experience, education, certification, and professional skills, and saving the data to a database. Finally, the timer is stopped, the database is saved, the connection is closed, and the results are also saved as a JSON file, ensuring an efficient and thorough extraction process. The application uses the professional social media profile URL addresses collected and recorded with the first application to access information about their careers that users share publicly. This information includes details such as education, work experience, certificates, skills, and current occupation. Once accessed, the data is anonymously stored in an SQL database and JSON file.

Figure 13

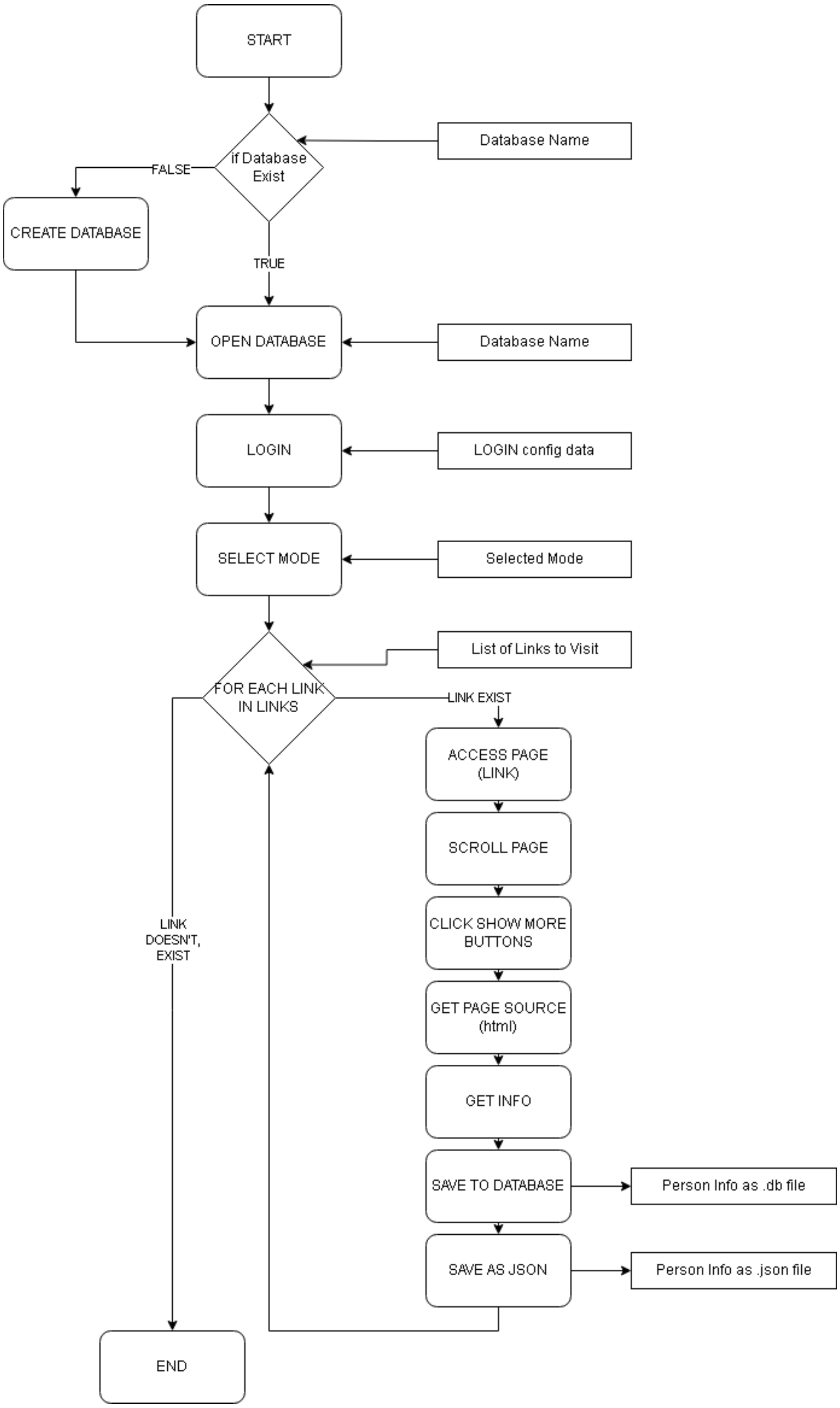
Automated Profile Information Extraction Process Flowchart



Before the data collection, a code reconstruction took place for the web scraping algorithm in Figure 13 because an update had rendered most of the codes of the previously tested data collection software unusable. During the code reconstruction, the basic libraries of the code written in Python remained constant, but the parts that made HTML Parsing had to be rewritten. Therefore, because of this change, the code structure of the data collection software was reconsidered, and most of the code was rewritten (Figure 14, Figure 15). Also, procedural improvements were made during the code rewriting phase for system stability.

Figure 14

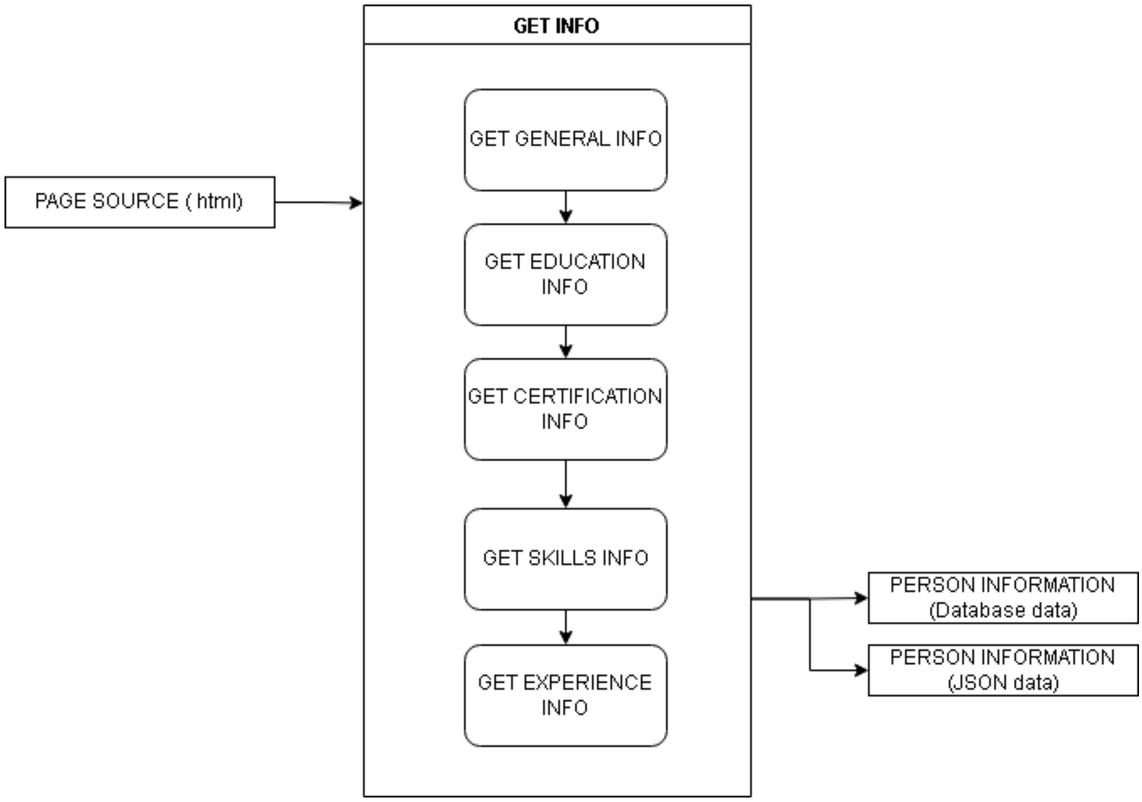
Automated Web Data Extraction and Storage Process Flowchart



The flowchart in Figure 14 outlines the automated process of extracting and saving profile information from web pages. The process starts by checking if a database exists; if not, a new database is created. The system logs in using configuration data and selects a mode, then accesses a list of links to visit. For each link, the system accesses the page, scrolls click the 'Show More' buttons, retrieves the page source, and extracts the information. The extracted data is then saved to the database and as a JSON file.

Figure 15

Detailed Profile Information Extraction Flowchart



The flowchart in Figure 15 demonstrates the process of extracting and storing profile information. Starting from the HTML page source, the system retrieves various types of information: general info, education, certification, skills, and experience. Each type of information is sequentially extracted and then saved into both database and JSON formats. This structured approach ensures comprehensive data collection and storage for each profile.

3.4.3.2.Data and the Data Structure

After the web scraping tool was developed, the data collection process took place between December 2022 and April 2023. As a result of the data collection process, a total of 12000 profiles were visited and scraped. However, 2743 of these profiles were information technology and information systems positions and were included in this study. These positions were grouped under 6 position categories: Data Analysis and Business

Intelligence, Software Development, Product and Project Management, Quality Assurance and Test, System Development and System Engineering, IS Consultancy, Strategy and Design.

A relational database stores the data obtained by the software developed with the algorithms described above. A relational database is a digital database whose organization is based on the relational data model. Various software systems used to maintain relational databases are known as relational database management systems. In Relational Database Management Systems (DBMS), data is kept in rows and columns in tables. After evaluating the data to be collected, the relational database structure in Figure 15 and Figure 16 was created, and the relationships between the tables were defined in the figures.

The diagrams in Figure 16 and Figure 17 depict an Entity-Relationship (ER) model designed for storing profile information. The central "Person" entity is connected to various other entities: Experience, Certification, Education, and Skills. Each entity contains specific attributes such as job title, certification name, school name, and skill type. Linking tables (Person_Exp, Person_Cert, Person_Edu, and Person_Skills) associate each person with their respective experiences, certifications, education, and skills through foreign keys.

Figure 16

PSMP Profile Information Relational Database Structure

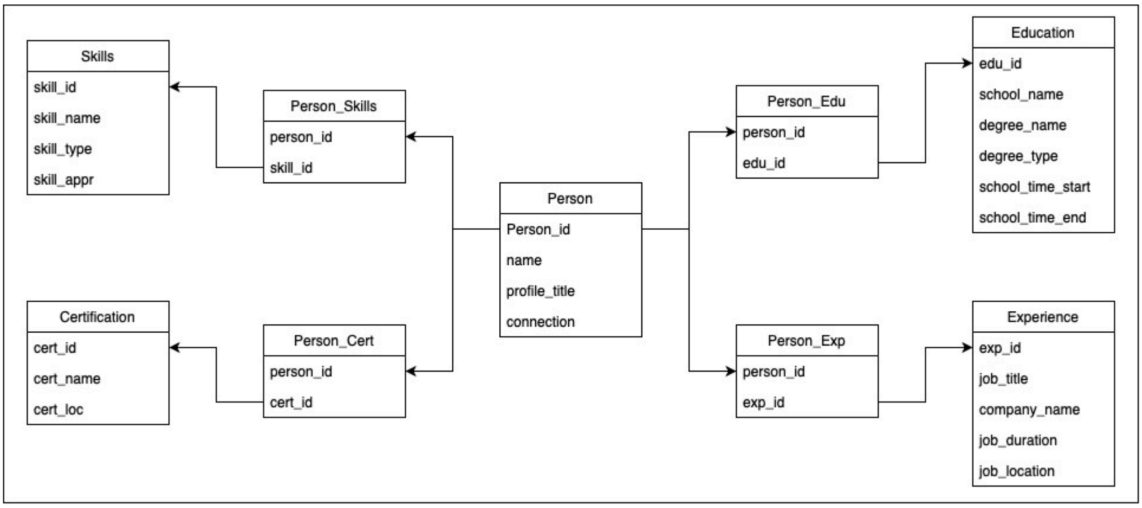
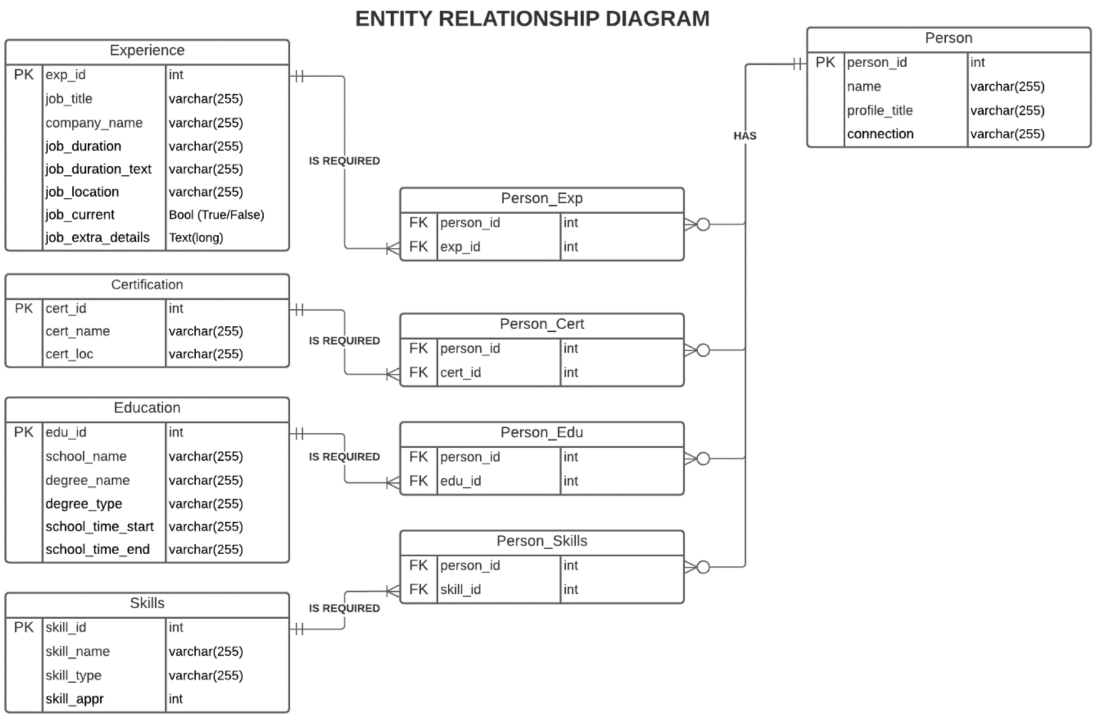


Figure 17

PSMP Profile Information Entity Relationship Diagram



The names, types, definitions, and sample data of the variables shared in the Database Structure and Entity Relationship (ER) diagram are provided in the following tables (Tables 2,3,4,5,6).

Table 2

Person Database Table

Variable Name	Variable Type	Variable Definition	Sample
person_id	Integer	ID number given to person in Database structure and JSON file structure	1
profile_title	Varchar (255)	Current job position that the person has shared	Research Assistant
connection	Varchar (255)	Number of PSMP connections	500

Table 3

Certification Database Table

Variable Name	Variable Type	Variable Definition	Sample
cert_id	Integer	The id number given to the certification in the database structure	1
cert_name	Varchar (255)	Name of certification	Google Analytics
cert_loc	Varchar (255)	Where did the certification come from?	Google

Table 4

Education Database Table

Variable Name	Variable Type	Variable Definition	Sample
edu_id	Integer	The number given to the education in the database structure	1
school_name	Varchar (255)	School Name	Ege University
degree_name	Varchar (255)	Degree Name	Electric Electronics Engineering
degree_type	Varchar (255)	Degre Type (Bachelor, Masters, PhD etc.)	Bachelor of Engineering
school_time_start	Varchar (255)	School start date	2009
school_time_end	Varchar (255)	School end date	2014

Table 5

Skills Database Table

Variable Name	Variable Type	Variable Definition	Sample
skill_id	Integer	The number given to the skill in the Database structure	1
skill_name	Varchar (255)	The full name of the skill	Python
skill_type	Varchar (255)	Skill type	Technical
skill_appr	Integer	Skill approval count	2

Table 6

Experience Database Table

Variable Name	Variable Type	Variable Definition	Sample
exp_id	Integer	Number given to experience in Database structure	1
job_title	Varchar (255)	Name of job or position	Full Stack Developer
company_name	Varchar (255)	company name	A Company
job_duration	Varchar (255)	How long the person has worked at the workplace	2 years 11 months
job_duration_text	Varchar (255)	The period of time the person works at the workplace	May 2013 – Mar 2016
job_location	Varchar (255)	location of the workplace	İzmir, Türkiye
job_current	Boolean (True/False)	Whether the person is still working at work	False

Following the database diagram, ER diagram, and data tables, it is important to note that PSMPs host valuable information such as company details, job openings, and user profiles with employment history, education, certifications, skills, and abilities, typically in an online resume style. Despite their value, these profiles lack semantic depth and ontological structure, limiting interoperability and machine actionability. Therefore, this study models publicly shared profiles using the ODCM approach and OntoUML to ensure semantic rigor and interoperability. The results of the ontological modeling are shared in Chapter 4.

3.4.3.3.Data Preparation

During the preparation phase of the thesis study, the collected textual data, intended for use in developing machine learning algorithms, was cleaned and standardized. This process involved removing irrelevant or redundant information, correcting errors, and converting the data into a consistent format to ensure it is suitable for analysis and modeling.

Given that the dataset intended for this research primarily consists of data entered by individuals, it is crucial to undertake the data cleaning steps. Some of the errors and inconsistencies needing correction include (1) the presence of values with spelling errors, (2) the incorrect use of English characters in place of Turkish characters, or vice versa, (3) some terms being entered in various forms (‘management information systems’, ‘mis, ‘ybs’ etc.), (4) inconsistent usage of abbreviations that may not be universally recognized, (5) mixed us of Turkish and Engish. To address these issues, a combination of manual review and algorithmic methods was employed to correct the dataset effectively.

Reference tables will be used to clean and standardize the data. During the process of standardizing educational data in the dataset, two key resources were employer: the standardized list of academic departments (Yüksek Öğretim Kurumu, n.d.-b) and the list of universities (Yüksek Öğretim Kurumu, n.d.-a), both provided by Turkish Higher Education Council (YOK) were used.

To standardize the names of occupations and positions in the dataset, two sources were used: the Turkish Labor and Employement Institution (ISKUR) occupations list (Türkiye İş Kurumu, n.d.) and the Vocational Qualifications Authority (MYK) occupations list (Mesleki Yeterlilik Kurumu, n.d.). These reference lists ensure that all job titles are aligned with nationally recognized standards, facilitating consistent categorization and analysis of occupational data across the dataset. This approach helps maintain the integrity of the data, particularly in analyses that require accurate and uniform job classification.

To address errors and to clean and standardize the dataset, a specific approach involving the use of lookup tables was implemented. A lookup table was created to include the values of the most frequently occurring items within the dataset. For each entry in the lookup table, the corrected value determined using the reference documents were entered. This manual process of preparing the lookup table is designed to serve as a foundational reference for subsequent stages of the data cleaning and standardization process, ensuring consistency and accuracy across the dataset.

For the values that are not feasible to manually correct, an algorithmic approach to data cleaning was implemented. Google Translate API was used to identify English terms and to translate them into Turkish. Additionally, the FuzzyWuzzy library in Python programming language was employed to calculate the similarity between the values in the dataset and those in the reference tables. This process involved the following steps:

- ‘LanguageDetect’ command from the Google Translate API was used to determine the language of the values in the database.
- Non-Turkish values were translated using the ‘Translate’ command of the Google Translate API. During this translation process, special attention was given to

instances where translations might lose their intended meaning when converted into Turkish. In such cases, manual intervention was employed to ensure the accuracy and appropriateness of the translations, preserving the original context and significance of the data. This careful review and adjustment process helps maintain the integrity and usability of the dataset for further analysis and application.

- The FuzzyWuzzy library in Python software language was used to compare the dataset with the data in the reference lists, obtaining a similarity index for each comparison. This index ranges from 0 to 1, where 0 indicates no similarity between two variables and 1 indicates identical variables. In this study, education and experience data within the dataset were replaced with the most similar value from the reference lists, provided the similarity was greater than 0.85. To ensure accuracy, values that exceeded the 0.85 similarity threshold were also subjected to a pre-review process before being accepted. This was done to ensure that the standardized data maintains high reliability and relevancy.
- The values in the skills data table were assessed based on the keywords listed in the Table 7. Using these keywords as a reference, corresponding codes and categories for the variables were determined. Manual coding was performed to cover 70% of the data based on frequency. After completing the manual coding for the remaining data, the FuzzyWuzzy library was used to consider similarities above 0.85 with a pre-review process. This approach was implemented to enhance the consistency and accuracy of the dataset.

Following the steps outlined above, a lookup table was created to identify the correct values that should replace the existing ones in the education, work experience, and skills data tables requiring cleaning. This lookup table was then used to replace the relevant values in the database, ensuring that the dataset was standardized. This process effectively aligned the data with recognized standards and reference values, improving the dataset's overall quality and usability for further analysis.

3.4.3.4. Feature Engineering

Feature engineering is an important and labor-intensive process in machine learning. The performance of machine learning models is highly affected by the dataset and the features. Hence, a significant portion of time and effort in machine learning model development is spent on preparing the dataset and engineering the features (Bengio et al., 2013). Creating new features from a given set of features is a frequent practice. These newly engineered features are designed to improve or replace parts of the existing features. This process involves adding new features calculated based on the other features. These features can be ratios, differences, and other mathematical or logical transformations of the values in the dataset (Heaton, 2016).

Before using machine learning algorithms, it is essential to transform raw data into features that can be effectively processed and made sense of. These features represent the attributes of the observed phenomenon, and they correspond to the columns in the data matrices that are used to train machine learning algorithms (Verdonck et al., 2021). The objectives of feature engineering can be expressed in two steps: (1) to prepare the input data so that machine learning algorithms can process and use it, (2) to transform the variables into features to enhance machine learning models' performance and to represent the domain better (Verdonck et al., 2021). Studies show that algorithms such as gradient-boosted machines, random forests, support vector machines, and deep neural networks significantly improved with appropriate feature engineering techniques (Heaton, 2016).

In this step, the dataset, which was gathered via web scraping and stored in a relational database, will be prepared for use with machine learning algorithms. The dataset in its current format includes both qualitative and categorical data. Machine learning algorithms typically require data in a specific numerical format, so the dataset will be transformed. Therefore, this step involved identifying key features of the dataset and organizing the data based on the identified features.

Firstly, the ontological model developed as a part of the research was examined for feature engineering. The ontological model (Figure 21) provided the following findings and insights. (1) Multiple skills combine to form a broader concept of competence. (2) Skills and competence are associated with a person and help characterize the person. (3) Competences are inherent to individuals, but they can be abstracted and generalized to be defined as a higher term, which is competence type. (4) Competence type is a higher-level, individual-independent definition of competence that applies across the labor market. (5) The ontology model also established the connection between skills, competencies, and position through the association relation between competence type and position type. (6) The ontology model helps distinguish between individuals' current and previous jobs and accurately represents the relationships. (7) Similarly to the previous finding, the model also distinguishes between active educational enrollment and graduation. According to the ontology model (Figure 21), graduation is what proves and improves an individual's skills and competencies. For this reason, only the graduation details will be utilized in the dataset.

Secondly, based on the aforementioned findings from the ontology model, the skills and capability categorization table of information technology employees proposed by (Aktaş et al., 2022) was used to group individual skills under certain categories to form competencies. This study compiles IS-related skills, defines related keywords, and groups them under categories. The provided keywords were referred to in order to determine the categorization of each skill (Table 7). However, since this study and the skill table were only aimed at IS-related skills of Management Information Systems jobs, the table of information system-related skills and competencies coined by (Aktaş et al., 2022) needed to be updated and extended for skills and competence.

Functional expertise was reassessed within the scope of extending the skill and competence category list. Functional expertise specified in job advertisements for management information systems-related jobs refers to competencies such as process analysis, business development, and process analysis (Aktaş et al., 2022) (Table 7). However, when analyzing the functional specialization in the dataset, it covered a wider range across different fields. Therefore, functional specializations were also extended and updated within the scope of the aforementioned extension. In this context, functional specializations are categorized into specific domains to avoid oversimplification by putting distinct and unique competencies under a single broad category. To maintain depth and clarity in the data, functional specializations have been expanded into the following sub-categories: business, finance, accounting, human resources, logistics and supply chain, production management and planning, quality control, information systems, marketing, sales, management information systems.

Several new categories have been added to better represent the data and minimize the data loss. Foreign language skills have been added under the social category to represent competencies in foreign languages. In the development methodologies category, user interface and experience design (UI and UX design) have been added to cover the design of software and software-related interfaces. Cyber security has been added under the

Software development category to represent the importance of security in software environments. Also, hardware has been found to be a comprehensive and loose term. To fix this issue, sub-categories such as server, network and storage, virtualization, and cloud computing are defined to detail different aspects of hardware and infrastructure technologies.

Table 7

Codebook for Skills

	Category	Codes	Keywords
Organizational Information	**Social**	Teamwork	Teamwork, interpersonal skills
		Communication Skills	Communication Skills, Effective Communication, Excellent Communication, Oral and Written, Verbal Communicating, Written Communication, Presentation Skills, Strong Communication
		Responsibility	Work Independent
	Managerial	Leadership	Team Leader, Management Skill, Planning Skills, Team Management, General Management
		Project Management	Agile, Scrum, Jira, Project Management, Manage Projects
	Organisational	Functional Expertise	Process Management, Process Development, Process Flow, Process Improvement, Process Analysis, Business Development
		Sectoral Expertise	Department of*, Industry Experience
System Information	**Problem Solver**	Analytical Thinking	Analytical Thinking, Analytical Skills, Attention to Detail
		Problem Solving	Problem Solving
		Innovative	Innovative, Creative
	Methodologies	Basic IT	MS Office, Microsoft Word, Excel, Power Point, Computer Skills
		System Methodologies	Software Development Methodology, System Development Methodology, System Infrastructure, Sdlc, Cobit, Mvp
		Enterprise Modules	SAP, ERP, CRM, SCM, BI, Business Inteligence
		Analysis Tools & Techniques	Tableau, Power BI, MicroStrategy, SSRS, Business Object, QlikView, Oracle BI, Microsoft SSIS, ETL, OLAP
		Documentation	Documentations, Create Reports
		Data	Data Analysis, Data Gathering, Data Modelling, Data Sets, Data Warehouse, Data Mining etc.
Technical Information	**Software**	Software Development	Java, Python, Object-Oriented, Jquery, Java_Script, Css, Angular, Php, Node, React, Html, Perl, Ruby, Asp.Net, .Net, C/C++, Swift, C#
		Database	Mysql, Sql, Oracle Sql, Pl-Sql, T-Sql, Postgre Sql, Nosql, Mongodb
	Hardware	Hardware	Server, Hardware, Network

Source: Aktaş et al. (2022)

The skills dataset was updated using the aforementioned expanded categorization as a reference. During this update process, each skill name in the dataset was compared with the keywords from the table and categorized under the appropriate categories. To ensure the accuracy of this process, the most frequently mentioned skill entries, consisting of over a thousand entries and constituting approximately 75% of the total frequency of the skill data, were manually coded. Following this, the similarity between words was calculated through the fuzzy matching technique, employing a script developed using the FuzzyWuzzy library in Python. This helped systematically align the remaining data with the defined categories. Entries with more than 95% similarity to a word already coded and categorized under a specific category were appropriately labeled after a thorough review by the researcher. These meta-definitions were coded as categories, which were also referred to as competencies. In the scope of this research, these categories will be saved as new data columns that were created with reference to the previously mentioned tables, and they will be utilized in the model training.

Focus groups emerged as a research method in the 1950s, as a form of open-ended interviews in the form of group discussions (Templeton, 1994). Focus groups are structured discussions designed to gather opinions from a group of participants on a specific subject. They are usually conducted with three to five participants, and the discussion is managed by a moderator. Participants may be selected based on their personal characteristics or expertise on the discussed topic. One of the benefits of conducting this activity as a group is that it allows participants to build upon other's responses, enriching the discussion (Langford & McDonagh, 2003).

The focus group method is frequently used in market research, product design, business process design, and system usability discussions. Recently, it has become a preferred method in software development due to its ability to quickly and cost-effectively gather qualitative feedback from practitioners (Kontio et al., 2004). In their paper, Kontio et al., (2004) reflect on three cases where the focus group method was used to obtain feedback and experiences from software engineering practitioners and application users. Their findings indicate that focus groups are highly effective in collecting qualitative insights and feedback, particularly in the context of software engineering

The following steps should be followed in focus group research (Edmunds, 1999): The first step is to define the research problem. Focus group research is useful for identifying the research problems. This method enables the researchers to gather expert opinions, receive feedback on new concepts, brainstorm new ideas, and uncover the underlying motivations behind certain issues (Edmunds, 1999; Kontio et al., 2004). Within the scope of this research, a focus group will be organized to discuss the following research questions: “What features from the collected dataset can be used to train the artificial intelligence platform to determine a person's job suitability?”, “What new and unique feature, other than the already existing features and columns, can be obtained by combining, shortening, and summarizing different columns.”

The second stage is the planning of the focus group. Typically, focus group meetings last between 2-3 hours. This limited timeframe ensures that discussions stay within the scope. The focus group study carried out within the scope of this study took place in a 2 and a half hour time frame, allowing sufficient time to explore the subject matter thoroughly without straying off topic.

In the third stage, participants are selected. It is important to choose participants who can offer valuable insights and significantly contribute to the discussion. This will enhance the overall quality of the study. Depending on the research question and the aim of the study, selected participants may be completely unfamiliar with the subject, or they may have substantial expertise. For this research, the focus group consisted of four individuals: three participants who are experts in human resources (with PhD) and one expert in information systems, ensuring a well-rounded discussion with depth.

The fourth stage is conducting the focus group session. Effective time management and allowing all participants to speak are critical. For this reason, the session is led by a moderator who guides the discussion and maintains focus. Taking notes during the session is also important to capture insights during the dialogue and discussion. In this focus group session, PhD candidate served as the moderator and led the discussion. Also, one additional participant acted as an observer and was responsible for taking notes throughout the session.

The following findings were obtained from reviewing the meeting notes collected during the session:

“Duration of work experience in the position where individuals were most recently employed or are currently employed (new feature created from the combination of the existing dataset and features)”

“Average duration of employment (new feature created from the combination of the existing dataset and features)”

"The most recent position held by the individuals (new feature created from the combination of the existing dataset and features)"

"Individual's possession of specific skills and capabilities (confirmation of an already existing feature)"

"Individual's educational achievements, such as whether they have graduated from specific levels of education (confirmation and an update of an already existing feature)"

"Individual's educational achievements, such as the undergraduate programs the individual graduated from (confirmation of an already existing feature)"

The final feature list that will be used to train the machine learning and artificial intelligence models is shared in the appendices.

3.4.3.5. Algorithms

Within the scope of the study, the aim is to address the research problem by developing solutions with machine learning and artificial intelligence methods using data collected from professional social media platforms. Several machine learning models will be designed to achieve this, and the outcomes will be evaluated using various criteria. In order to overcome any uneven data distribution, oversampling was employed. Synthetic Minority Over-sampling Technique (SMOTE) is used to oversample by creating synthetic samples rather than just duplicating existing samples. Algorithms are chosen from the approaches of decision trees, artificial neural networks, probability-based techniques, distance-based techniques, statistics-based methods, and gradient boosting. The specific algorithms chosen from each approach and the reasons for their involvement are detailed below.

Support Vector Machine (SVM)

Support vector machine (SVM), developed by Vladimir Vapnik in the 1990s, is a supervised learning method rooted in statistical learning theory (Vapnik, 1998). An SVM constructs hyperplanes to classify inputs in high-dimensional space. The support vectors are the closest values to the classification margin. SVM aims to maximize the margin between the hyperplane and the support vectors (Gove & Faytong, 2012). The function of the SVM approach can be expressed as shown:

$$Z = f(y) = sign(\sum_{i=1}^{N} y_i p_i K(x, x_i) + c) \tag{3.1}$$

In this equation, p_i and c are the parameters of the hyperplane, while $K(x, x_i)$ is the kernel radial basis function. The formulation *sign* is the expression of the signum function (Jabeur et al., 2021).

SVM can be used for classification and regression problems, as it is an effective classifier for both linear and nonlinear problems (Marneni & Vemula, 2022). SVMs are a widely used type of classifier that is commonly regarded as one of the best off-the-shelf choices (Gove & Faytong, 2012).

This study employed the SVM algorithm to assess an individual's suitability for a certain career field or position. The SVM model was developed in Python using the sklearn software library. The results of the model will be compared with the results of the other models developed as part of the study. As a result, the best-performing models will be included in the ensemble learning bucket of models. This ensemble approach will leverage the strengths of each model to enhance overall prediction accuracy and robustness.

Logistic Regression (LR)

Logistic regression is a fundamental method that was initially formulated by David Cox in 1958. It builds a logistic model that is advantageous in two aspects. It can be used for classification and class probability estimation because it is tied to logistic data distribution (Bartosik & Whittingham, 2021). Logistic regression is an algorithm utilized to classify data into distinct categories. In supervised classification scenarios where there are multiple classes, the algorithm's objective is to distinguish the decision boundaries separating the different groups. These decision boundaries effectively separate one class from another and can vary in complexity and shape depending on the specific context. The logistic regression model assigns weights to the features, which are used to calculate a value between 0 and 1 for each feature vector using the S-shaped logistic function. This value represents the probability of an example belonging to a specific class. The algorithm adjusts the weights to ensure the training examples are correctly classified (Gudivada et al., 2016).

The formula of LR is as follows:

$$F(x) = \frac{1}{1+e^{-(\beta_0+\beta_1 x)}} \qquad (3.2)$$

The output of the Logistic Regression, the score, can be expressed as shown:

$$Z = Log\frac{p_i}{1-p_i} = \sum_{I=1}^{N} \beta_i x_i + \beta_0 \quad (3.3)$$

Here $\beta_0 + \beta_1 x$ is similar to the linear model y = ax + b. The logistic function applies a sigmoid function to restrict the y value from a large scale to 0–1 (Sinnott et al., 2016).

This study employed the logistic regression algorithm to assess an individual's suitability for a certain career field or position. The logistic regression model was developed in Python using the sklearn software library. The results of the model will be compared with the results of the other models developed as part of the study, and if they are found to be successful, they will be included in the ensemble model.

K-Nearest Neighbour (KNN)

KNN is a pattern detection algorithm that calculates distances in the data space and produces a classification rule. In KNN, classification is based on the nearest neighbor rule. The nearest neighborhood is the detection of the closest point for each sample by using the locations in the data space with a distance measurement. As the size of the data set increases, the computation time will also increase as the distance of each sample to all other samples is measured.

The KNN algorithm performs classification based on the K number of nearest neighbors. The classification function of KNN is as follows:

$$P_r(Y = j|X = x_0) = \frac{1}{K}\sum_{i \in N_0} I(y_i = j) \quad (3.4)$$

In this formula, K refers to the number of nearest neighbors.

There is not a fixed value for K in KNN. The value of K varies according to the size of the sample, the number of attributes in the sample, the number of dependent variables, and the distribution of the sample in the data space. However, the value of K is usually chosen among odd numbers such as 3, 5, 7, and 9. This is to prevent the number of different class values from being equal. A common approach is to test different K values and to compare the model accuracies (Aksu & Yıldızçakar Sarıoğlu, 2022).

KNN will be one of the employed algorithms to assess an individual's fit for a certain position. The results of the model will be compared with the results of the other models developed as part of the study, and if they are found to be successful, they will be included in the ensemble model.

Random Forest Classifier (RFC)

Random forest (RF) is a nonparametric ensemble supervised machine learning model introduced by Breiman (Breiman, 2001). It contains thousands of decision trees, and each tree performs classification, and the class that receives the most votes is the output of the model. RF is one of the most popular learning algorithms because it can handle a large, complex, and multidimensional dataset, and it requires a smaller number of parameters to be used compared to other models (Belgiu & Drăgu, 2016; Dutta et al., 2022; Shahabi et al., 2019). Because of multiple trees and random sampling, a random forest is less likely to have problems such as overfitting in the training process, leading to high-accuracy results (Arabameri et al., 2019; Caie et al., 2020). RF classifier is easy to implement, fast responding, quick to deploy and has been used successfully in many domains (Caie et al., 2020).

The output function of RF is calculated as the following equation (Katuwal et al., 2020):

$$Z = argmax\,\frac{1}{T}\sum_{t=1}^{T} p_t(y/x) \qquad (3.5)$$

In this equation $p_t(y/x)$ shows the probability distribution of each tree *(t)* and *x* represents the test sample (Jabeur et al., 2021).

This study employed the random forest algorithm to assess an individual's suitability for a certain career field or position. The RF model was developed in Python using the sklearn software library. The results of the model will be compared with the results of the other models developed as part of the study, and if they are found to be successful, they will be included in the ensemble model.

Gradient Boosting Classifier (GBC)

Gradient Boosting Classifiers (GBCs) are a type of machine learning algorithm that sequentially improve the accuracy of output by fitting new models. The algorithm creates new base learners that are highly correlated with the negative gradient of the loss function associated with the ensemble. The core concept behind GBCs is to iteratively add models to the ensemble, which corrects errors made by previous models and results in a more accurate estimate of the response variable. There are various loss functions used in machine learning, which researchers can choose from depending on the specific task at hand. This flexibility allows for high customization of GBCs, making them ideal for data-driven tasks and providing significant freedom in model design. As a result, identifying the optimal loss function often requires trial and error (Natekin & Knoll, 2013).

The boosting approximation function of the gradient boosting approach is as follows:

$$Z = F(x_i) = \sum_{j=1}^{M} \beta_j h(x; b_j) \tag{3.6}$$

In this equation, $h(x;b_j)$ stands for base learners, x stands for explanatory variables, *ßj* stands for expansion coefficients, and *bj* stands for model parameters (Jabeur et al., 2021).

Gradient Boosting algorithms are easy to implement, which allows experimentation with different model configurations. Additionally, Gradient Boosting Machines (GBMs) have shown to be highly effective in practical applications and have successfully tackled various challenges in the fields of machine learning and data mining (Bissacco et al., 2007; O'Sullivan et al., 2009; Pittman & Brown, 2011).

This study employed the GBC algorithm to assess an individual's suitability for a certain career field or position. The GBC model was developed in Python using the sklearn software library. The results of the model will be compared with the results of the other models developed as part of the study. As a result, the best-performing models will be included in the ensemble learning bucket of models. This ensemble approach will leverage the strengths of each model to enhance overall prediction accuracy and robustness.

Extreme Gradient Boosting (XGBooost) Classifier

Extreme Gradient Boosting (XGBoost) is a machine learning approach that works by aplying gradient boosting to decision trees (Moslemi, 2023; Subasi et al., 2022). Gradient boosting framework enables parallel decision tree augmentation, solving data science problems faster (Schneider & Xhafa, 2022).

In more specific terms, XGBoost will iteratively build small and simple decision trees. These trees have high bias, low performance, and poor learning. Subsequently, XGBoost builds a new tree that aims to address the limitations of the proceeding tree and will do what the low-performing tree cannot do. This approach iteratively generates new trees that aim to correct errors until the criteria (such as the number of trees/estimators) are met (Subasi et al., 2022).

The predictive function of the XGBoost algorithm is as follows:

$$Z = F(x_i) = \sum_{t=1}^{T} f_t(x_i) \tag{3.7}$$

In this equation, x_i refers to the explanatory variables, $ft(x_i)$ refers to the output tree (Jabeur et al., 2021).

XGBoost has high training and execution speed. Its parallelization and distributed computing support enable it to be an agile approach with fast results for large projects. This approach has gained significant popularity in recent years and has been used by many ML competition-winning teams (Schneider & Xhafa, 2022; Subasi et al., 2022).

This study employed the XGBoost algorithm to assess an individual's suitability for a certain career field or position. The XGBoost model was developed in Python using the xgboost software library. The results of the model will be compared with the results of the other models developed as part of the study. As a result, the best-performing models will be included in the ensemble learning bucket of models. This ensemble approach will leverage the strengths of each model to enhance overall prediction accuracy and robustness.

Categorical Boosting (Catboost) Classifier

CatBoost is a novel gradient-boosting approach proposed by Prokhorenkova and Dorogush (Dorogush et al., 2018; Prokhorenkova et al., 2018). This approach excels at processing categorical data with minimal loss. Catboost differs from other gradient-boosting algorithms in three main points. (1) CatBoost uses ordered boosting, (2) it performs well in smaller datasets, and (3) it is capable of handling categorical data. The algorithm used by CatBoost has the added advantage of performing random permutations to estimate leaf values while selecting the tree structure. This helps overcome the overfitting caused by traditional gradient-boosting algorithms. The base predictor utilized by CatBoost is binary decision trees.

The estimated output algorithm of the CatBoost algorithm, as proposed by Dorogush, is as follows:

$$Z = H(x_i) = \sum_{j=1}^{J} c_j 1_{\{x \in R_j\}} \tag{3.8}$$

In this equation $H(x_i)$ represents the decision tree function formed by the explanatory variables *(xi)* and *Rj* represents the disjoint region corresponding to the leaves of the tree structure (Jabeur et al., 2021).

This study employed the CatBoost algorithm to assess an individual's suitability for a certain career field or position. The CatBoost model was developed in Python using the catboost software library. The results of the model will be compared with the results of the other models developed as part of the study. As a result, the best-performing models will be included in the ensemble learning bucket of models. This ensemble approach will leverage the strengths of each model to enhance overall prediction accuracy and robustness.

Adaptive Boosting (AdaBoost) Classifier

Adaptive Boosting, AdaBoost, was introduced by Yoav Freund and Robert Schapire (Freund & Schapire, 1997), and it is used for assembling multiple classifiers to merge them into a stronger classifier (Subasi, 2020). AdaBoost is based on a progressive structure that aims to predict values not correctly classified in the previous round. Adaboosts are resilient to overfitting, used in both binary and multiclass classification, and usually have high accuracy (Chakraborty et al., 2019; Subasi, 2020).

This study employed the AdaBoost algorithm to assess an individual's suitability for a certain career field or position. The AdaBoost model was developed in Python using the sklearn software library. These were chosen to optimize the model's performance for the given application. The results of the model will be compared with the results of the other models developed as part of the study, and if they are found to be successful, they will be included in the ensemble model.

Neural Networks

The concept of neural networks was first introduced by McCulloch and Pitts in 1943. Their proposal demonstrated that neural networks can be utilized to represent logical relationships, such as "and" or "or" operations. Later, they discovered that this model could also simulate and replicate the brain's natural pattern recognition and classification abilities. By implementing neural networks, machines can learn to classify specific patterns (Kiang, 2003).

Neural Networks are a machine-learning method formed by a combination of simple information-processing units called neurons. They are highly capable of learning non-linear relationships. The perceptrons, which were inspired by neural cells, are simple perceptrons with multiple binary inputs and outputs. The output of a perceptron is calculated by multiplying each input by the corresponding weights (representing the importance of each output) and summing all of the multiplications. If the result is above the threshold, the output will be 1. Otherwise, the perceptron will output 0.

The sigmoid neuro, which is a more sophisticated iteration of the perceptron model, differs from the perceptron. The output of the sigmoid neuron, which was calculated similarly to the perceptron, is obtained by additionally applying the sigmoid function. As a result, the output of the sigmoid neuron perceptron can generate any value between 0 and 1 rather than either 0 or 1. This enables its use as a non-linear classifier.

The neural network structure is formed by organizing neurons in layers. In this structure, there is an input layer, an output layer, and hidden layers. The layer where the data is received is called the input layer, and the layer where the processed data is output is called the output layer. Layers without input and output registers are called hidden layers. The number of neurons in the input and output layer is determined by the problem, while there are no strict rules for determining the number of neurons in the hidden layer (Ölmez & Er, 2022; Nilsson, 1998).

The output neuron in the neural network algorithm is expressed by the following function:

$$Z = \left(\sum_{j=1}^{M} g\left(\sum_{j=1}^{M} w_{ij}x_i + \alpha_j\right)w_{jk} + \alpha_k\right) \qquad (3.9)$$

In the equation, w_{ij} and w_{jk} are the weight values in the hidden and output layers. g is the activation function, x_i is the input ratio, α_j is the bias value of the hidden neurons and α_k is the bias value of the predicted variable.

This study employed the neural network algorithm to assess an individual's suitability for a certain career field or position. The neural network model was developed in Python using the TensorFlow software library. The results of the model will be compared with the results of the other models developed as part of the study, and if they are found to be successful, they will be included in the ensemble model.

3.4.3.6.Ensemble Learning

Ensemble learning is a machine learning approach that works by employing many machine learning algorithms in cooperation and combining their results with various methods (Schneider & Xhafa, 2022; S. Wang & Summers, 2012). This approach is based on training multiple base learning systems with different algorithmic approaches with a single data set and combining the results of these base learning systems with an ensemble approach to obtain a single output. This leads to better performance than a single ML method (Sugiyama, 2016; Yang, 2017) and eliminates the variance between results (Schneider & Xhafa, 2022).

The types of ensembling techniques are (1) Bayes optimal classifier, (2) Bayesian parameter averaging, (3) Boostrap aggregating (bagging), (4) Boosting, (5) Stacking and (6) Bucket of models (Gudivada et al., 2016). In this study, bucket of models approach will be used. The bucket of models approach is applied by selecting and combining the most successful algorithms among many models. In this context, the results obtained from Random Forest Classifier, Logistic Regression, K Nearest Neighbor (KNN), Support Vector Machine (SVM), Extreme Gradient Boosting Classifier (XGBoost), Gradient Boosting Classifier (GBC), Adaptive Boosting Classifier (AdaBoost) and Neural Networks will be evaluated, and the three (3) algorithms with the highest results according to the evaluation criteria will be selected for the ensemble. Then, the job fit value for each position will be obtained through 'soft' voting and used for the system.

3.4.4. Evaluation of AI/ML Models (Design Cycle, Step 8)

In this research, a distinctive methodological framework is employed, involving the development of individualized classification models for each distinct IT position group. This approach diverges significantly from traditional classification methods, which typically categorize each input into a single, predetermined group. Instead, separate models are constructed to assess the suitability of an individual's professional profile for each specific group independently. This multiple model focused strategy enhances the precision of job fit assessments by allowing each model to evaluate an individual's compatibility based on their unique set of skills, experiences, and educational background. Consequently, it is feasible for an individual to concurrently align with multiple position groups or potentially align with none, reflecting a more realistic and nuanced interpretation of career data. This methodological choice not only provides a granular understanding of an individual's placement within the IT sector but also aids in personalized career planning and development. This approach was shared in detail in Chapter 4.

The models developed with these artificial intelligence approaches will be compared according to the evaluation criteria, and the approaches with the best results will be adopted in the research process (Aksu & Yıldızçakar Sarıoğlu, 2022). Confusion Matrix is a tabular visualization of model predictions against their true labels. The matrix compares the actual target values with those predicted by the machine learning model. True Negative indicates the absence of a condition or characteristic. True Positive indicates the presence of a condition or characteristic. False Negative indicates the absence of a particular condition or characteristic. False Positive indicates that a particular condition or characteristic is present.

Table 8

Confusion Matrix

		Predicted	
		Negative	**Positive**
True	**Negative**	True Negative (TN)	False Positive (TP)
	Positive	False Negative (FN)	True Positive (FP)

Accuracy is a statistic that shows the correct prediction rate. This rate indicates the accuracy of the trained algorithm. However, this statistic will be reliable with the assumption that the observations used to train the algorithm are representative of all situations that can occur in real life. For this statistic to be reliable, the data set must be homogeneous and representative of the real world.

$$Accuracy = \frac{\sum TP + \sum TN}{All\ Observations} \tag{3.10}$$

Recall measures the model's ability to detect positive examples. It is the ratio of correctly predicted true observations to all known true observations in the data set.

$$Recall \ = \frac{TP}{TP+FN} \tag{3.11}$$

Precision assesses the proportion of accurately classified instances among those identified as positive outcomes, effectively measuring the accuracy of a model in identifying positive classifications.

$$Precision = \frac{TP}{TP+FP} \tag{3.12}$$

F-Score is a popular metric that combines precision and recall. It is called the F score, which is the harmonic mean of precision and recall, and takes a value between 0 and 1.

$$FScore = \frac{2*Precision*Recall}{Precision+Recall} \tag{3.13}$$

The metrics of accuracy, recall, precision, and F-score are used because they offer a comprehensive evaluation of the models' effectiveness in different aspects. Accuracy provides an overall performance metric that is easy to understand and useful when the importance of false positives and false negatives is about the same. Recall is important when it is crucial to identify as many suitable job roles as possible, such as ensuring no suitable job roles are overlooked. Precision is essential to avoid suggesting irrelevant job roles and ensure the quality of recommendations. The F score provides a single metric that accounts for both false positives and false negatives.

The evaluation of the algorithms developed within the scope of this study according to the evaluation criteria is shared in the findings section in Chapter 4.

3.4.5. Explaining the Prototype Development (Design Cycle, Steps 9 and 10)

In this thesis, within the scope of the deployment step, which is the final stage of the CRISP-DM approach and in line with the Design Science Research approach, a prototype was developed to demonstrate the application's basic functionality. This prototype also incorporates the active use of artificial intelligence models developed using the ensemble learning method.

The prototype application calculates a person's job fit (suitability) among the six IT positions within the scope of the research (Data Analysis and Business Intelligence, Software Development, Product and Project Management, Quality Assurance and Test, System Development and System Engineering, IS Consultancy, Strategy and Design) using the probability result of the ensemble model. The system presents these results in a visualized form, allowing the individual to select a target position. Following the selection of the target position, the individual is offered skill development opportunities relevant to that position. The system provides skill suggestions in two different categories personalized for the individual. First, the system lists the skills that are relevant to the chosen position and that the individual already possesses. These skills can be selected to be developed and improved further. The second section lists the skills that are relevant to the chosen position but that the individual does not yet possess.

The prototype was developed using the Firebase development platform. The front end was developed using HTML, CSS, and JavaScript to ensure compatibility across devices and browsers. AI/ML models, developed using ensemble learning techniques, were integrated to calculate job-fit scores and provide skill recommendations using Python programming language at the back end.

Two different use cases that show the functionality and feasibility of the system are shared in Chapter 4.

3.5.Documenting to the Knowledge Base (Rigor Cycle, Step 11)

The contribution of this design science research study to the knowledge base has been systematically documented through various academic publications and the PhD thesis itself. Throughout the thesis process, academic publications that are related to the thesis and contribute to its knowledge base are summarized as follows:

Aktaş and Akbıyık (2019) aim to demonstrate how creating social media profiles and engaging in professional social networking sites can influence an individual's social capital and, in turn, how this social capital impacts the value of their network. In another study, Aktaş et al. (2022) aim to clarify the expectations of Turkish employers from graduates of the management information systems (MIS) department. Additionally, the study highlights the essential skills and competencies that MIS graduates should possess based on labor market demands, as evidenced by job postings. The research further enhances the understanding of MIS in Turkey by comparing job postings between the United States of America (USA) and Turkey. In the latest publication, Aktaş and Akbıyık (2023) examine the use of Ontology-Driven Conceptual Modeling (ODCM) with OntoUML to tackle complex business issues in Management Information Systems (MIS). It demonstrates how ontologically grounded models can effectively address system integration and migration challenges.

Ultimately, the process has been thoroughly documented throughout the doctoral thesis, aiming to contribute to the knowledge base.

CHAPTER 4. RESULTS AND FINDINGS

This section presents the results of the ontology-driven conceptual modeling, evaluation of the machine learning and artificial learning models, and the prototype developed within the scope of this study and as the artifacts of the design science research.

4.1. Ontology-Driven Conceptual Model with OntoUML

PSMPs host a wealth of information, including company details, available job openings, and user professional profiles. The user profiles on such platforms feature a range of career-related data, like employment history, educational background, certifications, professed skills, and abilities, among other details that enhance their job prospects. Typically, this information is presented in an online resume style and is generally publicly available. Although this information is valuable for networking and employer screening, it lacks semantic depth and an ontologically well-founded structure. This implies that these abundant data do not guarantee semantic interoperability and are not machine-actionable.

Within the scope of this study, it is aimed to model the data on professional profiles of individuals shared publicly on PSMPs (Professional Social Media Platforms) using the ODCM (Ontology-Driven Conceptual Modeling) approach and OntoUML (Ontology-based Unified Modeling Language), ensuring semantic rigor and fostering semantic interoperability.

In this project, an ontology designed to serve as a conceptual common model with clearly defined concepts to aid communication among professionals in this field was developed. Additionally, it acts as a domain model for the consistent development of applications and algorithms. Given its dual purpose as both a conceptual framework and a development tool, the ontology was created using the SABiO (Fablo, 2014), which is an ontology development methodology that has been inspired by software engineering practices (**See Chapter 3**).

Competence, skills, and certifications are key elements in human resource management, educational planning, and personal career development. However, understanding the intricate relationships between these elements can be challenging. The project aims to use the OntoUML to develop an ontological model that represents and relates individual competencies, skills, and certifications. In this study, it is aimed to map the relationships between a person's skills, competencies, and certifications. This will be done along the following competency questions:

- What are the structural components of individual competencies, and how do they relate to a job position and industry-recognized certifications?
- How can we ontologically model the relationship between a person and their educational information?
- How can we ontologically model the relationship between a person and their job experience?

This study aims to model the domain of PSMP, with a particular emphasis on career-related information.

In the data models for this study, the individual is positioned at the core, with connections to their skills, certifications, education, and work experience. However, to achieve a deeper understanding and develop innovative solutions within the domain of PSMP, there is a need to implement ontology and semantics.

The following parts explain the development of the ontology-driven conceptual modeling of PSMP career data. This consists of defining the UFO stereotypes and defining the relationships using OntoUML. The following section consists of several smaller sub-ontologies to make the process more understandable and digestible. Later, it is shared the final ontology in Figure 21.

For information on UFO Stereotypes and OntoUML, please see Chapter 2.

4.1.1. ODCM Sub-ontology of a Position and Job Experience

This part of the study aims to define a person's current or past job experience, to define the position, and also to model intricate relationships between a person, a position, and an employer via employment.

In Figure 18, a person is defined as a *kind,* which exists independently and doesn't change its identity through time or across different contexts. In the case of experience sub-ontology, a person can be identified by two *roles:* employee and former employee. A *role* is context-dependent and typically temporary. Their existence is dependent on certain conditions or relationships to exist. Employee refers to a person who is actively and currently working in a position in a company. A former employee refers to a person who no longer works in a certain position anymore. A person can be an employee and a former employee at the same time. This means a person can be an employee of one employer, while it can be a former employer of another (or the same) company, and this role can be related to the same or a different position.

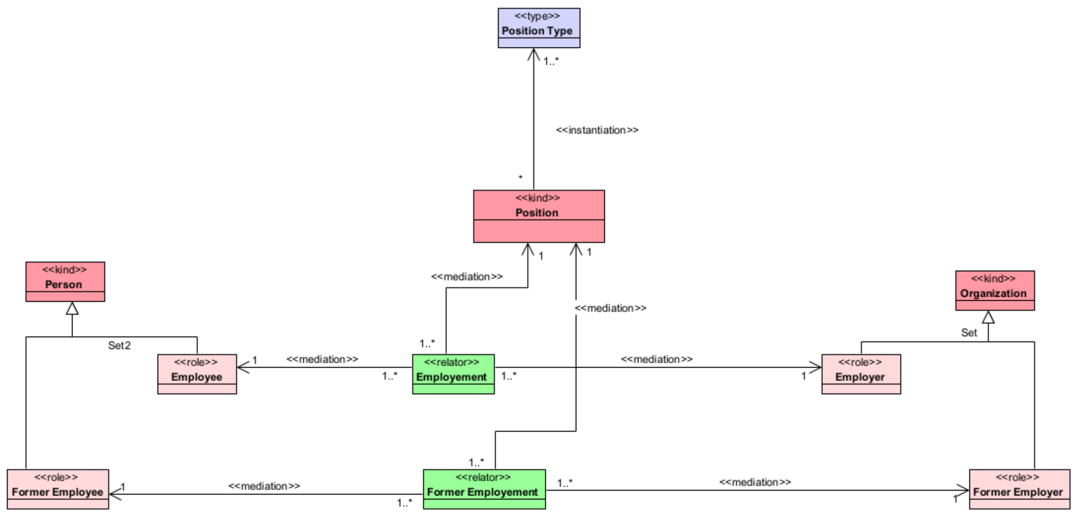
Organization is also defined as a *kind*, and it can be further detailed as two different *roles*, employer and former employer, depending on theis dependent on the context. Much like an employee, the definition of employer and former employer depends on the position's currency and activeness. A position, which is defined as a *kind,* refers to a job position that embodies the description of responsibilities and specific tasks. Position as a *kind* is an instantiation of a position type. A position type is a higher term that is beyond a specific company and is defined as a *type.*

The experience sub-ontology consists of two *relators*: employment and former employment. *Relator* helps encapsulate relationships and links the entities together. In this case, employment *relator* relates to an employee, employer, and a position with a mediation association. A former employment relator relates to a former employee, former employer, or a position with a mediation association.

This model helps us differentiate current and former employment, which was not evident in the database and ER models. It also guided us towards defining the position type, which would play an important role during ML model development.

Figure 18

Sub-Ontology of Experience

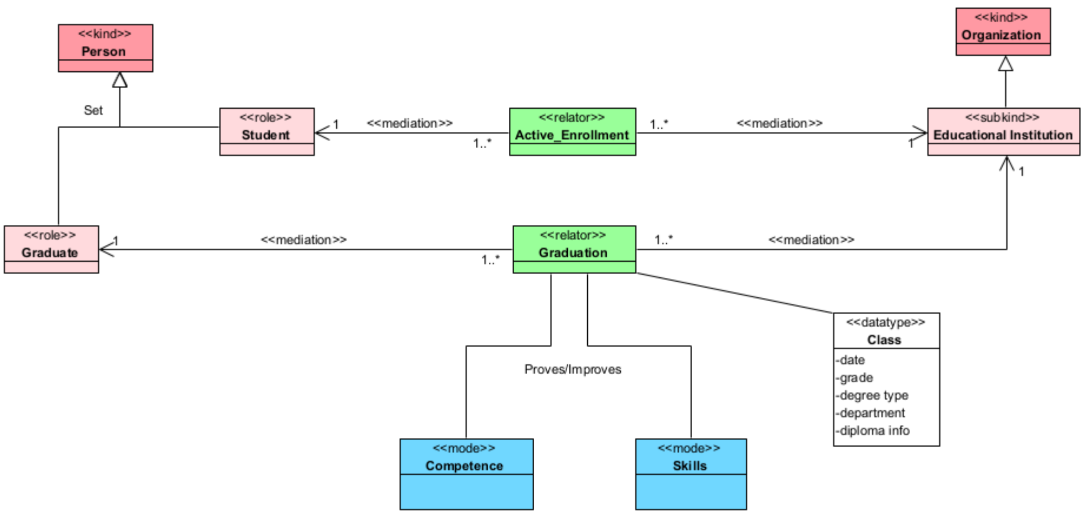


4.1.2. ODCM Sub-ontology of Education

This part of the study aims to define and ontologically model a person's education information. In Figure 19, a person is defined as a *kind,* which exists independently and doesn't change its identity through time or across different contexts. In the case of education sub-ontology, a person can be identified by two *roles*: student and graduate. A *role* is context-dependent and typically temporary. A student refers to a person who is actively enrolled in an educational institution. A graduate refers to a person who graduated from an educational institution. A person can be a student and a graduate at the same time. This means a person can be a student with an active enrollment in an educational institution while it has a graduation from another (or the same) educational institution. An educational institution is defined as a *sub-kind*, which is a kind of organization.

Figure 19

Sub-Ontology of Education



Education sub-ontology consists of two *relators*: an active enrollment and a graduation. *Relator* helps encapsulate relationships and links the entities together. In this case, the active enrollment *relator* relates to a student and an educational institution with a *mediation association*. The graduation *relator* relates a graduate and an educational institution with a *mediation association*.

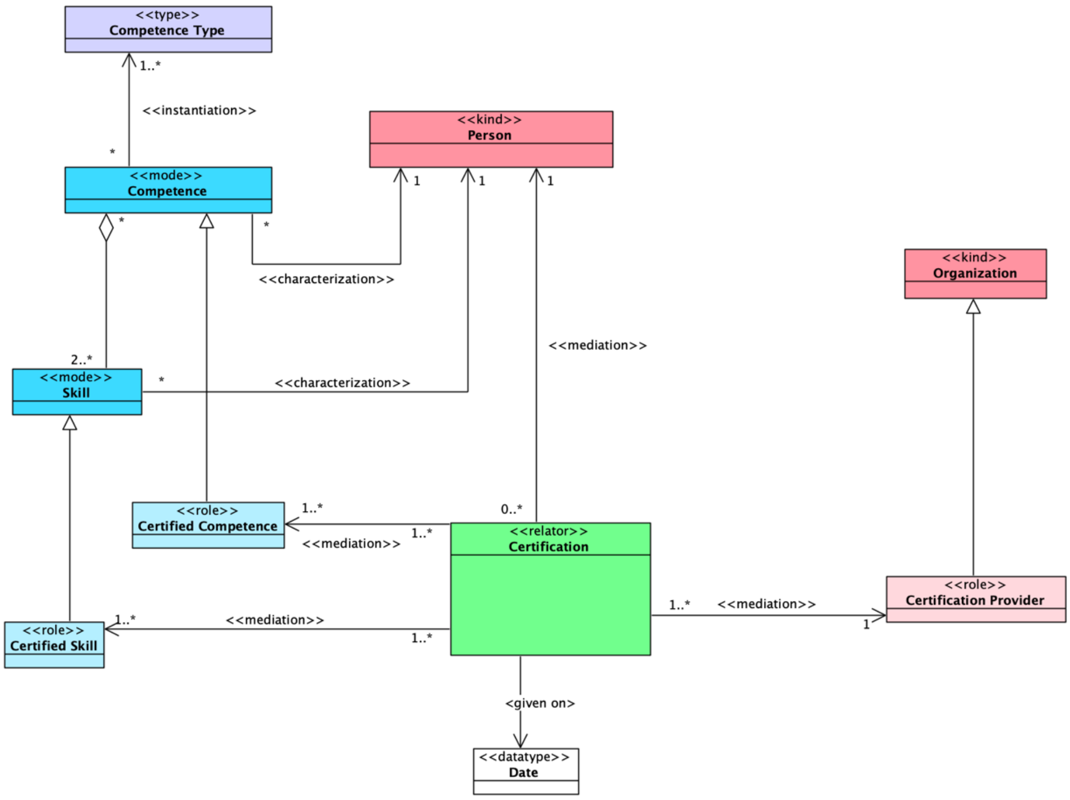
This model helps us differentiate active student enrollment and graduation, which was not evident in the database and ER models. It also forms a bridge with the competence and skills of a person, which will be discussed in the following sub-ontology.

4.1.3. ODCM Sub-ontology of Skills, Competence and Certification

This part of the study aims to define a person's skills, competence, and certification information in Figure 20. A person is defined as a *kind type,* which exists independently and doesn't change its identity through time or across different contexts.

Figure 20

Sub-Ontology of Skills and Certification



A person has competence and skill, both defined as a *mode*. Mode types are used to define intrinsic abilities. The association between a person and a competence and a skill is represented with a characterization. When skills and competencies are certified, they take the *role* of certified skills and certified competencies, respectively.

Competence as a *mode* inheres in an employee (it is a certain employee's competence), and it is intrinsic. However, the competence type as a *type* is the competence as a higher type. It is not intrinsic to a person; it is the general definition. For example, Bahadir's competence in Android App Development is inheres in Bahadır; thus, it is a *mode*. But the Android App Development competence itself, talking about competence and what it means, is a *type*. This model also indicates that competence consists of several skills.

The certification provider is defined as a *role*, which is an organization. Skills, competence, and certification sub-ontology has one *relator:* certification. *Relator* helps encapsulate relationships and links the entities together. Certification relates to a person, a certification provider, certified skills, and certified competencies. Certification also has a date on which it was given.

This sub-ontology helps us represent the skills and competencies of a person. It also lets us visualize the relationship between certification, skills, and competencies, which was not evident in the database and ER models.

4.1.4. Complete Ontology

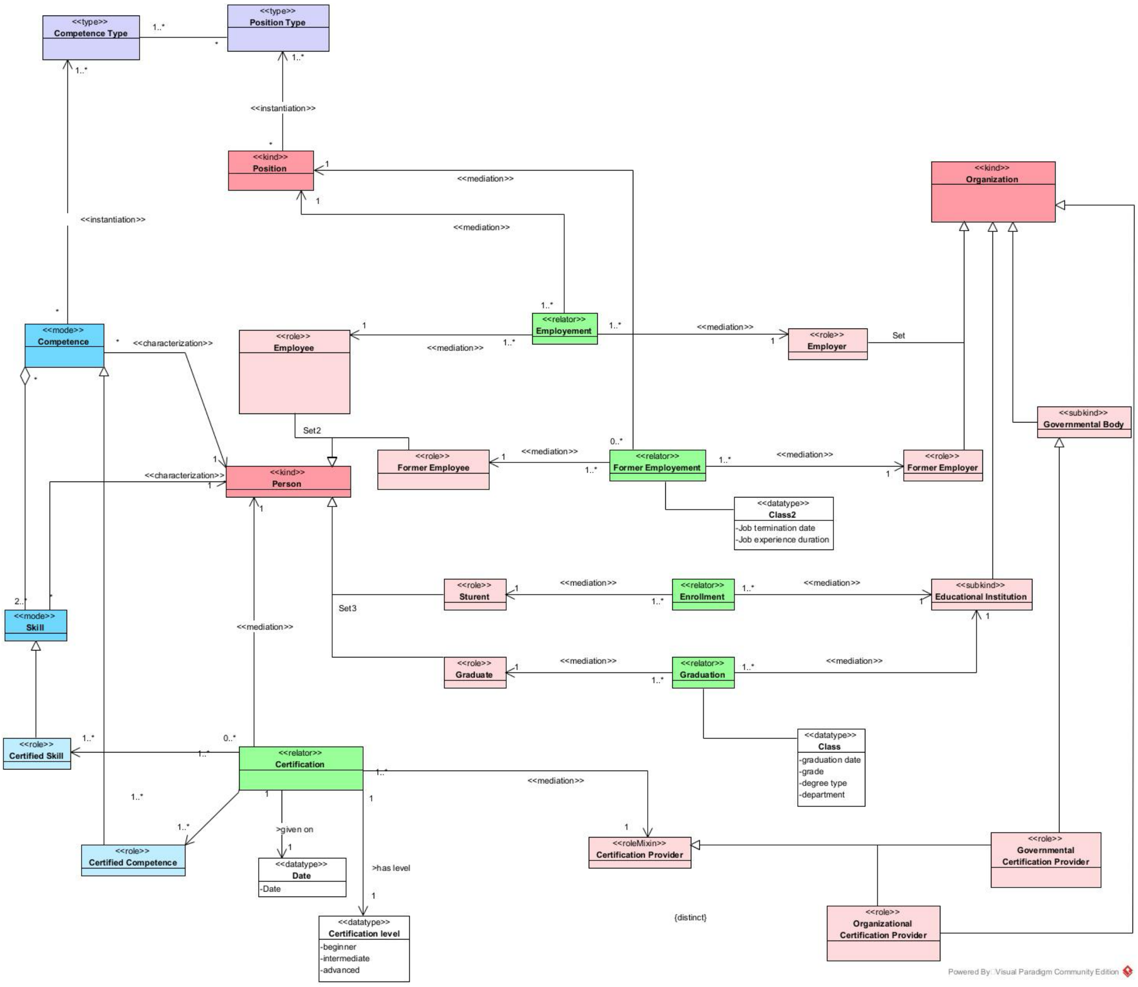
Figure 21 shows the complete ontology, which is an integration of all sub-ontologies already discussed beforehand. In addition, in this ontology, the competence type is associated with position type. This association explains that a position type (as a higher-level type) is related to one or many competence types (as a higher-level type).

This relationship indicates that a position type can be identified through the positions of an employee or a former employee. Also, competence types can be identified through the competencies and skills of a person. This line of thought indicates that a higher type of position can be identified via its instantiation in positions and its relationship with competence type. Thus, it can be argued that the aim of the thesis can be achieved following this line of thought and the findings gathered through ontological modeling.

The ontological model guides the study, especially in the data preparation and modeling phases.

Figure 21

Complete Ontology of Professional Data on PSMPs



4.1.5. Ontology Validation

Ontological modeling offers semantic richness and rigor that enhances the interoperability and machine-actionability of the domains or systems to which it is applied. In this study, ontology is a fundamental approach to organizing and structuring complex data structures. To validate the ontological modeling of professional data on PSMP developed in this study, consistency and syntax will be checked using the Visual Paradigm application, and literature-based validation will be performed. This validation process includes comparing the model with existing academic studies' outcomes.

Firstly, a syntax check was conducted to ensure there were no errors during the ontology modeling using OntoUML. The OntoUML Plugin in Visual Paradigm was utilized for this purpose, providing a tool for syntactical checks of the model and the diagram.

According to this tool, no syntactical issues were found in the diagrams and models created within the scope of this research.

For the literature-based validation, a comprehensive review was conducted of several studies to identify and discuss supporting evidence from these works.

Miranda et al. (2017) proposed an ontology-based model for competence management, offering a structured approach to representing professional competencies. This study views competencies as attributes that can be assessed, measured, and evaluated on an individual basis. This perspective aligns with and supports thesis's approach, which defines competencies as modes inherent within individuals. This model also posits that education enhances competencies and links competencies to various job roles and certifications. This approach mirrors the education sub-model (Figure19), which similarly connects training with the development of competencies and skills. The relationship between competencies, skills, and certifications in this study aligns with the definitions and the structure of certification, which are integral to the skills sub-ontology (Figure 20) and the complete ontology (Figure 21). The relationship established in this study between competence and job roles also corresponds to the higher-level relationship between competence types and position types in this research. This study made certain modeling choices that support a learning system, aligning with many aspects of the purposes of this study. This work differs from Miranda's in that work experiences are associated directly with individuals, and the modeling is structured around the data available on PSMP platforms.

Fazel-Zarandi and Fox (2012) aim to model human resources in dynamic environments. Their work and its findings critically overlap with and contribute to the validation of this research. In the ontological model developed within this study, skills and competencies are systematically defined and represented through a structural ontology. This structure demonstrates how these attributes are embedded in individuals. The ontology developed in Fazel-Zarandi and Fox (2012)'s work defines skills as attributes with varying levels of proficiency that enable individuals to perform specific activities. In this model, a skill is represented as a class with properties such as proficiency level and activity enabled, which aligns with the representation of skills as entities with defined attributes and relationships. This study's ontological model recognizes the evolving nature of skills and competencies, allowing updates and changes based on new information using PSMPs as a data source. Fazel-Zarandi and Fox (2012)'s work represents skills dynamically, allowing skill statements to change states (e.g., demonstrated, probable, possible, or refuted) based on new data and assessments. Therefore, both models incorporate the dynamic evaluation of skills. This study also identifies social networks as a source for the declaration of skills and competencies, which is of great importance in this study in the form of PSMP.

Azevedo et al. (2015) offer a comprehensive approach for modeling resources and capabilities within the context of enterprise architecture, employing a well-founded ontological framework. This study explains that resource elements represent type-level entities that capture the roles of objects within specific contexts, similar to how UFO was used to define the roles and relationships of competencies and skills. Both approaches emphasize the importance of capability-based planning. In their model, capabilities are connected to resources and behaviors that realize them. This aligns with this thesis's approach of defining skills as dynamic capabilities inherent in an individual, which can be reassessed and updated.

The work of Zaouga et al. (2019), which models human resource management using an ontology, also validates the ontological model developed in this study at several critical points. Zaouga's study defines roles within the human resources ontology, associates these roles with specific tasks and responsibilities, and explicitly models the competencies required to fulfill these roles. This study supports the ontology model in this research by associating roles with competencies. Both studies emphasize the importance of linking individuals to competencies. Another point of alignment is that Zaouga's study connects roles and competencies with the Person class to represent the qualifications that each individual possesses.

Tarasov (2012) presents a comprehensive approach to competence profile management using ontology. This approach plays a crucial role in validating this work, particularly in the representation and modeling decisions related to roles, workers, and competencies. In this ontology, the role class is associated with the competencies required to perform a job through worker. This model includes a direct relationship between the "Worker" class and the "Competency" class, where workers are linked to their competencies. The study underscores the importance of this approach for the effective management of an organization's workforce. Vladimir Tarasov's integration of roles and competencies, representation of worker and competency relationships, use of formal ontological structures, and formalization of operations on competence profiles align closely with the modeling choices.

Calhau and Almeida (2022) integrated competencies into enterprise architecture modeling with a zoom in approach and their outputs are aligned with this study in the following points and are useful for its validation. Calhau's work aligns with thesis' approach by emphasizing the integration of Knowledge, Skill and Attitude (KSA) to define competencies. In Calhau and Almeida's ontology, competencies are represented as dispositions that combine knowledge, skills, and attitudes, which mirrors this study's approach of modeling competencies as comprehensive constructs incorporating KSA. However, the model was limited to skills and left out knowledge and attitudes since they were not represented in PSMP data. They propose a hierarchical model for competencies, distinguishing between basic competencies and complex competencies as a part of personal competence, which is composed of multiple sub-competences and skills. This approach and modeling choices align with this thesis's modeling of skills as a mode, competence as a mode and competence type as a type in a hierarchical structure.

These alignments ensure that the model is both theoretically sound and practically applicable, providing a robust framework for competency management in professional social media platforms.

4.2. Evaluation of the AI/ML Models

The research employed the following approach to assess the suitability of an individual who has entered their professional career information into the system for a specific IT position (also referred to as a job fit in this study). Models were developed using the training dataset to determine how well an individual fits a particular job position requirement. This shows if a particular individual is related or unrelated to a position based on their professional career data. These models were individually developed for each group. This means the system calculates the job fit of the individual separately for each position group. This approach, which was described in a recent study (Cebeci, 2022), and was successfully implemented in several studies (Cebeci et al., 2023; Güner et al., 2023), diverges from traditional classification methods. Traditional classification methods try to find the right group for each input. If there are six classifications, as is in this study, these methods will try to label an input as one of the six classification groups.

However, this approach conducts a separate classification analysis for each group, determining whether an individual input belongs to each of these different groups. Consequently, this analysis may conclude that an individual does not belong to any group or, conversely, may belong to more than one group. However, if the analysis is conducted using a probabilistic machine learning method, it is possible to calculate the individual's level of fit with each group. This approach aligns with the predictive method intended to be implemented for this study, enabling a nuanced understanding of how closely individuals match different position groups.

This alignment is particularly evident in the following stages. (1) Within this study, it is likely that an individual using the proposed system will match with multiple position groups. In this case, in a business environment, where skills, competencies, and experiences in the IS sector are mostly transferable, the system may identify an individual as being suitable (or having a high job fit) for several positions. This potential outcome is significant as it allows individuals to realize their potential to transition between different positions and to compare their profiles with the requirements of various positions by inspecting job fit. This understanding can be instrumental in guiding career development and strategic job planning. (2) It is also possible that an individual using this proposed system does not achieve a high match (a high job fit) with any specific position, which would be normal for someone planning to enter the information system labor market without any previous experience. In such cases, the individual can utilize the second stage of the system, which focuses on micro-credential recommendations. This feature provides recommended micro-credentials required for the relevant information systems positions. Following this, the individual can then take steps toward their target positions by acquiring the required micro-credential through lifelong education, distance learning, boot camps, or certificate programs. This process aims to enhance their qualifications and align their skill set with the demands of the job market. (3) If the individuals using the system proposed in this study match with a single position group, they discover a significant level of compatibility and similarity between their skills and competencies and the requirements of the particular position. This discovery can be highly beneficial because it may identify where an individual's current abilities are best aligned. This potentially guides them toward a more focused career development process and job opportunities.

In the findings section of the study, a total of nine different machine learning models were developed for each position. The algorithms used to develop these models include Random Forest, CatBoost, K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Gradient boosting Classifier (GBC), AdaBoost, Logistic Regression, and Artificial Neural Networks. The models developed for each position were evaluated on accuracy, precision, recall, and F1-score. Then, the three most successful algorithms were selected to be included in the ensemble learning model.

4.2.1. Data Analysis and Business Intelligence Models

Within this study's scope, data analysts, data scientists, business analysts, and business intelligence specialists were grouped under the Data Analysis and Business Intelligence position. The machine learning and artificial intelligence models for this group, which focus on data analysis and business intelligence-oriented job positions, are detailed in Table 9 below.

Table 9

Results of Data Analysis and Business Intelligence Models

DA&BI Algorithm	accuracy	precicsion	recall	F-score
Random Forest Model	0,828302	0,839599	0,828302	0,826862
CatBoost Model (*)	0,876730	0,878226	0,876730	0,876608
XGBoost	0,836478	0,836555	0,836478	0,836469
KNN Model	0,784906	0,786883	0,784906	0,784534
SVM Model (*)	0,913836	0,914551	0,913836	0,913799
Gradient Boosting Clasifier	0,863522	0,863882	0,863522	0,863488
AdaBoost Clasifier	0,832075	0,843831	0,832075	0,830628
Logistic Regression (*)	0,914465	0,914633	0,914465	0,914457
Neural Networks	0,862264	0,869466	0,862264	0,861589

Based on the evaluation results of the machine learning and artificial intelligence models (Table 7), the CatBoost, Support Vector Machine, and Logistic regression algorithms are seen to be the three most effective models for predicting an individual's fit for data analysis and business intelligence positions. The Logistic Regression and Support Vector Machine based models, with an F-score above 0.91, were identified as the most successful. Three successful models were chosen to be a part of an ensemble model design to predict the fit of the individuals for data analysis and business intelligence positions. This approach aims for the creation of a validated and effective final model by integrating different successful models into a single model.

4.2.2. Software Development Models

Within this study, software developers, software engineers, mobile application developers, and game developers were grouped under Software Development. The performance of the ML and AI models, developed to assess an individual's fit for software development related positions, is presented in Table 10.

Table 10

Results of Software Development Models

Soft. Dev. Algorithm	accuracy	precicsion	recall	F1-score
Random Forest Model (*)	0,869433	0,869774	0,869433	0,869403
CatBoost Model (*)	0,874494	0,875701	0,874494	0,874393
XGBoost	0,856275	0,859963	0,856275	0,855906
KNN Model	0,823887	0,827516	0,823887	0,823397
SVM Model (*)	0,875506	0,877189	0,875506	0,875367
Gradient Boosting Clasifier	0,865385	0,865926	0,865385	0,865335
AdaBoost Clasifier	0,846154	0,846614	0,846154	0,846103
Logistic Regression	0,869433	0,869689	0,869433	0,869411
Neural Networks	0,829959	0,829981	0,829959	0,829956

According to the results presented in Table 8, the three most successful models for predicting an individual's fit for software development positions are CatBoost, Support Vector Machine, and Random Forest. Notably, the models developed using CatBoost and SVM algorithms stand out as the most effective, each achieving an F-Score above 0.87. Consequently, these three models - CatBoost, SVM, and Random Forest - were chosen to implement in an ensemble model. By integrating these successful models, researchers aim to provide a final model that is more robust and accurate than its parts.

4.2.3. Product Specialist and Project Management

Within this study's scope, the product manager, product owner, product specialist, and project manager were grouped under the Product Specialist and Project Management positions. The evaluation results of the machine learning and artificial intelligence models developed within the scope of this research, which aim to predict an individual's suitability for product specialist and project management positions based on their career information, are shared in Table 11.

Table 11

Results of Product Specialist and Project Management Models

Algorithm	accuracy	precision	recall	F1-score
Random Forest Model	0,849476	0,851793	0,849476	0,849228
CatBoost Model (*)	0,856675	0,858964	0,856675	0,856447
XGBoost	0,852094	0,852213	0,852094	0,852082
KNN Model	0,827225	0,827227	0,827225	0,827225
SVM Model (*)	0,887435	0,891515	0,887435	0,88714
Gradient Boosting Clasifier	0,846204	0,846932	0,846204	0,846123
AdaBoost Clasifier	0,829843	0,835592	0,829843	0,829111
Logistic Regression (*)	0,884162	0,886020	0,884162	0,884023
Neural Networks	0,837041	0,844579	0,837041	0,836145

According to the results, the three most effective algorithms are CatBoost, Support Vector Machine, and Logistic Regression. In particular, the models created using Logistic Regression and SVM algorithms performed well by achieving an F-score above 0.88. Therefore, these three models, including CatBoost, were chosen to be incorporated in the final ensemble model, which aims to predict the fit of individuals for information system job positions in product and project management.

4.2.4. Quality Assurance and Test

Within this study's scope, quality assurance specialists, software test specialists, software test engineers, and software test automation engineers were grouped under the quality assurance and test positions. The evaluation results of the machine learning models, which aim to assess the suitability of individuals for quality assurance and test positions, developed as a part of this research are presented in Table 12.

According to the results, the three most effective algorithms for predicting an individual's suitability (fit) for information system related quality assurance and testing positions were CatBoost, SVM, and Logistic Regression. The models developed with Logistic Regression and SVM performed especially well, each having an F-score over 0.90.

Meanwhile, the CatBoost model also performed well by obtaining an F-score of 0.82. Therefore, these models were selected to be integrated into an ensemble model designed to predict the suitability of individuals for the job position in quality assurance and testing.

Table 12

Results of Software Quality Assurance and Test Models

Algorithm	accuracy	precicsion	recall	F1-score
Random Forest Model	0,774194	0,807645	0,774194	0,767884
CatBoost Model (*)	0,823775	0,835897	0,823775	0,822171
XGBoost	0,810633	0,810957	0,810633	0,810584
KNN Model	0,783154	0,783709	0,783154	0,783048
SVM Model (*)	0,919355	0,921043	0,919355	0,919274
Gradient Boosting Clasifier	0,787336	0,801596	0,787336	0,784792
AdaBoost Clasifier	0,819594	0,835091	0,819594	0,817484
Logistic Regression (*)	0,921744	0,922251	0,921744	0,921721
Neural Networks	0,814528	0,829099	0,814528	0,812783

4.2.5. System Architecture, Development and Engineering

Within this study's scope, database administrator, DevOps and system engineer, system analyst, and software architect were grouped under the position group of System Architecture, Development, and Engineering. The evaluation results of the machine learning and artificial intelligence models designed to predict the suitability of individuals for positions in system architecture, development, and engineering are presented in Table 13.

According to these results, the most effective algorithms are logistic regression, support vector machine, and CatBoost. The logistic regression and SVM models outperform other models by achieving F-scores of 0.87 and 0.87, respectively. In addition, the CatBoost algorithm, with an F-score of 0.75, is the third selected algorithm for the ensemble model. Consequently, these three models were selected for the ensemble model designed to predict the job fit of individuals for information systems-related positions in system architecture, development and engineering.

Table 13

Results of System Development and Engineering Models

Algorithm	accuracy	precicsion	recall	F1-score
Random Forest Model	0,730573	0,748898	0,730573	0,725521
CatBoost Model (*)	0,752866	0,754454	0,752866	0,75248
XGBoost	0,729936	0,730476	0,729936	0,729778
KNN Model	0,676433	0,676443	0,676433	0,676428
SVM Model (*)	0,871338	0,879054	0,871338	0,870679
Gradient Boosting Clasifier	0,698726	0,698742	0,698726	0,69872
AdaBoost Clasifier	0,703822	0,723513	0,703822	0,697151
Logistic Regression (*)	0,870064	0,873754	0,870064	0,869742
Neural Networks	0,746878	0,756441	0,746878	0,744845

4.2.6. IS Consultancy, Strategy and Design

Within this study's scope, digital marketing specialists, SAP consultants, solution designers, key account managers, and UI/UX designers were grouped under the IS Strategy and Design positions. The evaluation of the model designed to predict the fit of an individual with specific positions based on their career and professional data is presented in Table 14.

Table 14

Results of IS Strategy and Design Models

Algorithm	accuracy	precicsion	recall	F1-score
Random Forest Model	0,865497	0,873611	0,865497	0,864763
CatBoost Model (*)	0,897661	0,899832	0,897661	0,897522
XGBoost	0,891813	0,892204	0,891813	0,891786
KNN Model	0,849708	0,850594	0,849708	0,849613
SVM Model (*)	0,950877	0,951233	0,950877	0,950868
Gradient Boosting Clasifier	0,870760	0,872784	0,87076	0,870585
AdaBoost Clasifier	0,877778	0,886859	0,877778	0,877056
Logistic Regression (*)	0,942105	0,942280	0,942105	0,942100
Neural Networks	0,892923	0,897643	0,892923	0,892642

According to the results shared in Table 14, the models that most effectively calculate an individual's fit for these job positions are CatBoost, SVM, and Logistic Regression Models. SVM and Logistic Regression models performed successfully by achieving F-scores of 0.95 and 0.94, respectively, while the model developed using the CatBoost algorithm achieved F-score above 0.89. Therefore, these three models were selected to be used in the ensemble model.

4.2.7. Ensemble Learning Results

In this research, the top three performing algorithms in each category are identified to be used in the ensemble learning approach. The identification of these algorithms is based on the evaluation results of each algorithm (accuracy, precision, recall, and F-Score), utilizing the bucket of models approach (Gudivada et al., 2016). Within the ensemble learning approach, for each position category, the three highest-performing models were combined to generate a single, unified output. This collaborative approach led to the final model that surpassed the performance of each individual algorithm included in the ensemble. Since this model is based on the collective decision-making of multiple algorithms, it is inherently more robust against overfitting and has a self-validating structure (Schneider & Xhafa, 2022; Sugiyama, 2016; Yang, 2017). Further details about the ensemble learning approach are provided in the methodology section of this study. Also, the best-performing algorithms for each position are detailed in the previous sections of the findings and listed in Tables 9 through 14.

The evaluation results of the machine learning model obtained by integrating multiple successful algorithms through the ensemble learning approach are presented in Table 15. According to these results, (1) the model predicting individuals' suitability for data analysis and business intelligence positions by calculating their job fit achieved an F-score of 0.917, an accuracy of 0.917, a precision of 0.918, and a recall of 0.917. (2) The model predicting individuals' suitability for software development positions by calculating their job fit achieved an F-score of 0.877, accuracy of 0.877, precision of 0.878, and recall of 0.877. (3) The ensembled model developed to predict the suitability of individuals for product specialist and project management positions by calculating their job fit achieved an F-score of 0.895, an accuracy of 0.895, a precision of 0.897, and a recall of 0.895. (4) The model predicting individuals' suitability for quality assurance and test positions by calculating their job fit achieved an F-score of 0.917, accuracy of

0.917, precision of 0.918, and recall of 0.917. (5) The ensembled model developed to predict the suitability of individuals for system architecture, development, and engineering positions by calculating their job fit achieved an F-score of 0.877, an accuracy of 0.877, a precision of 0.882, and a recall of 0.877. (6) The ensembled model developed to predict the suitability of individuals for IS consultancy, strategy, and design positions by calculating their job fit achieved an F-score of 0.953, an accuracy of 0. 953, a precision of 0. 953, and a recall of 0.953. Finally, the system's overall performance shows the models' average scores. It is seen that the system achieves an average accuracy value of 0.90659, a precision value of 0.90795, a recall value of 0.90659, and an F-score of 0.90649. The accuracy result shows that the model correctly predicts the outcome 90.659% of the time across all categories. The precision result shows that when the model predicts a positive class, it is correct 90.795% of the time. The recall results indicate that the model successfully identifies 90.659% of all actual positive cases. The value of the F-score indicated the model's balance between precision and recall, showing that it effectively integrates both aspects to evaluate the model's accuracy.

Table 15

Results of Ensemble Learning

	accuracy	**precicsion**	**recall**	**F1-score**
Data Analysis and Business Intelligence	0,917610	0,918166	0,917610	0,917583
Software Development	0,877530	0,878836	0,877530	0,877425
Product and Project Management	0,895942	0,897086	0,895942	0,895867
Quality Assurance and Test	0,917563	0,918253	0,917563	0,917529
System Development and System Engineering	0,877707	0,882082	0,877707	0,877356
IS Consultancy, Strategy and Design	0,953216	0,953278	0,953216	0,953215
Overall System Performance	0,90659	0,90795	0,90659	0,90649

Literature suggests that evaluation criteria ranging from 85% to 90% indicate successful solutions to social problems developed using textual data (Kay et al., 2015). Based on these benchmarks, the outcomes of this study, which involved tuning multiple successful machine learning algorithms to create an ensemble model that determines individuals' suitability for specific positions, are considered successful and adequate for addressing a social problem.

4.3. Prototype

In this thesis, within the scope of the deployment step, which is the final stage of the CRISP-DM approach, and in line with the Design Science Research approach, a prototype has been developed to demonstrate the basic functionality of the application. This prototype also incorporates the active use of artificial intelligence models developed using the ensemble learning method.

The prototype application calculates the job-fit (suitability) of a person among the six IT positions within the scope of the research. The system presents these results in a visualized form and then allows the individual to select a target position. Following the selection of the target position, the individual is offered skill development opportunities relevant to that position. The system provides skill suggestions in two different categories personalized for the individual. First, the system lists the skills that are relevant to the chosen position and that the individual already possesses. These skills can be selected to be developed and improved further. The second section lists the skills that are relevant to the chosen position but that the individual does not yet possess. In this section, the individual selects the skills they plan to acquire through lifelong education in the short and medium term. Since the focus is on micro-credentials and skills that can be acquired within the short and medium term through lifelong education, the skills selected for development or acquisition at this stage are limited to five. Consequently, the system recalculates the job-fit score, which indicates an individual's suitability for a particular position based on the user's selection. It then visualizes the job-fit of the user both before and after the proposed skill acquisition. Accompanying this visualization is a brief explanation of the skills that the individual has chosen, providing further insights into their career development process.

Individuals using the system can use it to enhance their suitability for their current position and stay up-to-date in their field or to improve their job fit for a different position, facilitating potential career changes and fostering an adaptable and flexible approach. For these distinct applications of the system, two separate use cases have been developed and are presented in this section. Each use case is illustrated with sequential screenshots from the system.

The first example features an individual employed in the industry as a software test engineer (Figures 22-30). In this use case, the individual utilizes the system to enhance their job fit within the Quality Assurance and Test field, which encompasses software test engineering. The process is detailed as follows: The individual first reviews how well their competencies match the positions within the application (Figures 23 and 24), selects the career path they wish to advance in (Figure 25), chooses from the skills recommended by the system to improve their proficiency in this position (Figures 26, 27, 28), views an informational screen about the selected skills (Figure 29), and finally, sees the potential increase in their suitability for the position upon acquiring these competencies (Figure 30).

Figure 22

Prototype Example 1- Profile Selection Page

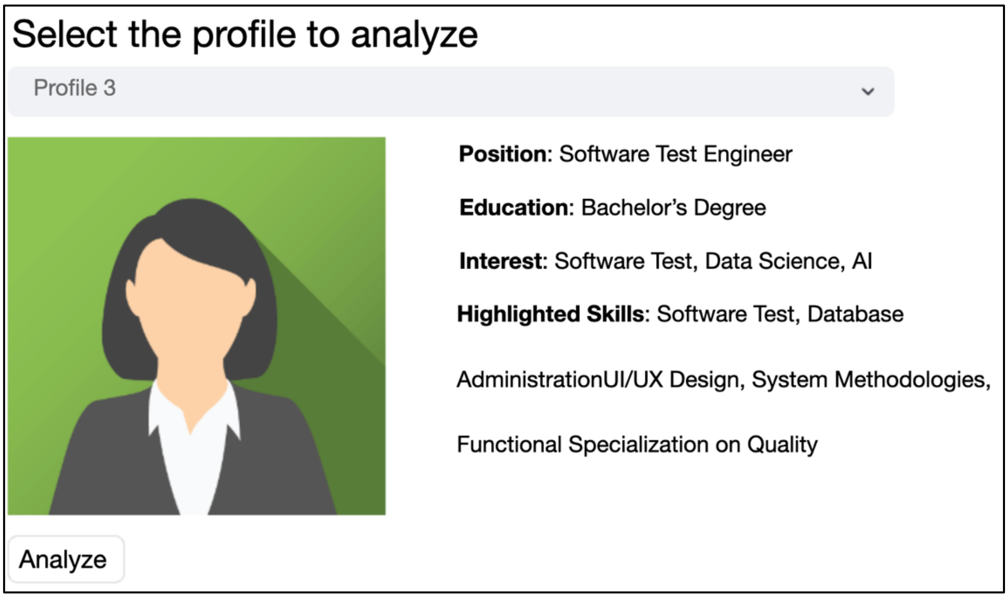


Figure 23

Prototype Example 1- Result Screen

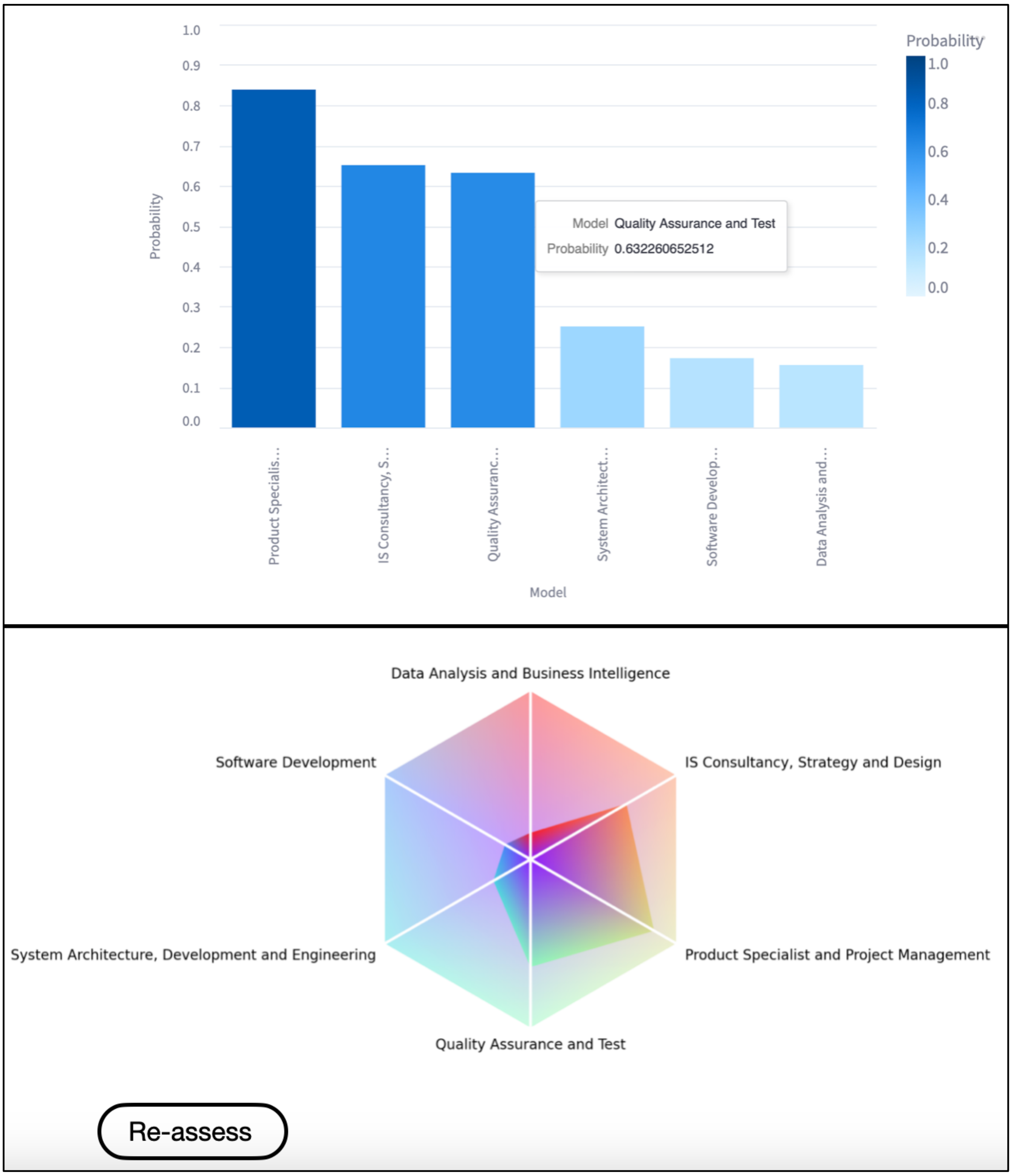


Figure 24

Prototype Example 1- Selection of Targeted Job Position

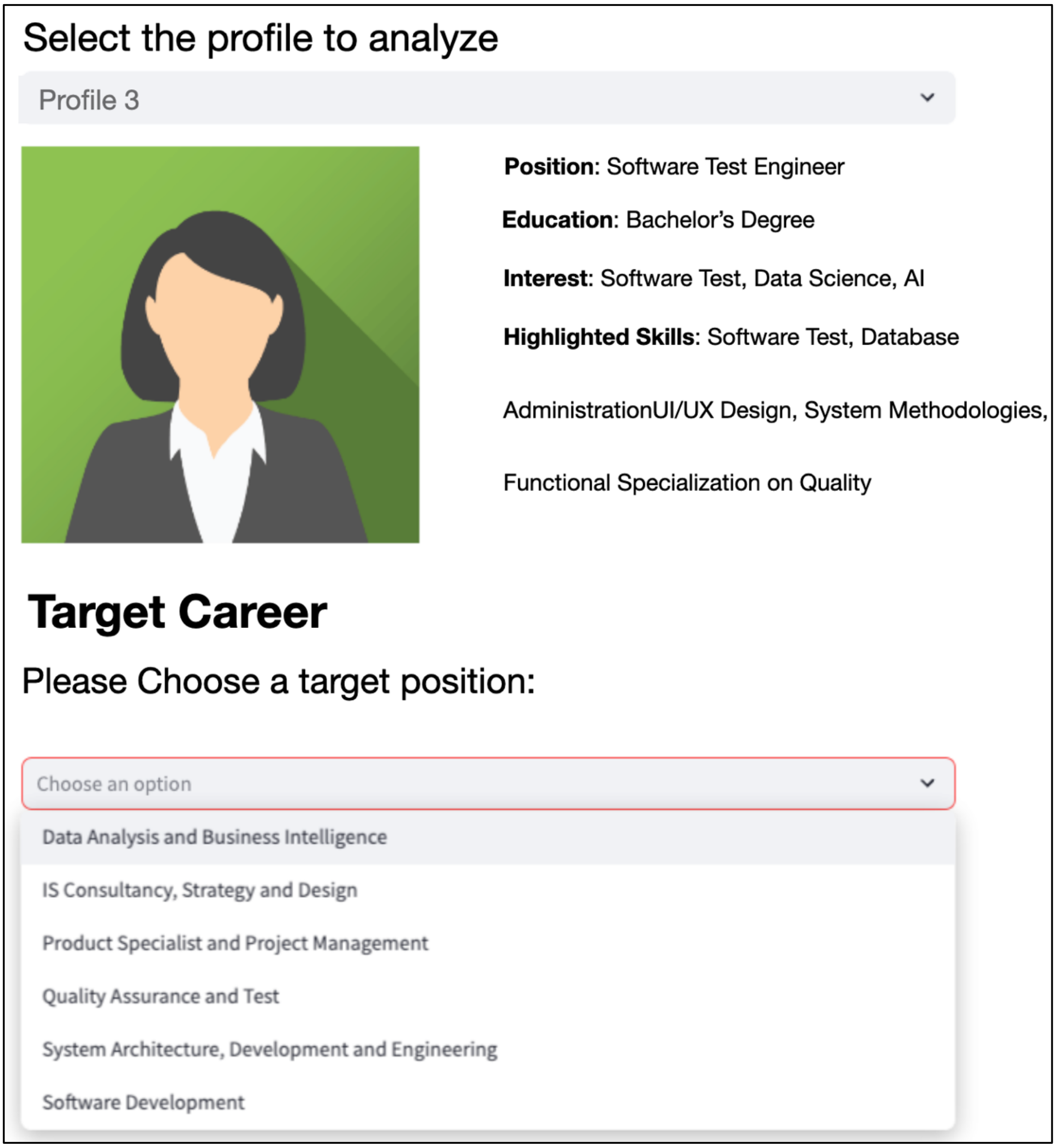


Figure 25

Prototype Example 1- Screen for Selection of Skills to Increase Job Fit

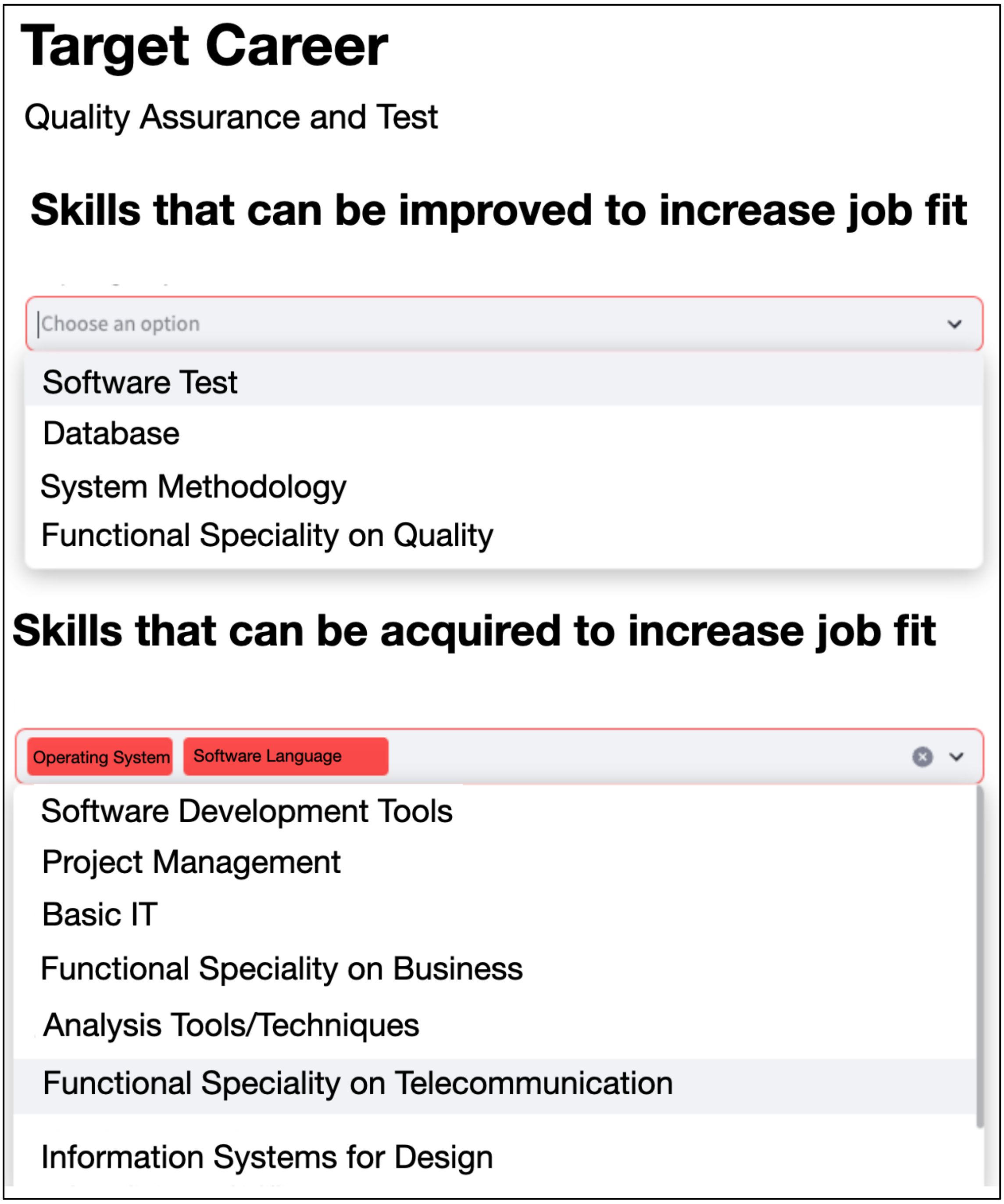


Figure 26

Prototype Example 1- Brief Overview of Selected Skill

Short Description of Selected Skills.

Functional Speciality of Quality

Knowledge and experience in quality management and improvement processes.

Key Terms: ISO Standards, Quality Control, Process Optimization

Software Language

Software language that are used for software development

Key Terms: Python, Java, C#, JavaScript, Angular, PhP, Reach, HTML, Swift

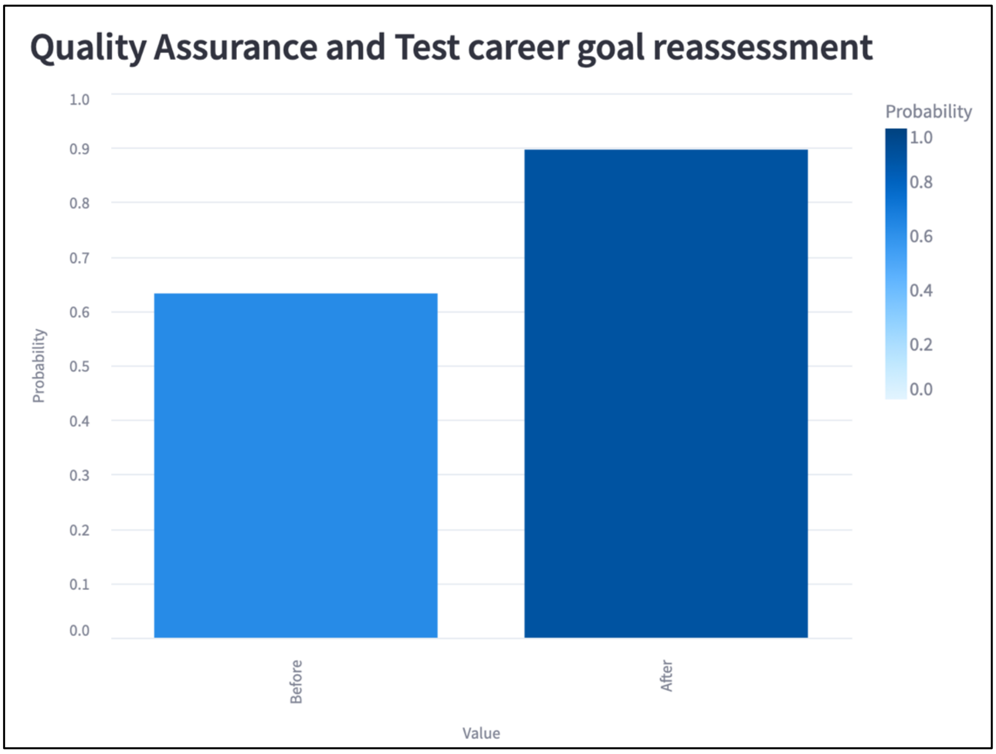
Analysis Tools and Techniques

Methods and tools used for data analysis.

Key Terms: SPSS, R, Tableau, Power BI, QlikView, OLAP

Figure 27

Prototype Example 1- Results of Re-Evaluation



The second example involves a user who is looking to transition between positions in information systems (Figures 31-36). This user employs the system to get help for his shift from a software development role to a position in system development and engineering. Through the system, the user can see how the specific skills they choose to develop and acquire could enhance their suitability for the new position.

In this use case, the individual utilizes the system to enhance their job-fit within the system development and engineering. The individual uses the system to analyze their profile and review how well their competencies match the positions within the application (Figures 28 and 29). The user then selects the career path they wish to advance in and chooses from the skills recommended by the system to improve their proficiency in this position (Figures 30). Lastly, the user views the potential increase in their suitability for the position upon acquiring these competencies (Figure 31) and an informational screen about the selected skills (Figure 32).

Figure 28

Prototype Example 2- Profile Selection Page

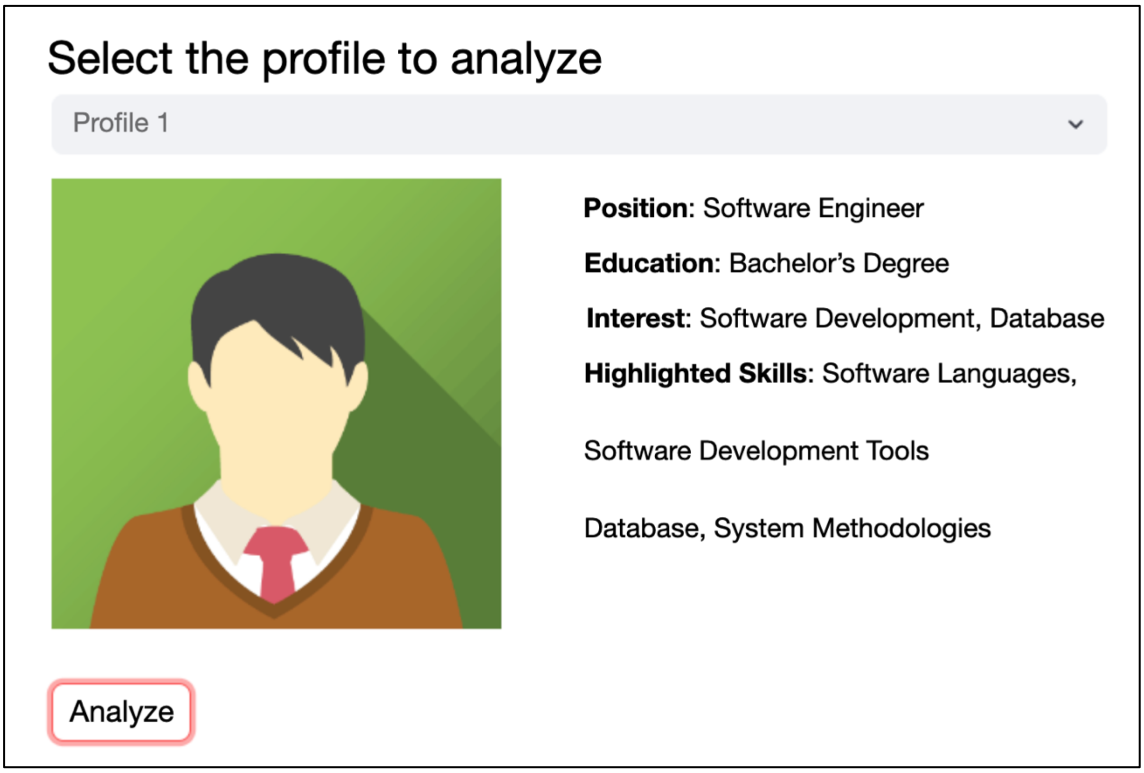


Figure 29

Prototype Example 2 – Results Screen

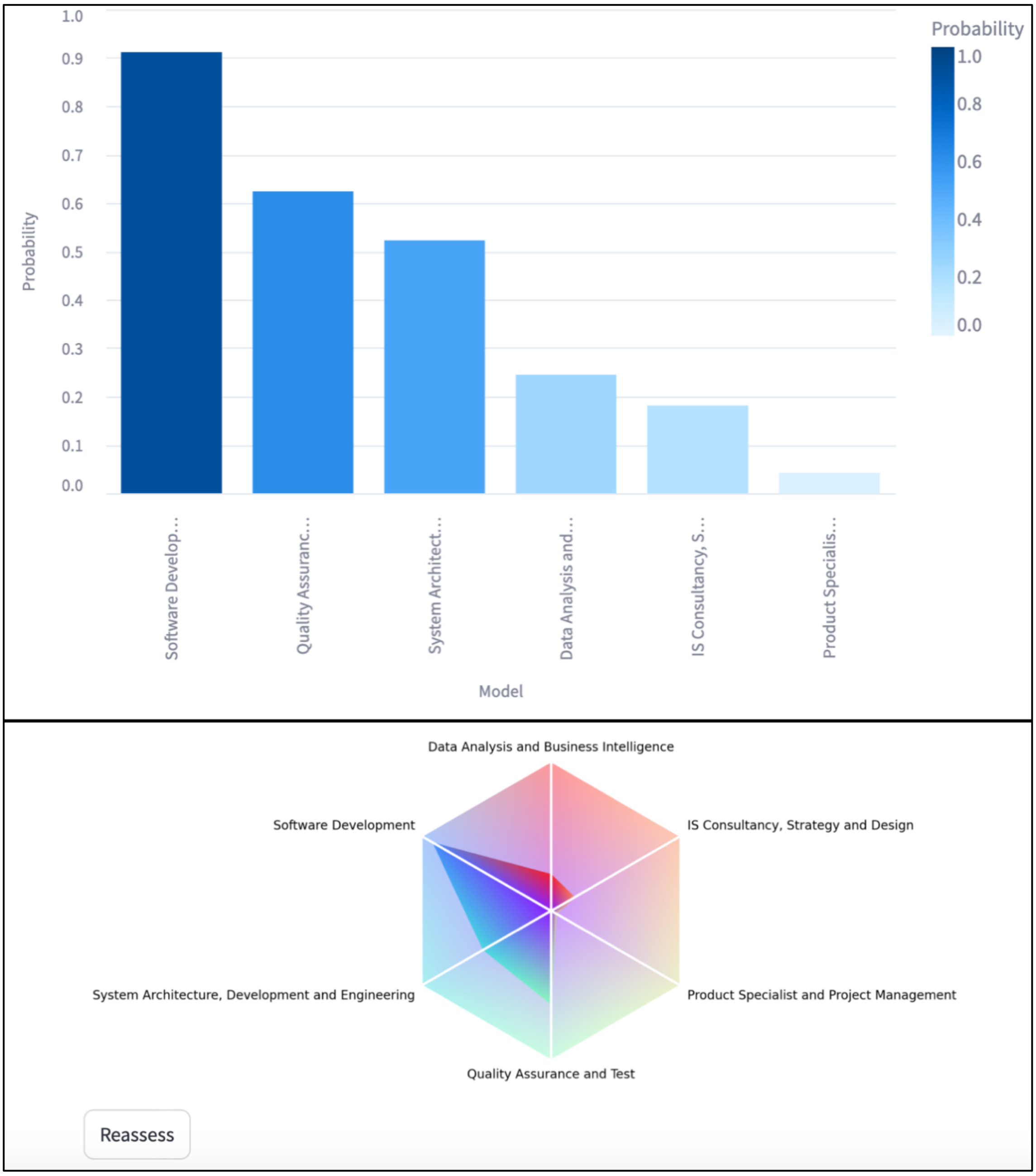


Figure 30

Prototype Example 2 - Selection of Targeted Job Position and Skills

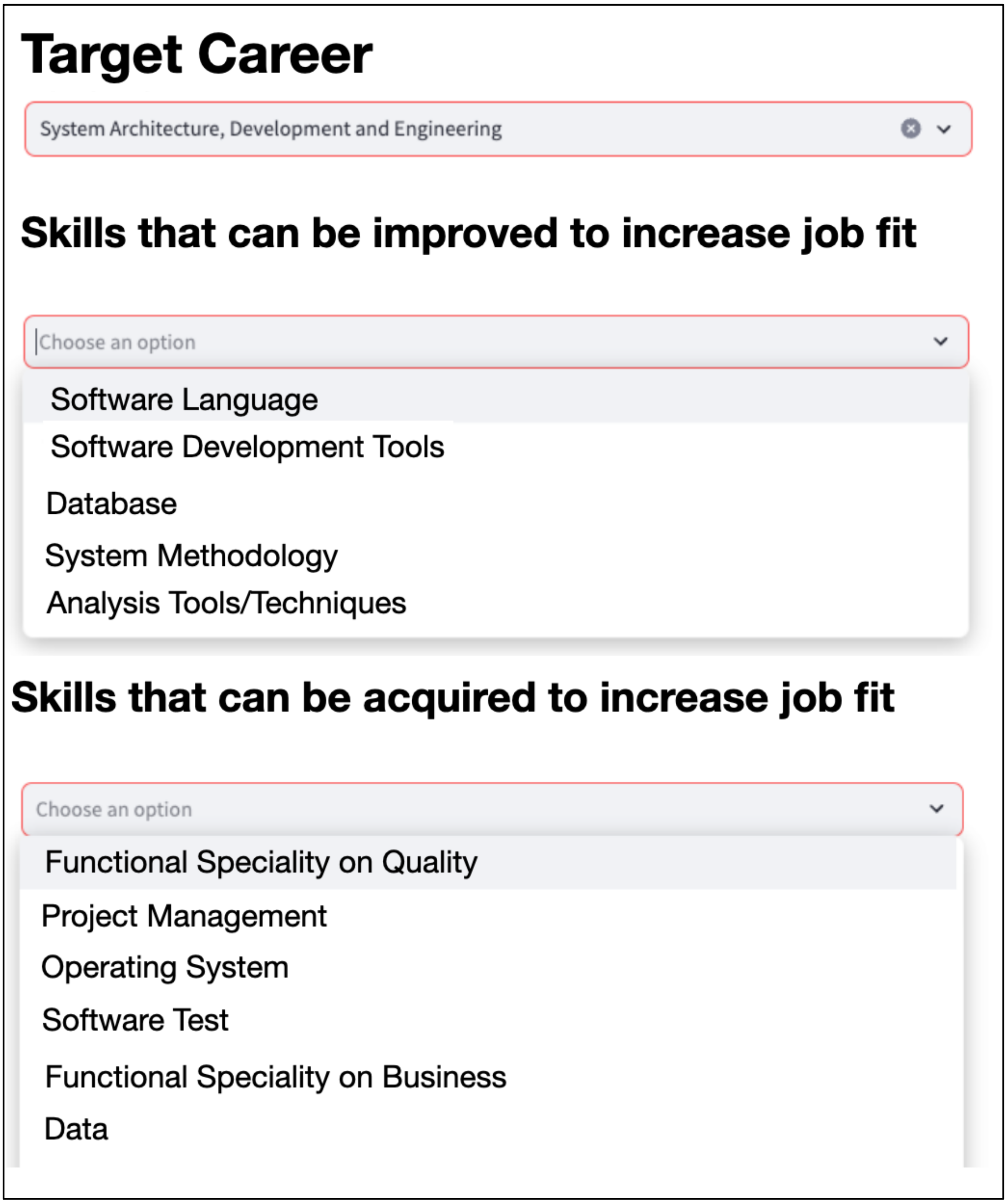


Figure 31

Prototype Example 2 - Results of Re-Evaluation

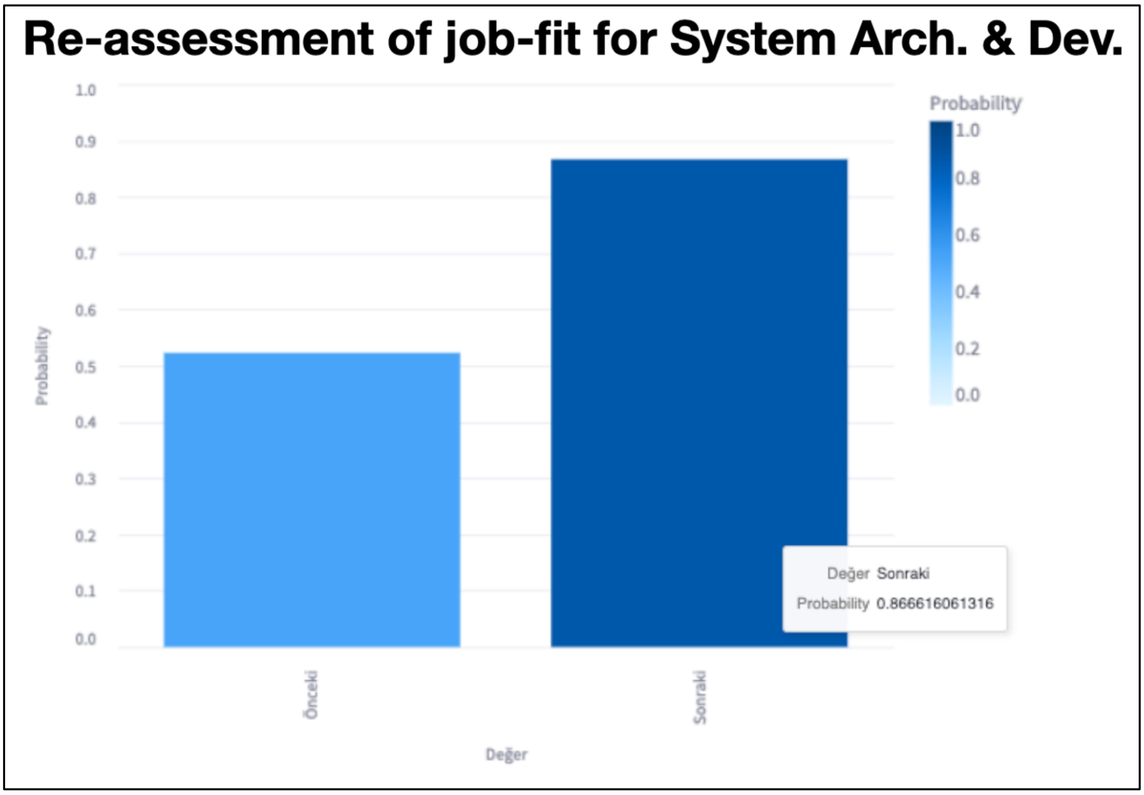


Figure 32

Prototype Example 2 - Brief Overview of Selected Skill

In conclusion, this thesis has successfully demonstrated the practical application of a career planning system that integrates the principles of the CRISP-DM approach and

Design Science Research methodology. By developing a prototype that utilizes artificial intelligence models, specifically through ensemble learning, this work provides a robust tool for individuals to evaluate and enhance their career trajectories. The system's capacity to personalize skill development recommendations based on the individual's current competencies and desired career path illustrates its innovative approach to career advancement. This tool not only facilitates immediate career development strategies but also encourages lifelong learning and adaptability in the rapidly evolving IT sector.

DISCUSSION

This thesis aims to resolve the research question of "How does an artificial intelligence-based career planning system, utilizing ontology, data science, and machine learning techniques, affect the career development and job alignment of information technology professionals, career changers, and NEET individuals?". To achieve the research objectives, firstly, an ontology model was developed to verify that PSMP data aligns with the research purposes, Professionally Oriented Social Network Platform (PSMP) profile data was collected using the data collection tool developed, and secondly, machine learning models were developed to calculate the job-fit score, which indicated individuals' suitability for a position, and finally, a prototype of the AI-based career planning system was designed and developed.

An ontology-driven conceptual model was developed using UFO foundational ontology and OntoUML modeling language to effectively integrate the data into the system to be developed (Figure 21). The findings from the ontological model suggest that position types can be identified through the positions held by current or former employees, and types of competencies can be determined from individuals' skills and competencies. Building on this, higher position types can be identified through their manifestation in specific positions and their relationship with competence types. Consequently, it has been shown that the objectives of this thesis can be achieved using the insights obtained from the ontological model.

Within the scope of the investigation of the utility of PSMPs and the suitability of professionals' profile data on these platforms to support the career development process. The findings of the studies conducted by the researcher and thesis advisor during the thesis indicate that PSMP platforms are positioned differently compared to other social media platforms by professionals, and individuals' participation is goal-oriented (Aktaş & Akbıyık, 2019). Also, PSMP users view their profiles on these platforms as digital CVs (Aktaş & Akbıyık, 2019; Mashayekhi & Head, 2018, 2022). Based on these findings, within the scope of this thesis, the machine-learning models were created using PSMP profile data collected with the data collection tool developed by the researcher.

Additionally, ensemble models were developed composed of three different machine-learning algorithms that were identified as the most successful out of the ten evaluated algorithms. These models were evaluated based on their performance metrics. The models achieved average accuracy, precision, recall, and F-score exceeding 0.90. This level of accuracy underscores the robustness and reliability of the AI-based career planning system in providing accurate job-fit calculations (Kay et al., 2015).

Another significant finding and the outcome of this thesis is the development of the prototype. A prototype of an AI-based career planning system, which is intended for use by IT professionals, aspiring employees seeking IT positions, and NEETs (Not in Education, Employment, or Training), was designed and developed. The system, developed using a design science research approach (Hevner et al., 2004), ontology-driven conceptual modeling (Guizzardi, 2005), and machine learning methods, calculates a job-fit score that indicates individuals' suitability for positions. Then, it provides personalized career development suggestions aimed at skill enhancement within the framework of the life-span, life-space approach and lifelong learning (Super, 1980), and facilitates self-assessment within the framework of individual career planning (Anafarta, 2001; Jaffe & Scott, 1991), thus guiding individuals in their personal development.

Theoretical Alignment

The findings of this study provide significant insights into the utility of professionally oriented social media platform data and the application of artificial intelligence and machine learning in career planning. To contextualize these results, the implications and relevance of the research findings were discussed by drawing on established career theories. This will help validate the study's outcomes and demonstrate the alignment of the study's results with the literature on career development and lifelong learning. The study's theoretical contributions will be highlighted by analyzing the results through the lenses of individual career planning theories, boundaryless career theory, career construction theory, and lifelong learning theories.

Individual Career Planning Theories

The individual career planning process consists of self-assessment, exploring possibilities, creating a plan, taking action, evaluating action, and receiving feedback (Anafarta, 2001; Jaffe & Scott, 1991). The proposed approach and the developed system in this study were created to support and follow the steps of individual career planning, and they align with individual career planning processes as follows.

Self-assessment: An ontology-driven, AI-based career planning system helps users calculate their job-fit score that will indicate their suitability for IT positions.

Explore Possibilities: The system will visualize individuals' suitability for multiple IT positions, representing the results in the bar chart and hexagonal plot format. This will enable users to compare and explore their job fit for various positions.

Create a Plan: Once the user selects a position, the system will present skill recommendations for the specific position so that the user can create a plan to improve oneself in the scope of lifelong learning.

Take Action: The system encourages the user to take action by enabling them to select skills to be improved or to be acquired among the ones that are recommended.

Evaluate Outcome: Lastly, the system will calculate the expected job fit after the user selects the skills to be improved or to be acquired. Thus, users can simulate the process and be informed about the potential outcome.

Boundaryless Career Approach

The boundaryless career is a modern approach that advocates for individuals' career paths not being limited to a single employer or a company. It emphasized that individuals should have control over their own career planning process (Arthur, 1994). This approach promotes more flexible and adaptable career paths. It challenges rigid assumptions about traditional career paths and argues that such constraints hinder both organizational and individual development (Arthur, 1994; Forret & Sullivan, 2002).

The AI-based career planning system aims to support individuals' flexibility, mobility, and adaptability by allowing them to assess their suitability for desired IT positions. It provides suggestions to enhance this suitability and promotes lifelong learning. In addition, the system utilizes the cumulative data of the labor market rather than the data of a single company. Using PSMP data modeled with ontology-driven conceptual modeling helps individuals break free from rigid patterns and assumptions within organizations. In summary, the system aims to empower users to take control of their career paths within the boundaryless career landscape by providing the necessary tools.

Career Construct Theory (CTT)

Career construct theory describes career construction as a dynamic process in which individuals shape their careers by ascribing personal meaning to their experiences and future plans (Savickas et al., 2009; D. Wang & Li, 2024). Three main components of this theory are vocational personality, career adaptability, and life themes. Vocational personality refers to individuals' skills, needs, values, and interests related to their careers. Career adaptability refers to the ability to develop necessary strategies to cope with the career transitions and career development process. Life themes, on the other hand, represent the personal meaning and direction individuals attribute to their careers (Savickas, 2013; Savickas et al., 2009).

This thesis aligns with career construct theory in the following ways. By using ontologically modeled PSMP data, the system offers a data-driven approach to the acquisition of skills and competencies, enabling individuals to understand and develop their vocational personalities. The system enhances career adaptability by enabling individuals to perform self-assessment through the calculation of job-fit scores. By providing providing personalized recommendations, it helps individuals successfully navigate career transitions and challanges during career development. Consequently, it emphasizes continuous and adaptive growth in career development.

Lifelong Learning Theories

Individuals need to continually enhance and develop their skills throughout their adult lives to meet the demands of modern life and handle its challenges. Lifelong learning is essential to achieve this continuous development. Lifelong learning is engaging in educational activities at all stages of life. It is a life-wide approach where learning is incorporated into all phases and aspects of life, from home to school to the workplace and beyond (Laal, 2011). Lifelong learning refers to the activities individuals engage in their lives to develop their knowledge, skills and competencies in a particular area. These activities are motivated by personal, social, or professional reasons (Laal, 2011). Several theories and approaches explain the role of lifelong learning for individuals, especially adults. In the scope of this discussion, two main approaches will be examined. The first one is Jarvis' adult education and lifelong learning theory, and the latter is Knowles' adult learning theory.

Jarvis' theory of adult education and lifelong learning presents a comprehensive framework that emphasizes the importance of continuous learning throughout life. According to this framework, learning activity occurs in a variety of contexts, such as personal, social, and Professional, and it spans the life of the individual. This theory highlights the significance of learning events and activities outside formal education. It focuses on non-formal and informal learning approaches. Non-formal learning encompasses educational activities that occur outside of traditional educational institutions, while informal learning refers to learning that happens through work, family interactions, or daily activities (Jarvis, 2006).

The AI-based career planning system offers skills suggestions following personal assessments, encouraging individuals to develop their skills. With these suggestions, people have the opportunity to develop themselves outside of formal education. The system encourages the acquisition of skills and micro-credentials. This acquisition process may take place through online courses, certifications, or practical experiences. This helps individuals enhance their capabilities through various learning opportunities outside traditional educational paths, aligning with non-formal and informal learning approaches.

Knowles's adult learning theory, also known as Andragogy, focuses on the ways adults learn and how it differentiates from the learning process of children and younger individuals. Knowles's theory outlines the distinct characteristics of the adult learning process and seeks to identify the most effective adult learning styles. Accordingly, the theory makes six assumptions about the adult learning process; (1) adults want to understand the reason behind their learning, (2) adults benefit from building on their existing experiences, (3) adults should feel responsible for their own learning, (4) adults are motivated to learn when they need to solve real-life problems, (5) adults prefer learning to be problem-oriented, (6) adults learn best when they are intrinsically motivated (Knowles et al., 2020).

The AI-based career planning system aligns with Knowles' Andragogy approach by empowering individuals to take control of their career planning and talent development process. It also supports self-directed learning, where individuals determine what they need to learn and understand the reasons behind it. Adopting a problem-oriented approach, the system provides individuals with personalized skill recommendations for the position they have identified, aiming to increase their suitability for these roles. Additionally, the system presents visualized their results results compared to the realities of the labor market, aiming to foster intrinsic motivation for lifelong learning through these personalized suggestions.

Practical Implications

This study offers significant practical implications for both individuals and organizations. Additionally, the integration of ontology-driven conceptual modeling aims to achieve these contributions in a scalable, interoperable, and scalable manner.

Practical Implications for Individuals

The system aims to provide an up-to-date and consistent career assessment based on the competencies of professionals who are already working in selected IT positions. This approach ensured that the assessment accurately reflects the current state of the labor market. The system calculates individuals' suitability for a position and provides a personalized list of skills that would increase their suitability for the position they are interested in. The system also allows individuals to foresee the potential impacts of the skills that they selected among the list of items. This personalized approach aims to create a career development process that individuals can internalize, thus enabling intrinsic motivation. Unlike traditional methods, the ability to deliver evaluations and feedback quickly and in real time makes this service more accessible to a broader audience.

The system's data-driven decision-making approach utilized a novel dataset. It ensured that assessments and the results were based on the current labor market trends and requirements. This enables the system to guide and assist users in career transition processes such as job changes, sector changes, or transition to different positions. It enables individuals to plan effectively by allowing them to see the impact of potential skill acquisition on their position fit, thus encouraging lifelong learning. Through lifelong learning, individuals can continuously improve themselves in line with the demands of the changing labor market and increase their employability and adaptability by taking control of their own career development.

Practical Implications for Organizations

The system also offers significant practical implications for organizations. By integrating the system into their employees' career development process, organizations can make career development more accessible to a larger number of employees. By providing instant feedback and suggestions to users, the system is more time-effective and faster for the organizations compared to traditional approaches. Thus, the organizations can ensure that their workforce remains competitive in the international job market. With a system that supports continuous learning and adaptability, organizations can integrate continuous learning activities into their career development processes, and they can enable their employees to move between units and positions, achieving the flexibility needed to adapt to the rapidly changing industry.

Practical Implications of Ontological Foundations

The use of a foundational ontology, Unified Foundational Ontology (UFO), in the case of this study, ensures that the system is ontologically well-founded, supports interoperability, and facilitates integration. The ontological model represents the domain of professionally oriented social media platform data, with a particular emphasis on career-related information. With ontological rigor provided by the ontology-driven conceptual model, a common understanding can be established in the process of upkeeping, updating, and extending the system, thereby future-proofing it. This ensures that the system keeps up with the evolving labor market requirements and, if needed, can be integrated into other working systems with the help of ontological understanding.

Limitations

This study was conducted with several limitations and within a certain scope. Firstly, the study focused on the analysis and modeling of positions related to information systems (IS) and information technologies (IT). IS and IT are fields primarily affected by rapidly changing job requirements. These sectors are also where micro-credentials acquired through boot camps, certifications, and other lifelong learning activities are already used in the recruitment process. This scope limitation is partly due to the high PSMP adaptation rate among IS and IS workers (J. Davis et al., 2020). By focusing on these sectors, the study can achieve more comprehensive access to the labor market during the data collection process. Thus, this research and the prototype developed are limited to specific job positions within the field of information technologies and systems.

Another limitation of the research is related to the geographical diversity of the data. Since the majority of the companies with Great Place to Work employee satisfaction awards and certificates in Turkey are located in Istanbul, the data is geographically concentrated in this city. Consequently, it was not possible to use geographical location as a criteria in this study.

CONCLUSION

This thesis aims to address the challenges of individual career development for information systems professionals in the rapidly changing job market through the AI-based career planning system developed within the scope of this research. In this context, the research sought the answer to whether an AI-based system can be implemented in the career planning process for the career development of information systems professionals. To address this, three main artifacts were developed within the framework of design science research, providing resolution to the research problems and sub-problems.

Firstly, to answer the sub-problem of “Is ontology-based conceptual modeling an effective method for ensuring the scalability, adaptability, and interoperability of an artificial intelligence-based career planning system?”, the research domain was modeled using an ontology-driven conceptual modeling approach, employing the Unified Foundational Ontology and the OntoUML ontological modeling language. This ensured that the connection between the research domain, the dataset, the research problems, and the system developed within the scope of the research with ontological rigor. This approach increases the interoperability of the thesis study's results, facilitates the integration of the system with different systems, and guides the possible integration processes. It also laid the foundation for the system's internationalization, enhanced the system's understandability, and supported its further development (Gemino & Wand, 2005; Guizzardi & Proper, 2021; Olivé, 2007).

Secondly to answer the research sub-problem of “Are data science methods effective in accurately determining job alignment for specific information technology positions based on an individual's current skills, experiences, education, and qualifications?”, artificial intelligence models were developed to calculate the job-fit score, which indicates the suitability of individuals for the IT positions included in the research. Using an ensemble learning approach, the three most successful models for each position were selected from ten different machine learning and artificial intelligence algorithms. These models were combined to create an ensemble model that makes decisions collectively. The study resulted in successful models that performed above 0.90 in average accuracy, recall, precision, and F-score criteria.

Finally, to answer the research sub-problems of “How effectively can an artificial intelligence-based system provide personalized, accurate recommendations on the skills and competencies individuals need to acquire or improve to align with targeted information technology positions?" and "How effectively does the prototype of an ontology-driven, artificial intelligence-based career planning system demonstrate its feasibility and functionality?" the third artifact of the design science research process, which is a prototype of the system, is designed and developed. This prototype leverages the developed artificial intelligence models to calculate an individual's job-fit for various positions based on their professional information. It also recommends skills for the position chosen by the user and calculates a new job-fit value that updates when the selected skills are improved or acquired. This system is presented in the findings section of the thesis, with screenshots illustrating two different scenarios. The first scenario involves an individual aiming to improve themselves in their current position and stay up-to-date with the job market, while the second scenario involves an individual seeking to increase their suitability for a different position. These three artifacts align with the design science research process and the research goals and objectives.

The practical implications of the research are considered from individual, organizational, and ontological points of view. Individuals can use the system to learn about the potential impact of their chosen skills, which increases intrinsic motivation and helps them to internalize the career development process. The system guides the individual career development process with a data-driven approach based on current labor market trends and encourages continuous learning. Unlike traditional methods, it offers fast and real-time assessment and feedback, enabling the system to reach a wider audience. Career development can be made more accessible within organizations, and continuous learning activities can be integrated into the company's training processes, helping employees adapt to the changing and evolving industry. The use of Unified Foundational Ontology ensures ontologically sound grounding of the system, interoperability, and easy integration with other systems. This ensures that the system can be updated and extended, making it future-proof.

The study alignings with major career and lifelong learning theories. The study results and outputs support the individual career development process (Anafarta, 2001; Jaffe & Scott, 1991) by allowing users to perform the steps of assessment, exploration of possibilities, planning, action, and outcome evaluation. Additionally, the outputs of the study support the flexibility and adaptability of users within the scope of boundaryless career theory (Arthur, 1994; Guan et al., 2019), enabling career structuring beyond traditional approaches. Career construct theory (Savickas, 2013; D. Wang & Li, 2024) is addressed by enhancing vocational personality understanding, career adaptability, and the personal meaning of careers through personalized, data-driven insights. Lifelong learning theories, particularly those of Jarvis (Jarvis, 2006) and Knowles (Knowles et al., 2020), are supported by encouraging continuous skill development through non-formal and informal learning opportunities and fostering intrinsic motivation and problem-oriented learning in adults. Overall, the system empowers individuals to take control of their career development in line with contemporary career and learning theories.

REFERENCES

Adekola, B. (2011). Career management as correlates for career development and job satisfaction a case study of Nigerian bank employees. *Australian Journal of Business and Management*, *1*(2), 100–112. http://www.ajbmr.com/articlepdf/ajbmr_v01n02_07.pdf

Aksu, Ç., & Yıldızçakar Sarıoğlu, C. (2022). Makine öğrenmesi temelli yöntemler - olasılık ve uzaklığa dayalı teknikler. In A. Hızıroğlu, H. İ. Cebeci, & K. Seçkin Codal (Eds.), *Veri modelleme ve veri analitiği: Sağlık ve işletme uygulamaları* (pp. 258–283). Nobel Akademik Yayıncılık.

Aktaş, B., & Akbıyık, A. (2019). Social connectedness in professionally-oriented social network sites and its effect on social network value. *6. International Management Information Systems Conference*, *October 2020*.

Aktaş, B., & Akbıyık, A. (2023, October 18). The potential of the use of ontology-driven conceptual modeling and OntoUML in management information systems studies. *10th International Management Information Systems Conference*. https://2023.imisc.net/tr/imisc23_bildiri_ozet_kitapcigi.pdf

Aktaş, B., Baz Aktaş, N., & Akbıyık, A. (2022). İşgücü piyasalarında yönetim bilişim sistemleri programlarının farkındalığı: ABD ve Türkiye'deki iş ilanları üzerinden bir değerlendirme. *Journal of Research in Business*, *7*(1), 60–79. https://doi.org/10.29228/JRB.1024997

Aliaga, O. A. (2001). Human capital, HRD and the knowledge organization. In O. A. Aliaga (Ed.), *Academy of Human Resource Development 2001 Conference*.

Altan, Z. (2022). An ontology model oriented to the career planning of university students. *HORA 2022 - 4th International Congress on Human-Computer Interaction, Optimization and Robotic Applications, Proceedings*. https://doi.org/10.1109/HORA55278.2022.9799963

Anafarta, N. (2001). Orta düzey yöneticilerin kariyer planlamasına bireysel persfektif. *Akdeniz İ.İ.B.F. Dergisi*, *2*, 1–17.

Arabameri, A., Pradhan, B., & Rezaei, K. (2019). Gully erosion zonation mapping using integrated geographically weighted regression with certainty factor and random forest models in GIS. *Journal of Environmental Management*, *232*, 928–942. https://doi.org/10.1016/J.JENVMAN.2018.11.110

Ardıç, K., & Özdemir, Y. (2017). *İnsan kaynakları yönetimi (Ders Notları)*. Sakarya Yayıncılık.

Arthur, M. B. (1994). The boundaryless career: A new perspective for organizational inquiry. *Journal of Organizational Behavior*, *15*(4), 295–306. https://doi.org/10.1002/job.4030150402

Arthur, M. B. (2008). Examining contemporary careers: A call for interdisciplinary inquiry. *Https://Doi.Org/10.1177/0018726707087783*, *61*(2), 163–186. https://doi.org/10.1177/0018726707087783

Aytaç, S. (2006). *Yorumlar 2 paylaş çalışma yaşamında kariyer yönetimi planlaması gelişimi ve sorunları*. Ezgi Kitabevi.

Azevedo, C. L. B., Iacob, M. E., Almeida, J. P. A., Van Sinderen, M., Pires, L. F., & Guizzardi, G. (2015). Modeling resources and capabilities in enterprise architecture: A well-founded ontology-based proposal for ArchiMate. *Information Systems*, *54*, 235–262. https://doi.org/10.1016/J.IS.2015.04.008

Barber, A. E., & Roehling, M. V. (1993). Job postings and the decision to interview: A verbal protocol analysis. *Journal of Applied Psychology*, *78*(5), 845–856. https://doi.org/10.1037/0021-9010.78.5.845

Bartosik, A., & Whittingham, H. (2021). Evaluating safety and toxicity. *The Era of Artificial Intelligence, Machine Learning, and Data Science in the Pharmaceutical Industry*, 119–137. https://doi.org/10.1016/B978-0-12-820045-2.00008-8

Behrendt, P., Göritz, A. S., & Heuer, K. (2019). Career counseling process quality promotes reemployment. *Https://Doi.Org/10.1177/0894845319884731*, *48*(5), 599–606. https://doi.org/10.1177/0894845319884731

Belgiu, M., & Drăgu, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, *114*, 24–31. https://doi.org/10.1016/j.isprsjprs.2016.01.011

Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(8), 1798–1828. https://doi.org/10.1109/TPAMI.2013.50

Bergenhenegouwen, G. J. (1996). Competence development - a challenge for HRM professionals: Core competences of organizations as guidelines for the development of employees. *Journal of European Industrial Training*, *20*(9), 29–35. https://doi.org/10.1108/03090599610150282

Bissacco, A., Yang, M. H., & Soatto, S. (2007). Fast human pose estimation using appearance and motion via multi-dimensional boosting regression. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 4270154. https://doi.org/10.1109/CVPR.2007.383129

Boland Jr., R. J. (2022). 12 Design in the punctuation of management action. *Managing as Designing*, 106–112. https://doi.org/10.1515/9780804767439-014/HTML

Boudi, Z., Ait Wakrime, A., Toub, M., & Haloua, M. (2020). Building valid career ontologies with B-CPNs. *Communications in Computer and Information Science*, *1207 CCIS*, 33–46. https://doi.org/10.1007/978-3-030-45183-7_3

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324/METRICS

Brott, P. E. (2001). The storied approach: A postmodern perspective for career counseling. *The Career Development Quarterly*, *49*(4), 304–313. https://doi.org/10.1002/J.2161-0045.2001.TB00958.X

Caers, R., & Castelyns, V. (2011). Linkedin and facebook in Belgium: The influences and biases of social network sites in recruitment and selection procedures. In *Social Science Computer Review* (Vol. 29, Issue 4, pp. 437–448). https://doi.org/10.1177/0894439310386567

Caie, P. D., Dimitriou, N., & Arandjelović, O. (2020). Precision medicine in digital pathology via image analysis and machine learning. *Artificial Intelligence and Deep Learning in Pathology*, 149–173. https://doi.org/10.1016/B978-0-323-67538-3.00008-7

Calhau, R. F., & Almeida, J. P. A. (2022). Zooming in on competences in ontology-based enterprise architecture modeling. *International Conference on Enterprise Design, Operations, and Computing* , *466 LNBIP*, 198–213. https://doi.org/10.1007/978-3-031-26886-1_12

Calhau, R. F., Azevedo, C. L. B., & Almeida, J. P. A. (2021). Towards ontology-based competence modeling in enterprise architecture. *Proceedings - 2021 IEEE 25th International Enterprise Distributed Object Computing Conference, EDOC 2021*, 71–81. https://doi.org/10.1109/EDOC52215.2021.00018

Cebeci, H. İ. (2022). Metin madenciliği ve analitiği. In A. Hızıroğlu, H. İ. Cebeci, & K. Seçkin Codal (Eds.), *Veri modelleme ve veri analitiği: Sağlık ve işletme uygulamaları* (pp. 460–494). Nobel Akademik Yayıncılık.

Cebeci, H. İ., Güner, S., Arslan, Y., & Aydemir, E. (2023). Barriers and drivers for biking: What can policymakers learn from social media analytics? *Journal of Transport & Health*, *28*, 101542. https://doi.org/10.1016/J.JTH.2022.101542

Chakraborty, D., Ghosh, A., & Saha, S. (2019). A survey on internet-of-thing applications using electroencephalogram. *Emergence of Pharmaceutical Industry Growth with Industrial IoT Approach*, 21–47. https://doi.org/10.1016/B978-0-12-819593-2.00002-9

Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Rudiger, W. (2000). Crisp-Dm 1.0. *CRISP-DM Consortium*, 76.

Corcho, O., Fernández-López, M., & Gómez-Pérez, A. (2006). Ontological engineering: Principles, methods, tools and languages. In C. Calero, F. Ruiz, & M. Piattini (Eds.), *Ontologies for Software Engineering and Software Technology*. Springer.

Cumbler, E., Yirdaw, E., Kneeland, P., Pierce, R., Rendon, P., Herzke, C., & Jones, C. D. (2018). What is career success for academic hospitalists? A qualitative analysis of early-career faculty perspectives. *Journal of Hospital Medicine*, *13*(6), 372–377. https://doi.org/10.12788/JHM.2924

Cumhurbaşkanlığı Strateji ve Bütçe Başkanlığı, T. (n.d.). *On birinci kalkınma planı 2019-2023*.

Çarkıt, E. (2019). Kariyer danışmanlığı ve mesleki rehberlik alanında türkiye'de yapılan lisansüstü tezlerin araştırma eğilimleri. *MANAS Sosyal Araştırmalar Dergisi*, *8*(2). https://www.academia.edu/download/55986276/Manas_Sosyal_Arastirmalar_Dergisi_31._Sayi.pdf

Davis, C., Baty, B. J., Hippman, C., Trepanier, A., & Erby, L. (2020). Genetic counselors with advanced skills: II. A new career trajectory framework. *Journal of Genetic Counseling*, *29*(5), 771–785. https://doi.org/10.1002/JGC4.1204

Davis, J., Wolff, H. G., Forret, M. L., & Sullivan, S. E. (2020). Networking via LinkedIn: An examination of usage and career benefits. *Journal of Vocational Behavior*, *118*, 103396. https://doi.org/10.1016/J.JVB.2020.103396

Dini, S. S., Sharwin, X. R., Sherin, S. J., & Shobana, G. (2022). Career guidance application: A digital approach to career planning for secondary level students. *I-Manager's Journal on School Educational Technology*, *18*(2), 29. https://doi.org/10.26634/JSCH.18.2.19253

Dneprovskaya, N., Chris Kang, S.-B., & Shevtsova, I. (2022). Evolution of the competencies to embrace digital technology for sustainable development. *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3491101.3519730

Dorogush, A. V., Ershov, V., & Yandex, A. G. (2018). *CatBoost: gradient boosting with categorical features support*. https://arxiv.org/abs/1810.11363v1

Drewery, D., Woodside, J., & Eppel, K. (2022). Artificial intelligence and résumé critique experiences. *Canadian Journal of Career Development*, *21*(2), 28–39. https://doi.org/10.53379/CJCD.2022.338

Duffy, R. D., & Sedlacek, W. E. (2007). What is most important to students' long-term career choices: Analyzing 10-year trends and group differences. *Journal of Career Development*, *34*(2), 149–163. https://doi.org/10.1177/0894845307307472

Dutta, P., Paul, S., Cengiz, K., Anand, R., & Majumder, M. (2022). A predictive method for emotional sentiment analysis by machine learning from electroencephalography of brainwave data. *Implementation of Smart Healthcare Systems Using AI, IoT, and Blockchain*, 109–130. https://doi.org/10.1016/B978-0-323-91916-6.00008-4

Edmunds, H. (1999). The focus group research handbook. *The Bottom Line*, *12*(3), 46–46. https://doi.org/10.1108/BL.1999.12.3.46.1

Enache, M., Sallán, J. M., Simo, P., & Fernandez, V. (2013). Organizational commitment within a contemporary career context. *International Journal of Manpower*, *34*(8), 880–898. https://doi.org/10.1108/IJM-07-2013-0174/FULL/XML

Erdoğmuş, N. (2003). *Kariyer geliştirme: Kuram ve uygulama*. Nobel Yayınevi.

Fablo, R. A. (2014). SABiO: Systematic approach for building ontologies. *Ontologies in Conceptual Modeling and Information Systems Engineering (ONTO-COM- ODISE 2014)*.

Falbo, R. A., Guizzardi, G., Duarte, K. C., & Natali, A. C. C. (2002). Developing software for and with reuse: an ontological approach. *International Conference on Computer Science, Software Engineering, Information Technology, E-Business and Applications, CSITeA 2002*, 311–316. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.18.9318&rep=rep1&type=pdf

Fazel-Zarandi, M., & Fox, M. S. (2012). An ontology for skill and competency management. *Frontiers in Artificial Intelligence and Applications*, *239*, 89–102. https://doi.org/10.3233/978-1-61499-084-0-89

Felaco, C., & Parola, A. (2022). Subjective well-being and future orientation of NEETs: Evidence from the italian sample of the european social survey. *Social Sciences 2022, Vol. 11, Page 482*, *11*(10), 482. https://doi.org/10.3390/SOCSCI11100482

Felstead, A., Gallie, D., Green, F., & Inanc, H. (2015). Fits, misfits and interactions: learning at work, job satisfaction and job-related well-being. *Human Resource Management Journal*, *25*(3), 294–310. https://doi.org/10.1111/1748-8583.12071

Fettke, P. (2009). How conceptual modeling is used. *Communications of the Association for Information Systems*, *25*(1), 571–592. https://doi.org/10.17705/1cais.02543

Fonseca, C. M., Guizzardi, G., Almeida, J. P. A., Sales, T. P., & Porello, D. (2022). Incorporating types of types in ontology-driven conceptual modeling. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *13607 LNCS*, 18–34. https://doi.org/10.1007/978-3-031-17995-2_2

Fonseca, C. M., Sales, T. P., Bassetti, L., & Viola, V. (2021). *OntoUML plugin for visual paradigm*. https://github.com/OntoUML/ontouml-vp-plugin

Forret, M. L., & Sullivan, S. E. (2002). A balanced scorecard approach to networking: A guide to successfully navigating career changes.

Organizational Dynamics, *31*(3), 245–258. www.organizational-dynamics.com

Forrier, A., Sels, L., & Stynen, D. (2009). Career mobility at the intersection between agent and structure: A conceptual model. *Journal of Occupational and Organizational Psychology*, *82*(4), 739–759. https://doi.org/10.1348/096317909X470933

Freund, Y., & Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, *55*(1), 119–139. https://doi.org/10.1006/JCSS.1997.1504

Gedrimiene, E., Celik, I., Kaasila, A., Mäkitalo, K., & Muukkonen, H. (2024). Artificial Intelligence (AI)-enhanced learning analytics (LA) for supporting career decisions: advantages and challenges from user perspective. *Education and Information Technologies*, *29*(1), 297–322. https://doi.org/10.1007/s10639-023-12277-4

Gemino, A., & Wand, Y. (2005). Complexity and clarity in conceptual modeling: Comparison of mandatory and optional properties. *Data and Knowledge Engineering*, *55*(3), 301–326. https://doi.org/10.1016/j.datak.2004.12.009

Ghosh, A., Woolf, B., Zilberstein, S., & Lan, A. (2020). Skill-based career path modeling and recommendation. *Proceedings - 2020 IEEE International Conference on Big Data, Big Data 2020*, 1156–1165. https://doi.org/10.1109/BigData50022.2020.9377992

Gouda, H. (2022). Exploring the effects of learning abilities, technology and market changes on the need for future skills. *Higher Education, Skills and Work-Based Learning*, *12*(5), 900–913. https://doi.org/10.1108/HESWBL-10-2021-0200/FULL/XML

Gove, R., & Faytong, J. (2012). Machine learning and event-based software testing: Classifiers for identifying infeasible gui event sequences. *Advances in Computers*, *86*, 109–135. https://doi.org/10.1016/B978-0-12-396535-6.00004-1

Granrose, C. S., & Portwood, J. D. (1987). Matching individual career plans and organizational career management. *Academy of Management Journal*, *30*(4), 699–720. https://doi.org/10.5465/256156

Greenwood, B. (2009). Facebook: The next great vetting tool. *Information Today*, *26*(8).

Guan, Y., Arthur, M. B., Khapova, S. N., Hall, R. J., & Lord, R. G. (2019). Career boundarylessness and career success: A review, integration and guide to future research. *Journal of Vocational Behavior*, *110*, 390–402. https://doi.org/10.1016/J.JVB.2018.05.013

Guarino, N. (1998). Formal ontology and information systems. *Proceedings of FOIS'98*, 3–15. https://doi.org/10.1007/978-3-642-00834-4_23

Gudivada, V. N., Irfan, M. T., Fathi, E., & Rao, D. L. (2016). Cognitive analytics: Going beyond big data analytics and machine learning. *Handbook of Statistics*, *35*, 169–205. https://doi.org/10.1016/BS.HOST.2016.07.010

Guizzardi, G. (2005). *Ontological foundations for structural conceptual models*. Centre for Telematics and Information Technology. https://doi.org/10.1007/978-3-642-31095-9_45

Guizzardi, G., Benevides, A. B., Fonseca, C. M., Porello, D., Almeida, J. P. A., & Sales, T. P. (2022). UFO: Unified foundational ontology. *Applied Ontology*, *17*(1), 167–210. https://doi.org/10.3233/AO-210256

Guizzardi, G., & Proper, H. A. (2021). On understanding the value of domain modeling. *CEUR Workshop Proceedings*, *2835*(February), 51–62.

Guizzardi, G., Wagner, G., Almeida, J. P. A., & Guizzardi, R. S. S. (2015). Towards ontological foundations for conceptual modeling: The unified foundational ontology (UFO) story. *Applied Ontology*, *10*(3–4), 259–271. https://doi.org/10.3233/AO-150157

Güner, S., Taşkın, K., Cebeci, H. İ., & Aydemir, E. (2023). Service quality in rail systems: Listen to the voice of social media. *Transportation Research Record: Journal of the Transportation Research Board*. https://doi.org/10.1177/03611981231200225

Hagendorff, T. (2020). The ethics of AI ethics: An evaluation of guidelines. *Minds and Machines*, *30*(1), 99–120. https://doi.org/10.1007/S11023-020-09517-8/TABLES/1

Hall, D. T. (1986). *Career development in organizations*. Jossey-Bass.

Hansen, J. M., Jackson, A. P., & Pedersen, T. R. (2016). Career development courses and educational outcomes. *Http://Dx.Doi.Org/10.1177/0894845316644984*, *44*(3), 209–223. https://doi.org/10.1177/0894845316644984

Heaton, J. (2016). An empirical analysis of feature engineering for predictive modeling. *Conference Proceedings - IEEE SOUTHEASTCON*, *2016-July*. https://doi.org/10.1109/SECON.2016.7506650

Herr, E. L., Cramer, S. H., & Niles, S. G. (2004). *Career guidance and counseling through the lifespan: Systematic approaches* (6th ed.). Pearson Education.

Herr, E. L., & Shahnasarian, M. (2001). Selected milestones in the evolution of career development practices in the twentieth century. *The Career Development Quarterly*, *49*(3), 225–232. https://doi.org/10.1002/j.2161-0045.2001.tb00564.x

Hevner, A. R. (2007). A three cycle view of design science research. In *Scandinavian Journal of Information Systems* (Vol. 19, Issue 2).

Hevner, A. R., & Chatterjee, S. (2010). Design research in information systems: Theory and practice. In *Springer* (Vol. 2).

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information research. *MIS Quarterly*, *28*(1), 75.

Hirschi, A. (2018). The fourth industrial revolution: Issues and implications for career research and practice. *Career Development Quarterly*, *66*(3), 192–204. https://doi.org/10.1002/cdq.12142

Hoek, J., O'Kane, P., & McCracken, M. (2016). Publishing personal information online: How employers' access, observe and utilise social networking sites within selection procedures. *Personnel Review*, *45*(1), 67–83. https://doi.org/10.1108/PR-05-2014-0099

Türkiye İş Kurumu (n.d.). *Türkiye iş kurumu - meslek meslekleri taniyalim*. Retrieved April 12, 2024, from https://esube.iskur.gov.tr/Meslek/MeslekleriTaniyalim.aspx

İstanbullu Dinçer, F., Akova, O., & Kaya, F. (2013). Meslek yüksekokulu turizm ve otel işletmeciliği programı öğrencilerinin kariyer planlaması üzerine bir araştırma: İstanbul Üniversitesi ve Gümüşhane Üniversitesi örneği. *Elektronik Mesleki Gelişim ve Araştırma Dergisi, Ejoir*, 42–56.

Jabeur, S. Ben, Gharib, C., Mefteh-Wali, S., & Arfi, W. Ben. (2021). CatBoost model and artificial intelligence techniques for corporate failure prediction. *Technological Forecasting and Social Change*, *166*, 120658. https://doi.org/10.1016/J.TECHFORE.2021.120658

Jackson, D. A. (2017). Using work-integrated learning to enhance career planning among business undergraduates. *Australian Journal of Career Development*, *26*(3), 153–164. https://doi.org/10.1177/1038416217727124

Jaffe, D. T., & Scott, C. D. (1991). Career development for empowerment in a changing work world. In J. M. Kummerow (Ed.), *New Directions in Career Planning* (pp. 33–59). CPP Books.

Jarvis, P. (2006). *Towards a comprehensive theory of human learning : lifelong learning and the learning society*. New York. https://www.routledge.com/Towards-a-Comprehensive-Theory-of-Human-Learning/Jarvis/p/book/9780415355414

Kantar, N., & Bynum, T. W. (2021). Global ethics for the digital age – flourishing ethics. *Journal of Information, Communication and Ethics in Society*, *19*(3), 329–344. https://doi.org/10.1108/JICES-01-2021-0016/FULL/XML

Karakatsanis, I., AlKhader, W., MacCrory, F., Alibasic, A., Omar, M. A., Aung, Z., & Woon, W. L. (2017). Data mining approach to monitoring the requirements of the job market: A case study. *Information Systems*, *65*, 1–6. https://doi.org/10.1016/J.IS.2016.10.009

Katuwal, R., Suganthan, P. N., & Zhang, L. (2020). Heterogeneous oblique random forest. *Pattern Recognition*, *99*, 107078. https://doi.org/10.1016/J.PATCOG.2019.107078

Kay, M., Patel, S. N., & Kientz, J. A. (2015). How good is 85%? A survey tool to connect classifier evaluation to acceptability of accuracy. *Conference on Human Factors in Computing Systems - Proceedings*, *2015-April*, 347–356. https://doi.org/10.1145/2702123.2702603

Khaouja, I., Kassou, I., & Ghogho, M. (2021). A survey on skill identification from online job ads. *IEEE Access*, *9*, 118134–118153. https://doi.org/10.1109/ACCESS.2021.3106120

Kiang, M. Y. (2003). Neural networks. *Encyclopedia of Information Systems*, 303–315. https://doi.org/10.1016/B0-12-227240-4/00121-0

Kluemper, D. H., & Rosen, P. A. (2009). Future employment selection methods: Evaluating social networking web sites. *Journal of Managerial Psychology*, *24*(6), 567–580. https://doi.org/10.1108/02683940910974134

Knowles, M. S., Holton III, E. F., Swanson, R. A., & Robinson, P. A. (2020). *The adult learner: The definitive classic in adult education and human resource development*. https://doi.org/10.4324/9780429299612

Kontio, J., Lehtola, L., & Bragge, J. (2004). Using the focus group method in software engineering: Obtaining practitioner and user experiences. *Proceedings - 2004 International Symposium on Empirical Software Engineering, ISESE 2004*, 271–280. https://doi.org/10.1109/ISESE.2004.1334914

Krunal A, V. (2014). Content evocation using web scraping and semantic illustration. *IOSR Journal of Computer Engineering*, *16*(3), 54–60. https://doi.org/10.9790/0661-16395460

Kuijpers, M. A. C. T., Schyns, B., & Scheerens, J. (2006). Career competencies for career success. *The Career Development Quarterly*, *55*(2), 168–178. https://doi.org/10.1002/J.2161-0045.2006.TB00011.X

Kuzgun, Y. (2021). *İlköğretimde rehberlik* (8th ed.). Nobel Akademik Yayıncılık.

Laal, M. (2011). Lifelong learning: What does it mean? *Procedia - Social and Behavioral Sciences*, *28*, 470–474. https://doi.org/10.1016/J.SBSPRO.2011.11.090

Langford, J., & McDonagh, D. (2003). *Focus groups: Supporting effective product development*. Taylor and Francis.

Lau, P. L., Anctil, T., Ee, G. T., Jaafar, J. L. S., & Kin, T. G. (2020). Self-concept, attitudes toward career counseling, and work readiness of malaysian vocational students. *The Career Development Quarterly*, *68*(1), 18–31. https://doi.org/10.1002/CDQ.12210

Lawson, R. (1989). Web scraping with python. In *Journal of Chemical Information and Modeling* (Vol. 53). https://doi.org/10.1017/CBO9781107415324.004

Le Deist, F. D., & Winterton, J. (2005). What is competence? *Human Resource Development International*, *8*(1), 27–46. https://doi.org/10.1080/1367886042000338227

Lee, D., Kim, M., & Na, I. (2018). Artificial intelligence based career matching. *Journal of Intelligent & Fuzzy Systems*, *35*(6), 6061–6070. https://doi.org/10.3233/JIFS-169846

Leibowitz, Z. B., Farren, C., & L. Kaye, B. (1991). *Designing career development systems*. Pfeiffer.

Lindland, O. I., Sindre, G., & Solvberg, A. (1994). Understanding quality in conceptual modeling. *IEEE Software*, *11*(2), 42–49. https://doi.org/10.1109/52.268955

Luo, Y. (2016). Research on career planning of college students. *Proceedings of the 2016 International Conference on Economics, Social Science, Arts, Education and Management Engineering*, 856–861. https://doi.org/10.2991/ESSAEME-16.2016.172

Marneni, D., & Vemula, S. (2022). Analysis of Covid-19 using machine learning techniques. *Statistical Modeling in Machine Learning: Concepts and Applications*, 37–53. https://doi.org/10.1016/B978-0-323-91776-6.00006-3

Martinez-Plumed, F., Contreras-Ochando, L., Ferri, C., Hernandez-Orallo, J., Kull, M., Lachiche, N., Ramirez-Quintana, M. J., & Flach, P. (2019). CRISP-DM twenty years later: From data mining processes to data science trajectories. *IEEE Transactions on Knowledge and Data Engineering*, *33*(8), 3048–3061. https://doi.org/10.1109/TKDE.2019.2962680

Mashayekhi, M., & Head, M. (2018). Taking advantage of using p-snss taking advantage of using professionally-oriented social network sites: The role of users' actions and profiles. *AMCIS2018*. https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1452&context=amcis2018

Mashayekhi, M., & Head, M. (2022). Developing social capital through professionally oriented social network sites. *Information & Management*, *59*(6), 103664. https://doi.org/10.1016/J.IM.2022.103664

Masolo, C., Borgo, S., Gangemi, A., Guarino, N., & Oltramari, A. (2003). Ontology library. In *WonderWeb Deliverable D18* (p. 343). Laboratory For Applied Ontology - ISTC-CNR.

McDaniels, C., & Gysbers, N. C. (1991). *Counseling for career development : theories, resources, and practice*. Jossey-Bass.

Milev, P. (2017). Conceptual approach for development of web scraping application for tracking information. *Economic Alternatives*, *3*, 475–485.

Miller, J. H. (2004). Extending the use of constructivist approaches in career guidance and counselling: Solution focused strategies. *Australian Journal of Career Development*, *13*(1), 50–59. https://doi.org/10.1177/103841620401300108

Miranda, S., Orciuoli, F., Loia, V., & Sampson, D. (2017). An ontology-based model for competence management. *Data & Knowledge Engineering*, *107*, 51–66. https://doi.org/10.1016/J.DATAK.2016.12.001

Montgomery, B. L. (2017). Mapping a mentoring roadmap and developing a supportive network for strategic career advancement. *SAGE Open*, *7*(2). https://doi.org/10.1177/2158244017710288/ASSET/IMAGES/LARGE/10.1177_2158244017710288-FIG3.JPEG

Moslemi, A. (2023). A tutorial-based survey on feature selection: Recent advancements on feature selection. *Engineering Applications of Artificial Intelligence*, *126*. https://doi.org/10.1016/j.engappai.2023.107136

Mesleki Yeterlilik Kurumu (n.d.). *Ulusal meslek standartları*. Retrieved April 12, 2024, from https://portal.myk.gov.tr/index.php?option=com_meslek_std_taslak&view=taslak_listesi_yeni&msd=2&Itemid=432

Mylopoulos, J. (1998). Information modeling in the time of the revolution. *Information Systems*, *23*(3–4), 127–155. https://doi.org/10.1016/s0306-4379(98)00005-2

Nafukho, F. M., Hairston, N. R., & Brooks, K. (2004). Human capital theory: Implications for human resource development. *Human Resource Development International*, *7*(4), 545–551. https://doi.org/10.1080/1367886042000299843

Natekin, A., & Knoll, A. (2013). Gradient boosting machines, a tutorial. *Frontiers in Neurorobotics*, *7*(DEC), 63623. https://doi.org/10.3389/FNBOT.2013.00021/BIBTEX

Nicholson, N., Giusti, F., & Martos, C. (2023). Ontology-based AI design patterns and constraints in cancer registry data validation. *Cancers 2023, Vol. 15, Page 5812*, *15*(24), 5812. https://doi.org/10.3390/CANCERS15245812

Nilsson, N. J. (1998). Neural networks. *Artificial Intelligence: A New Synthesis*, 37–57. https://doi.org/10.1016/B978-0-08-049945-1.50008-3

Nunamaker, J. F., Chen, M., & Purdin, T. D. M. (1991). Systems development in information systems research. *Journal of Management Information Systems*, *7*(3), 86–106.

Olivé, A. (2007). Conceptual modeling of information systems. In *Conceptual Modeling of Information Systems*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-39390-0

On ikinci kalkınma planı. (n.d.). Retrieved May 12, 2024, from https://onikinciplan.sbb.gov.tr/

OntoUML. (n.d.). Retrieved April 11, 2024, from https://ontouml.org/

Osborn, D. S., & Baggerly, J. N. (2004). School counselors' perceptions of career counseling and career testing: Preferences, priorities, and predictors. *Http://Dx.Doi.Org/10.1177/089484530403100104*, *31*(1), 45–59. https://doi.org/10.1177/089484530403100104

O'Sullivan, D., Abela, A. V., & Hutchinson, M. (2009). Marketing performance measurement and firm performance: Evidence from the European high-technology sector. *European Journal of Marketing*, *43*(5–6), 843–862. https://doi.org/10.1108/03090560910947070

Ölmez, E., & Er, O. (2022). Makine öğrenmesi temelli yöntemler - yapay sinir ağlarına dayalı teknikler. In A. Hızıroğlu, H. İ. Cebeci, & K. Seçkin Codal (Eds.), *Veri modelleme ve veri analitiği: Sağlık ve işletme uygulamaları* (pp. 316–355). Nobel Akademik Yayıncılık.

Özdemir, Y., & Kibar, A. (2018). Kariyer yolculuğunda akıllı bir adım. *5th International Management Information Systems Conference Kariyer*.

Özgen, H., & Yalçın, A. (2017). *İnsan kaynakları yönetimi stratejik bir yaklaşım*. Akademisyen Kitabevi.

Pandya, S. S., & Wang, J. (2024). Artificial intelligence in career development: a scoping review. *Human Resource Development International*. https://doi.org/10.1080/13678868.2024.2336881

Paquette, G., Marino, O., & Bejaoui, R. (2021). A new competency ontology for learning environments personalization. *Smart Learning Environments*, *8*(1), 1–23. https://doi.org/10.1186/S40561-021-00160-Z/FIGURES/9

Perrone, K. M., Ægisdóttir, S., Webb, L. K., & Blalock, R. H. (2006). Work-family interface. *Journal of Career Development*, *32*(3), 286–300. https://doi.org/10.1177/0894845305283002

Petkov, R. (2021). Digital media literacy, artificial intelligence and modernization of youth career services. *2021 30th International Scientific Conference Electronics, ET 2021 - Proceedings*. https://doi.org/10.1109/ET52713.2021.9579955

Pittman, S. J., & Brown, K. A. (2011). Multi-scale approach for predicting fish species distributions across coral reef seascapes. *PLOS ONE*, *6*(5), e20583. https://doi.org/10.1371/JOURNAL.PONE.0020583

Polat, F. B., Ardıç, K., & Özdemir, Y. (2016). Bireysel kariyer planlamada etkili olan faktörlerin belirlenmesine yönelik bir araştırma sakarya devlet okulları örneği. *İşletme Bilimi Dergisi*, *4*(1), 29–65. https://doi.org/10.22139/ibd.13907

Polidoro, F., Giannini, R., Conte, R. Lo, Mosca, S., & Rossetti, F. (2015). Web scraping techniques to collect data on consumer electronics and airfares for Italian HICP compilation. *Statistical Journal of the IAOS*, *31*(2), 165–176. https://doi.org/10.3233/sji-150901

Presbitero, A., & Teng-Calleja, M. (2023). Job attitudes and career behaviors relating to employees' perceived incorporation of artificial intelligence in the workplace: a career self-management perspective. *Personnel Review*, *52*(4), 1169–1187. https://doi.org/10.1108/PR-02-2021-0103/FULL/XML

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: unbiased boosting with categorical features. *Advances in Neural Information Processing Systems*, *31*. https://github.com/catboost/catboost

Quigley, N. R., & Tymon, W. G. (2006). Toward an integrated model of intrinsic motivation and career self-management. In *Career Development International* (Vol. 11, Issue 6, pp. 522–543). https://doi.org/10.1108/13620430610692935

Racene, A., oec, M., ed, M., & Dislere, V. (2019). Choice of a Value Identification Method in Career Counselling. *Environment. Education. Personality*, *12*, 10–11. https://doi.org/10.22616/REEP.2019.038

Ren, Y., & Ren, Y. (2018). A framework of petroleum information retrieval system based on web scraping with python. *2018 15th International Conference on Service Systems and Service Management, ICSSSM 2018*, 1–6. https://doi.org/10.1109/ICSSSM.2018.8465013

Robertson, P. J. (2018). Developing career capabilities in "NEET" young people: experiences of participants in the Prince's trust team programme. *British Journal of Guidance & Counselling*, *46*(6), 752–764. https://doi.org/10.1080/03069885.2018.1434130

Saltz, J. S. (2021). CRISP-DM for data science: Strengths, weaknesses and potential next steps. *Proceedings - 2021 IEEE International Conference on Big Data, Big Data 2021*, 2337–2344. https://doi.org/10.1109/BIGDATA52589.2021.9671634

Savickas, M. L. (2013). Career construction theory and practice. In S. D. Brown & R. W. Lent (Eds.), *Career development and counseling: Putting theory and research to work* (second, pp. 144–180). John Wiley and Sons, Inc.

Savickas, M. L. (2014). Career counseling paradigms: Guiding, developing, and designing. *APA Handbook of Career Intervention, Volume 1: Foundations.*, 129–143. https://doi.org/10.1037/14438-008

Savickas, M. L., Nota, L., Rossier, J., Dauwalder, J. P., Duarte, M. E., Guichard, J., Soresi, S., Van Esbroeck, R., & van Vianen, A. E. M. (2009). Life designing: A paradigm for career construction in the 21st century. *Journal of Vocational Behavior*, *75*(3), 239–250. https://doi.org/10.1016/J.JVB.2009.04.004

Schmidt, A., & Kunzmann, C. (n.d.). *Sustainable competency-oriented human resource development with ontology-based competency catalogs*. http://www.professional-learning.eu/

Schneider, P., & Xhafa, F. (2022). Machine learning. *Anomaly Detection and Complex Event Processing over IoT Data Streams*, 149–191. https://doi.org/10.1016/B978-0-12-823818-9.00019-5

Scholl, M. B., & Cascone, J. (2010). The constructivist résumé: Promoting the career adaptability of graduate students in counseling programs. *The Career Development Quarterly*, *59*(2), 180–191. https://doi.org/10.1002/J.2161-0045.2010.TB00061.X

Schröer, C., Kruse, F., & Gómez, J. M. (2021). A systematic literature review on applying CRISP-DM process model. *Procedia Computer Science*, *181*, 526–534. https://doi.org/10.1016/J.PROCS.2021.01.199

S.C.M. de S Sirisuriya. (2015). A comparative study on web scraping. *8th International Research Conference KDU, November*, 135–140. http://www.kdu.ac.lk/proceedings/irc2015/2015/com-020.pdf

Seger, Y. R. (2016). Individual development plans: Determining your strengths, identifying areas for improvement, and initiating conversations about your career. *The FASEB Journal*, *30*(S1), 116.1-116.1. https://doi.org/10.1096/FASEBJ.30.1_SUPPLEMENT.116.1

Sevinç, A., & Siyez, D. M. (2018). Öğrenci kariyer yapılandırma envanteri'nin türkçe uyarlaması: Geçerlik ve güvenirlik çalışması. *İş ve İnsan Dergisi*. https://doi.org/10.18394/iid.349410

Shahabi, H., Jarihani, B., Piralilou, S. T., Chittleborough, D., Avand, M., & Ghorbanzadeh, O. (2019). A semi-automated object-based gully networks detection using different machine learning models: A case study of bowen catchment, Queensland, Australia. *Sensors 2019, Vol. 19, Page 4893*, *19*(22), 4893. https://doi.org/10.3390/S19224893

Simon, H. (1996). *The sciences of the artificial* (Third Edition). MIT Press.

Sinnott, R. O., Duan, H., & Sun, Y. (2016). A case study in big data analytics: Exploring twitter sentiment analysis and the weather. *Big Data: Principles and Paradigms*, 357–388. https://doi.org/10.1016/B978-0-12-805394-2.00015-5

Slamet, C., Andrian, R., Maylawati, D. S., Suhendar, Darmalaksana, W., & Ramdhani, M. A. (2018). Web scraping and naïve bayes classification for job search engine. *IOP Conference Series: Materials Science and Engineering*, *288*(1). https://doi.org/10.1088/1757-899X/288/1/012038

Smith, B., & Welty, C. (2001). Ontology: Towards a new synthesis. *Formal Ontology in Information Systems: Collected Papers from the Second International Conference*, *January 2001*.

Smith-Ruig, T. (2009). Exploring career plateau as a multi-faceted phenomenon: Understanding the types of career plateaux experienced by accounting professionals. *British Journal of Management*, *20*(4), 610–622. https://doi.org/10.1111/J.1467-8551.2008.00608.X

Sterner, W. R. (2012). Integrating existentialism and super's life-span, life-space approach. *The Career Development Quarterly*, *60*(2), 152–162. https://doi.org/10.1002/J.2161-0045.2012.00013.X

Subasi, A. (2020). Machine learning techniques. *Practical Machine Learning for Data Analysis Using Python*, 91–202. https://doi.org/10.1016/B978-0-12-821379-7.00003-5

Subasi, A., Panigrahi, S. S., Patil, B. S., Canbaz, M. A., & Klén, R. (2022). Advanced pattern recognition tools for disease diagnosis. *5G IoT and Edge Computing for Smart Healthcare*, 195–229. https://doi.org/10.1016/B978-0-323-90548-0.00011-5

Sugiyama, M. (2016). Ensemble learning. *Introduction to Statistical Machine Learning*, 343–354. https://doi.org/10.1016/B978-0-12-802121-7.00041-8

Sullivan, S. E. (2001). Careers in the 21st Century. *Group & Organization Management*, *26*(3), 252–254. https://doi.org/10.1177/1059601101263002/ASSET/1059601101263002.FP.PNG_V03

Sullivan, S. E., & Baruch, Y. (2009). Advances in Career Theory and Research: A Critical Review and Agenda for Future Exploration. *Journal of Management*, *35*(6), 1542–1571. https://doi.org/10.1177/0149206309350082

Super, D. E. (1963). The definition and measurement of early career behavior: A first formulation. *The Personnel and Guidance Journal*, *41*(9), 775–780. https://doi.org/10.1002/j.2164-4918.1963.tb02394.x

Super, D. E. (1980). A life-span, life-space approach to career development. *Journal of Vocational Behavior*, *16*(3), 282–298. https://doi.org/10.1016/0001-8791(80)90056-1

Tarasov, V. (2012). Ontology-based approach to competence profile management. *Journal of Universal Computer Science*, *18*(20), 2893–2919.

Templeton, J. F. (1994). *The focus group: A strategic guide to organizing, conducting and analyzing the focus group interview*. McGraw-Hill Professional Publishing.

Tomy, S., & Pardede, E. (2019). Map my career: Career planning tool to improve student satisfaction. *IEEE Access*, *7*, 132950–132965. https://doi.org/10.1109/ACCESS.2019.2940986

Unschold, M., & Gruninger, M. (1996). Ontologies: Principles, methods and applications. *Knowledge Engineering Review*, *11*(2).

Vaishnavi, V. K., & Kuechler, W. (2007). Design science research methods and patterns: Innovating information and communication technology. In *Design Science Research Methods and Patterns: Innovating Information and Communication Technology*. https://doi.org/10.1201/9781420059335

Van der Horst, A. C., Klehe, U. C., & Van der Heijden, B. I. J. M. (2017). Adapting to a looming career transition: How age and core individual differences interact. *Journal of Vocational Behavior*, *99*, 132–145. https://doi.org/10.1016/J.JVB.2016.12.006

Van Dijck, J. (2013). "You have one identity": Performing the self on Facebook and LinkedIn. *Media, Culture and Society*, *35*(2), 199–215. https://doi.org/10.1177/0163443712468605

Van Dijk, M. S. (2004). Career development within HRD: Foundation or fad? *Academy of Human Resource Development International Conference*, 771–778.

Vapnik, V. N. (1998). *The nature of statistical learning theory* (Second Edition). Springer.

Verbruggen, M., Dries, N., & Van Laer, K. (2016). Challenging the uniformity myth in career counseling outcome studies. *Journal of Career Assessment*, *25*(1), 159–172. https://doi.org/10.1177/1069072716657797

Verdonck, M., Gailly, F., Pergl, R., Guizzardi, G., Martins, B., & Pastor, O. (2019). Comparing traditional conceptual modeling with ontology-driven conceptual modeling: An empirical study. *Information Systems*, *81*(February 2021), 92–103. https://doi.org/10.1016/j.is.2018.11.009

Verdonck, T., Baesens, B., Óskarsdóttir, M., & vanden Broucke, S. (2021). Special issue on feature engineering editorial. *Machine Learning*, 1–12. https://doi.org/10.1007/S10994-021-06042-2/METRICS

Vergiliel Tüz, M. (2003). Kariyer planlamasında yeni yaklaşımlar. *Uludağ Üniversitesi Fen-Edebiyat Fakültesi Sosyal Bilimler Dergisi*, *4*(4), 169–176.

Volmer, J., & Spurk, D. (2011). Protean and boundaryless career attitudes: relationships with subjective and objective career success. *Zeitschrift Fur Arbeitsmarktforschung*, *43*(3), 207–218. https://doi.org/10.1007/S12651-010-0037-3/TABLES/3

Waddell, J., & Bauer, M. (2005). Career planning and development for students: Building a career in a professional practice discipline. *Canadian Journal of*

Career Development, *4*(2), 24–39. http://ceric.ca/cjcd/book/CJCD10thAnniversaryFullBook.pdf

Wang, D., & Li, Y. (2024). Career construction theory: Tools, interventions, and future trends. *Frontiers in Psychology*, *15*, 1381233. https://doi.org/10.3389/FPSYG.2024.1381233

Wang, S., & Summers, R. M. (2012). Machine learning and radiology. *Medical Image Analysis*, *16*(5), 933–951. https://doi.org/10.1016/J.MEDIA.2012.02.005

Welcome to Python.org. (n.d.). Retrieved April 11, 2024, from https://www.python.org/

Westman, S., Kauttonen, J., Klemetti, A., Korhonen, N., Manninen, M., Mononen, A., Niittymäki, S., & Paananen, H. (2021). Artificial intelligence for career guidance - current requirements and prospects for the future. *IAFOR Journal of Education*, *9*(4), 43–62. https://doi.org/10.22492/IJE.9.4.03

Willis, L., & Wilkie, L. A. (2009). Digital career portfolios: Expanding institutional opportunities. *Journal of Employment Counseling*, *46*(2), 73–81. https://doi.org/10.1002/j.2161-1920.2009.tb00069.x

Wilson, M., Robertson, P., Cruickshank, P., & Gkatzia, D. (2022). Opportunities and risks in the use of AI in career development practice. *Journal of the National Institute for Career Education and Counselling* , *48*(1), 48–57. https://doi.org/10.20856/JNICEC.4807

Wirtz, J., Patterson, P. G., Kunz, W. H., Gruber, T., Lu, V. N., Paluch, S., & Martins, A. (2018). Brave new world: Service robots in the frontline. *Journal of Service Management*, *29*(5), 907–931. https://doi.org/10.1108/JOSM-04-2018-0119/FULL/PDF

Yang, Y. (2017). Ensemble learning. *Temporal Data Mining Via Unsupervised Ensemble Learning*, 35–56. https://doi.org/10.1016/B978-0-12-811654-8.00004-X

Yeşilyaprak, B. (2019). Türkiye'de mesleki rehberlik ve kariyer danışmanlığı hizmetleri: Güncel durum ve öngörüler. *Kariyer Psikolojik Danışmanlığı Dergisi*, *2*(2), 73–102. https://doi.org/10.2019/10.21.2019

Yeşilyaprak, B. (2022). Mesleki rehberlik ve kariyer danışmanlığına giriş. In Binnur Yeşilyaprak (Ed.), *Mesleki rehberlik ve kariyer danışmanlığı* (14th ed.). Pegem Akademi.

Yüksek Öğretim Kurumu. (n.d.-a). *YÖK üniversiteler*. Retrieved April 12, 2024, from https://www.yok.gov.tr/universiteler/universitelerimiz

Yüksek Öğretim Kurumu. (n.d.-b). *YÖK yükseköğretim bölüm ve program isimleri*. Retrieved April 12, 2024, from https://www.yok.gov.tr/Sayfalar/Kurumsal/IdariBirimler/egitim_ogretim_daire_bsk/yuksekogretim-bolum-ve-program-isimleri.aspx

Zaouga, W., Rabai, L. B. A., & Alalyani, W. R. (2019). Towards an ontology based-approach for human resource management. *Procedia Computer Science*, *151*, 417–424. https://doi.org/10.1016/j.procs.2019.04.057

Zhong, A., Qiu, R., Zhang, X., Lv, S., & Song, L. (2023). Characterizing the job-task-skill pattern of job requirements with job advertisement mining. *ACM International Conference Proceeding Series*, 139–147. https://doi.org/10.1145/3584816.3584838

APPENDIX

Appendix 1. *Feature Set*

Features Used for ML/AI Training
Duration of the most recent job
Total Job Experience Duration
Average Job Duration
Last Job Position
Graduated Education Degree (Bachelor’s, Master’s, PhD)
Graduated Bachelor’s Degree Name (if exist)
The Number (count value) of each Skill Categories declared, adopted from Aktaş et al., (2022) (Social, Managerial, Organizational, Problem Solver, Methodologies, Software, Hardware)
The Number (count value) of each Skills declared, adopted from Aktaş et al., (2022) (Teamwork, Communication Skills, Responsibility, Leadership, Project Management, Functional Expertise (and its variants of business, finance, accounting, human resources, logistics and supply chain, production management and planning, quality control, information systems, marketing, sales, management information systems), Sectoral Expertise, Analytical Thinking, Problem Solving, Innovative, Basic IT, System Methodologies, Enterprise Modules, Analysis Tolls and Techniques, Documentation, Data, Software Development, Database, Foreign Language, User Interface and User Experience Design, Cyber Security, Server Related Hardware, Virtualization Related Hardware, Cloud Related Hardware)

CURRICULUM VITAE

Name Surname: Bahadır AKTAŞ

EDUCATION STATUS

Degree of Education	Place of Study	Year of Study
Doctorate	Sakarya University / Graduate School of Business / Management Information Systems	2018-Currently
Master's	Dokuz Eylül University / Graduate School of Social Sciences/ Management Information Systems	2015-2017
Bachelor's	Ege University / Faculty of Engineering / Electrical and Electronics Engineering	2009-2014
High School	Bornova Anatolian High School	2005-2009

JOB EXPERIENCE

Year	Location	Position
2018 - Currently	Sakarya University	Research Assistant
2023 – 2023	Twente University	Visiting Researcher
2017 – 2018	PiaQa	Software Test Engineer

FOREIGN LANGUAGE

English (91.25/100 YDS Score)

PUBLISHED WORKS

Aktaş, B. (2024). Design science research. In A. Akbıyık (Eds.), *Current and prospective approaches, methods, and techniques for management information systems research* (pp. 17–31). Sakarya Yayıncılık.

Aktaş, B., Akbıyık, A., Nusair, K., & Okumus, F. (2024). Travellers' social media postings during protests and mass demonstrations. *Current Issues in Tourism*, *27*(10), 1513–1529. https://doi.org/10.1080/13683500.2023.2214359

Karayel, T., Aktaş, B. and Akbıyık, A. (2024), Human factors in remote work: Examining cyber hygiene practices, *Information and Computer Security*, Vol. ahead-of-print No. ahead-of-print. https://doi.org/10.1108/ICS-11-2023-0215

Aktaş, B., & Akbıyık, A. (2023, October 18). *The potential of the use of ontology-driven conceptual modeling and ontouml in management information systems studies*. 10th International Management Information Systems Conference. https://2023.imisc.net/tr/imisc23_bildiri_ozet_kitapcigi.pdf

Aktaş, B., & Uzavcı, M. (2022). Veri görselleştirme. In A. Hızıroğlu, H. İ. Cebeci, & K. Seçkin Codal (Eds.), *Veri modelleme ve veri analitiği: Sağlık ve işletme uygulamaları* (pp. 122–149). Nobel Akademik Yayıncılık.

Kurt, K., Aktaş, B., & Turan, A. (2022). Understanding online shopping continuance intention of turkish users: An empirical assessment. *Pamukkale Üniversitesi Sosyal Bilimler Enstitüsü Dergisi* (50), 144-160. https://doi.org/10.30794/pausbed.1003073

Aktaş, B., Baz Aktaş, N., & Akbıyık, A. (2022). İşgücü piyasalarında yönetim bilişim sistemleri programlarının farkındalığı: ABD ve Türkiye'deki iş ilanları üzerinden bir değerlendirme. *Journal of Research in Business*, *7(IMISC2021 Special Issue)*, 60-79. https://doi.org/10.54452/jrb.1024997

Türkmen Barutçu, M., Aktaş B. (2022). Online yemek siparişlerinde müşteri şikayetleri. In N. Sütütemiz, M. Saygılı (Eds.), *Yeni normalde tüketici davranışları teori - araştırma – uygulama* (pp. 253-277). Gazi Kitabevi.

Karayel, T., Aktaş, B., & Akbıyık, A. (2022, December). *Kullanıcıların bilgisayar güvenliği davranışını inceleme: Siber hijyen*. 7th International Management Information Systems Conference, İzmir, TURKIYE.

Baz Aktaş, N., Aktaş, B., & Akbıyık, A. (2021). Koronavirüs'ün (covid-19) Türkiye'de e-ticaret müşteri memnuniyetine etkisi: Trendyol örneği. *Journal of Information Systems and Management Research*, 3(1), 39-50.

Aktaş, B., Kurt, K., & Turan, A. H. (2019). Üniversite öğrencilerinin sosyal medya kullanım tercihlerinin incelenmesi. *Yönetim Bilişim Sistemleri Dergisi*, 5(1), 1-15.

Aktaş, B., & Akbıyık, A. (2019, October). *Social connectedness in professionally-oriented social network sites and its effect on social network value*. 6th International Management Information Systems Conference.

Aktaş, B., & Aydın, C. (2018). Talaşlı imalat sektöründe zaman serileri kullanarak üretim etkililiğinin tahmini. *Bilişim Teknolojileri Dergisi*, 11(4), 407-416. https://doi.org/10.17671/gazibtd.383339

Aktaş, B., Ediz, Ç., Alma Çallı, B., Coşkun, E. (2018, October). *Transforming a truly traditional dairy product company's logistics and transportation related decision making processes*. 5th International Management Information Systems Conference, Ankara, Turkiye.

Aydın, C., & Aktaş, B. (2018). Talaşlı imalat sektöründe nesnelerin interneti ve anlık veri analizi yöntemleri kullanarak üretim etkinliğinin artırılması. *İstanbul Üniversitesi İşletme Fakültesi Dergisi*, 47(1), 33-45.

Baz, N., Aktaş, B., Akcura, N., Sokullu, R., Uyar, E. (2015). *BikeEge - bicycle sharing system for ege university students.* International Scientific Conference on Information, Communication and Energy Systems and Technologies, ICEST 2015, Sofia, Bulgaria.