Data scientists are new knowledge workers representing the knowledge economy era. Knowledge workers perform unstandardized works that solve ambiguity-intensive problems. Therefore, self-leadership, which emphasizes self-motivated, autonomous judgment and execution, significantly influences their work-related outcomes. Even knowledge workers have high occupational commitment, they usually show low organizational commitment. Knowledge workers’ intention to leave is also relatively high due to this reason. This study focused on data scientists’ self-leadership, predicted that self-leadership would increase an organization’s commitment and intention to leave. Based on the trait activation theory (TAT), the author also confirmed how perceived job autonomy enhances self-leadership influences. Results showed that data scientists’ self-leadership significantly lowered intention to leave through organizational commitment and this mediating effect was moderated by perceived job autonomy. This study broadened the theoretical understanding of the effects of knowledge workers’ self-leadership and presented practical implications for managing data scientists.
I. Introduction

The scope and depth of data analysis have recently been expanding due to the rapid development of related technologies. McKinsey[1] predicted that machine learning techniques, combined with advances in data collection technology and increased computational power, will create disruptive innovation in all industries. To cope with this revolutionary change, data scientists are recruiting competitively in various fields. Several research organizations have referred to data scientists as the most promising job in the future—Davenport and Patil[2] refer to data scientists as “the sexist job of the 21st century.” Data scientists are new knowledge workers in the knowledge economy era[3].

Previous studies typically defined knowledge workers as highly educated and the input and output of their work are concerned mainly with the handling and production of information[4]. Edgar et al.[5] focused on the job characteristics of knowledge workers and considered it as “any worker whose job involves a significant amount of gathering, creating, and dissemination of knowledge.” Other researchers pointed out that knowledge workers are characterized by the ability to handle and produce information with theoretical and analytical knowledge[6].

Similar to other knowledge workers, data scientists require a strong educational background. To become data scientists, they usually require a high level of knowledge in mathematics, statistics, and/or computer science. A data scientist is “an expert who possesses advanced skills to collect, analyze, interpret, and visualize data and derive meaningful implications based on the results of analysis”[7]. According to a related survey, 88% of data scientists had a master’s degree or higher and 46% had a doctorate[8]. Gehl[9] referred to data scientists as “specialized and hard-earned training talents that are not easily replicable.”

A prominent feature of the knowledge workers’ job is that it deal with ambiguity-intensive information or knowledge, has a low level of standardization, and requires less direct supervision[10,11]. According to the self-leadership theory, in the information age, value is created through the autonomous judgment and execution of employees, not the leader’s direction and control[12]. For this reason, knowledge workers’ self-leadership positively influences work-related outcomes[13]. Self-leadership should be considered an essential characteristic of data scientists in that it has a significant influence on the work-related outcomes for those who perform non-routine and under-designed work[14]. A data science project is a challenging process that solves unstandardized problems. Simultaneously, it is necessary to pay attention to working situations that could activate self-leadership such as job autonomy[15][16].

Knowledge workers have low organizational commitment[17]. Previous researchers have pointed out that knowledge workers have a high occupational commitment as experts, which acts as a cause of conflict with employers and nonprofessional coworkers[17][18]. Knowledge workers’ low organizational commitment negatively affects work outcomes, especially as a direct influencing factor that increases intention to leave[19]. Data scientists share the same problems of low organizational commitment and high intention to leave[20].
Considering that perceived alternative job opportunities (PAJO) is a powerful factor in increasing knowledge workers’ intention to leave [21], retaining data scientists is a challenging since they are in higher demand than ever.

In the knowledge economy era, where knowledge is the key to value creation, knowledge workers determine companies’ competitive advantages [22]. In addition, the turnover of knowledge workers means the outflow of knowledge, which is a company’s core asset. Under the industry 4.0 paradigm, where artificial intelligence and big data are leading innovation in business [23], low organizational commitment and high turnover rate of data scientists directly lead to weakened corporate future competitiveness. Academic researches on the influences of self-leadership—an important trait required for data scientists—on work-related outcomes such as organizational commitment and intention to leave are still insufficient. Previous studies have focused on defining data scientists [24], organizing common qualities and skills [25][26], and establishing plans to train them [27–30].

Data scientists can be called new knowledge workers, but it is not clear whether they can be explained in the traditional knowledge workers’ framework. First, collaboration is essential for data scientists because the activity of data scientists creating knowledge from data and information is a practice that requires a team with an interdisciplinary background [31]. In addition, knowledge sharing and collaboration tend to be conducted via peer-to-peer knowledge sharing across organizational boundaries rather than between traditional knowledge workers [32][33]. There may also be generation differences in that there is a higher proportion of Generation X and Millennials, with different work values and needs, than in other professional jobs [34][35].

The purpose of the study is to demonstrate the underlying mechanism of how knowledge worker’s self-leadership lowers their intention to leave. Although the high turnover rate is a characteristic of knowledge workers [19], there are still no studies on the effect of self-leadership on knowledge workers’ turnover as far as the authors know. It is particular meaningful in that an empirical analysis was conducted for data scientists with high perceived alternative job opportunities [24].

Based on self-leadership theory, self-leadership is expected to lower intention to leave through organizational commitment [36][37]. In addition, based on trait activation theory (TAT) [38][39], the authors predicted that perceived job autonomy would control the direct effect of self-leadership on organizational commitment and the mediating effect on intention to leave through organizational commitment.

The importance of data scientists is growing as essential knowledge workers influencing an organization’s competitiveness [40]. Under the situation, this research provides a theoretical understanding of data scientists and practical implications for retention them successfully.

II. Literature Review

1. Self-Leadership and intention to leave

Self-leadership is defined as the responsible behavior and attitude that employees take to lead themselves [39]. The construct refers to the
self-influence exercised to achieve the self-direction and self-motivation necessary for a task's successful execution, providing a comprehensive understanding of the self-influencing processes. Self-leadership is a new type of leadership that maximizes the potential of employees themselves, which explains the changed modern organization and employees' values [41].

The self-leadership theory emphasizes that in an knowledge economy autonomous judgment, execution, and management are essential, rather than the instruction and control of others, for employees to demonstrate their potential and creativity [39]. Therefore, previous studies argued that self-leadership influences knowledge workers' outcomes who must manage a under-designed work [4]. In many cases, knowledge work is non-routine, and what to do and how to perform the task is not clearly defined. Therefore, substantial responsibility for planning, prioritizing, coordinating, and executing work rests with individual knowledge workers [4]. For this reason, knowledge workers' self-leadership showed that it increases empowerment and commitment [13]; creates work intensity and productivity [42]; and strengthens organizational commitment, trust, and team productivity [43]. Taken together, self-leadership can be an essential quality for performing data science tasks that involve solving challenging, under-designed, and non-routine work [44].

Self-leadership is closely related to the activation of intrinsic motivation in that it places importance on the natural rewards obtained by the work itself [12]. The self-leadership theory, focusing on natural rewards, is based on cognitive evaluation theory (CET) [45]. According to cognitive evaluation theory, the two factors that promote intrinsic motivation are the feeling of competence and self-determination, which form the core of the natural rewards of self-leadership [46]. In other words, intrinsic motivation behavior reinforces natural rewards such as the feeling of competence and self-determination [45], and self-leadership controls behavior and the effects on the environment by itself to obtain these natural rewards [47]. Intrinsic motivation, increased through natural rewards, promotes spontaneous, creative task performance [46]. For this reason, the natural rewards due to self-leadership have been recognized as an antecedent that improves work-related outcomes. Self-leadership theorists have revealed that employees with high self-leadership have opportunity-oriented mental constructs, strive to overcome challenging situations, and are willing to solve self-problems [48]. Therefore, employees with high self-leadership tend to focus on the positive side of work, orient the mental process to become a part of their work, willingly strive to achieve organizational goals, and commit to the organization they belong to [49].

Subsequently, this research predicted that self-leadership would be a factor that lowers intention to leave. As mentioned earlier, self-leadership's natural rewards strategy focuses on the satisfying side of the situation and triggers the correct actions [49]. In other words, the employees themselves build positive characteristics for the task so that it is a reward in itself and intentionally places the rewarding characteristics by focusing on the ungratifying features of the task. According to the attraction–selection–attrition (ASA) theory, members who
have trouble adjusting to the organization because they think they are not suitable for the organization tend to leave[50]. On the other hand, employees with high self-leadership perform tasks more innovatively and creatively[51] and endeavor to contribute to the organization through non-task behaviors such as organized citizenship behavior[52]. These efforts and achievements make employees more adaptable to the organization, resulting in a lower intention to leave. Previous studies have shown that self-leadership is a significant factor that lowers intention to leave[53,54]. Accordingly, the author hypothesized that data scientists’ self-leadership would significantly lower the intention to leave.

H1: The data scientists’ self-leadership will lower their intention to leave.

2. Mediating Role of Organizational Commitment

Organizational commitment refers to the emotional attachment and unity that members feel toward the organization they belong to[55]. Organizational commitment is characterized by strong trust and acceptance of the organization’s goals and values, the will to dedicate itself to the organization’s interests, and the desire to maintain the status of the organization’s members[56]. Allen and Meyer[57] explained that organizational commitment is a psychological state consisting of: 1) affective commitment, which indicates emotional attachment, identification, and involvement to the organization; 2) normative commitment, which indicates a moral obligation to remain in the organization; and 3) continuance commitment, which indicates a perception of rising cost when leaving the organization. Some researchers suggested that affective commitment is the only dimension of organizational commitment with credibility and validity that a factor toward an organization[58-60].

Joo[61] argued that organizational commitment is a critical factor in understanding individual work-related behaviors in a knowledge-based economy. He explained that an organization’s commitment is more stable and less sensitive to routine fluctuations than factors such as job satisfaction or job commitment, making it suitable for predicting employees’ behavior in a volatile modern business environment. In the same context, prior studies have demonstrated that organizational commitment is a factor predicting work-related outcomes such as employees’ work effort and performance, organizational citizenship behavior, workplace stress, and turnover intention[62][63].

Concerning knowledge workers, it should be noted that they have been recognized as having low organizational commitment[17]. The reason is that the high level of professional occupational commitment of knowledge workers causes conflicts with employers and nonprofessional colleagues, and this experience lowers their organizational commitment[18]. As a result, knowledge workers are more likely to leave the organization to resolve such conflicts. In the knowledge economy, knowledge workers’ turnover directly influences the weakening of a company’s competitiveness and increasing of cost[22]. Previous studies suggested innovation behavior, career fulfillment, role significance, worker–workplace fit, and knowledge management practices as antecedents that increase knowledge workers’ organizational commitment[64][65]. In addition, researchers such as Joo[61] and Issahaka and Lines[43]
argued that leadership is a major factor of knowledge workers' organizational commitment. The researchers emphasized that companies should increase knowledge workers' organizational commitment through appropriate leadership.

Accordingly, the authors predicted that data scientists' self-leadership would increase their organizational commitment. As mentioned earlier, self-leadership leads to the use of behavioral strategies that make given tasks and situations more positive or cognitive strategies that deviate from negative perspectives to maximize natural rewards[46]. Employees with high self-leadership would show more commitment to their organization because they focus on the positive aspects of the situation and the achievement of goals[48]. Sesen et al.[36] found that self-leadership increased organizational commitment in a quantitative study conducted on primary school teachers. In addition, researches conducted on various knowledge workers such as faculty members at universities, R&D researchers, bankers, nurses, registered individuals in sports organizations, and high school teachers confirmed that self-leadership affects the promotion of organizational commitment[66-69].

On the other hand, organizational commitment is a more significant factor than job commitment for predicting intention to leave[70]. Intention to leave refers to the extent to which a particular employee consciously considers plans to leave. Organizational commitment has proven to be the best predictor of intention to leave[70][71]. The reason is that organizational commitment is an attitude towards the organization, so its impact is more clearly reflected in their intention to leave the organization rather than job commitment. Employees who do not commit to the organization they belong to tend to solve problems on their own by leaving the organization[62]. On the other hand, employees with high organizational commitment continue to stay in the organization under favorable or unfavorable circumstances and make every effort to protect the organization's assets[72].

Prior researches have demonstrated that organizational commitment mediates the effects of leadership. Almutairi[73] found that transformational leadership influences job performance through organizational commitment. Jabbar et al.[74] also showed that organizational commitment mediates the relationship between transformational leadership and job satisfaction. Yousef[75] proved that leadership behavior improves job performance through organizational commitment. He explained the reason is that employees recognize their superiors by participating in consultative leadership behavior and then show higher performance by becoming more committed to their organizations. In the same vein, Hulpia et al.[76] also showed that when leadership is decentralized, people have a higher organizational commitment, which leads to more positive work outcomes. We could infer the mediating effect of organizational commitment will be the same in self-leadership. As discussed earlier, knowledge workers generally have low organizational commitment, which leads to a high intention to leave[77]. Therefore, the author here hypothesizes that the data scientist's organizational commitment is a key factor influencing intention to leave, and confirms whether data scientists' high self-leadership makes them commit more to the
organization and, as a result, lowers intention to leave.

**H2**: Organizational commitment will mediate the relationship between self-leadership and intention to leave.

### 3. Moderating Role of Perceived Job Autonomy

According to TAT, a situation, event, or intervention activates an individual's specific trait[37]. That is, individual characteristics are activated when a trait-relevant situational cue is given. The sources of this cue are divided into task, social, and organizational levels, each acting as a moderation variable.

Prior studies based on TAT proved that the trait-relevant situational cue moderates the effect of leadership. Luria et al.[78] revealed that an organization's level of centralized structure moderates the relationship between leadership attributes and leadership potential in a study conducted on combat soldiers. Phaneuf et al.'s[79] survey of 643 followers and 89 leaders showed that the arise of transformation leadership was moderated by the organizational context. Researchers have shown that the leadership effectiveness of transformational leadership is moderated by the perceived dynamic work environment[80]. Consistent with these findings, it has been demonstrated that the effectiveness of self-leadership is also moderated by perceived job autonomy[81]. As Ho and Nesbit's[81] research results of 407 employees in Hong Kong and China, both self-leadership effects on supervisor performance rating and objective performance measure were significantly moderated by job autonomy. Therefore, the author inferred that the data scientists' perceived job autonomy would moderate the effect of self-leadership.

In the natural rewards of self-leadership, crucial component is self-determination which refers to fulfilling the basic psychological needs of human beings' autonomy[45]. Autonomy is a core dimension of job characteristics and can be defined as "the degree to which the job manager allows freedom, independence, and discretion in determining the work schedule and work method necessary for performing the job"[82]. According to self-determination theory, environmental support that strengthens autonomy promotes effective self-regulation[83]. Therefore, the author inferred that self-leadership promotes desirable behavior (e.g., organizational commitment) in a situation where job autonomy is given. In a similar context, previous studies proved that job autonomy moderates the relationship between employees' proactivity and job performance[15]. Accordingly, it is hypothesized that the influence of self-leadership on organizational commitment will be moderated by perceived job autonomy.

**H3**: Perceived job autonomy will moderate relationship between self-leadership and organizational commitment.

In this study, the author further proposed a moderated-mediation model in which job autonomy moderates the mediating effect of organizational commitment. This model is based on the job demand–resource model(JD–R)[84]. According to the JD–R, when job resources corresponding to job demands are provided, negative influences decrease and positive influences are strengthened[85][86]. The moderating effect, according to the response of these job demands and job resources, is called the matching hypothesis[87].
Xanthopoulou et al. expanded the JD-R model to demonstrate that the response of personal resources and job resources influences work engagement. Shin and Jeung also showed that from the JD-R model’s perspective, the active personality (a personal resource) responds to job autonomy (a work resource), increases work engagement, and lowers intention to leave. Here, the personal resource is “positive self-assessment related to the perception and resilience that individuals can successfully control their environment.” This concept is very similar to self-leadership, a cognitive and behavioral strategy that seeks to control the environment, set goals, and reinforce intrinsic motivation; in addition, self-leadership is strongly related to personal resources such as self-efficacy.

Job autonomy is a representative job resource and is expected to act as an essential resource for the exercise of self-leadership. Therefore, when job autonomy as a job resource is high, self-leadership is strongly exerted and maximizes the effect of lowering intention to leave through organizational commitment. By synthesizing these facts, the author presents the moderated-mediation hypothesis as follows (Figure 1).

**H4:** Perceived job autonomy will moderate the effect of self-leadership on intention to leave through organizational commitment.

### III. Research Methods

#### 1. Characteristics of Sample

Research data were collected through an online survey of data scientists belonging to related departments, i.e., the data science team and big data team in large Korean companies. The author surveyed data scientists from affiliates through executives and team leaders of the companies taking the Big Data Executive MBA (Master of Business Administration). Data scientists belonged to 14 companies: 5 manufacturing, 3 IT, 3 distribution, 2 finance, and 1 telecommunications industry. The survey was conducted for two weeks in July 2020. A total of 221 data scientists participated in the questionnaire, among them, 203 responses were used for analysis, excluding incomplete responses.

The characteristics of the sample are as follows. The proportion of males in the sample was remarkably high, at 84.23%. In terms of age, 30 to 40 years old (51.23%) accounted for the greatest proportion, followed by 20 to 30 years old (38.91%). Above all, the proportion of masters (55.17%) and doctoral (27.58%) degree holders were remarkably high. Respondents’ work experience was the highest at 51.23% in the more than one year and less than five years category. A total of 10.70% answered that they had more than five years and less than ten years.

![Figure 1. Research Model](image-url)
years of work experience. The survey confirmed that data scientists’ work experience is relatively short.

2. Measurement Validity and Reliability

The author adopted measurement items whose reliability and validity were verified in several preceding studies. All variables were measured on a 5-point Likert scale (1 = strongly disagree; 5 = strongly agree).

Nine self-leadership measurement items used in the study of Houghton and Dawley[51] were adopted. Organizational commitment was measured by Fu et al.’s[89] four affective commitment items. For the measurement of intention to leave, four items were introduced from Bluedorn’s[90] research. Additionally, the author adopted age, gender, education, and work experience as control variables which are widely used in leadership studies[91]: age (between 20 and 30 = 1; 30 to 40 = 2; 40 to 50 = 3; more than 50 = 4), gender (male = 0; female = 1), highest degree held (bachelors = 1; masters = 2, doctors = 3), work experience (less than 1 year = 1; 1 – 5 years = 2; 5 – 10 years = 3; 10 – 15 years = 4; more than 15 years = 5).

Prior to analysis, the variables’ reliability and validity were verified through confirmatory factor analysis (CFA) with R’s ‘semTools’, and ‘lavaan’ packages. The measurement model has $\chi^2(98) = 279.728$, comparative fit index (CFI) = 0.918, Tucker–Lewis index (TLI) = 0.900, Bollen’s incremental fit index (IFI) = 0.919, and root mean square error of approximation (RMSEA) = 0.066, confirming that it has a desirable fit.

As a result of checking the internal consistency for reliability analysis, Cronbach’s $\alpha$ value for all latent variables was 0.863 or higher, exceeding the standard value of 0.70. The composite reliability (CR) value was also found to be at least 0.866 or higher, exceeding the standard value of 0.70. The reliability of the measure items is evaluated to be secured only when the standardized loading of all items exceeds 0.70. From the analysis results, the author confirmed that the factor loading $\lambda$ of all measurement items exceeded 0.640[92][Table 1].

<table>
<thead>
<tr>
<th>Construct</th>
<th>Items</th>
<th>$\lambda$</th>
<th>t</th>
<th>Cronbach’s $\alpha$</th>
<th>CR</th>
<th>AVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>SL1</td>
<td>0.810</td>
<td>13.833</td>
<td></td>
<td>0.958</td>
<td>0.959</td>
<td>0.725</td>
</tr>
<tr>
<td>SL2</td>
<td>0.781</td>
<td>12.963</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SL3</td>
<td>0.830</td>
<td>14.137</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SL4</td>
<td>0.908</td>
<td>16.256</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SL5</td>
<td>0.826</td>
<td>14.037</td>
<td>0.958</td>
<td>0.959</td>
<td>0.725</td>
<td></td>
</tr>
<tr>
<td>SL6</td>
<td>0.766</td>
<td>12.624</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SL7</td>
<td>0.800</td>
<td>13.402</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SL8</td>
<td>0.929</td>
<td>16.899</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SL9</td>
<td>0.930</td>
<td>16.927</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JA1</td>
<td>0.872</td>
<td>17.030</td>
<td></td>
<td>0.912</td>
<td>0.909</td>
<td>0.717</td>
</tr>
<tr>
<td>JA2</td>
<td>0.925</td>
<td>18.365</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JA3</td>
<td>0.821</td>
<td>15.014</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JA4</td>
<td>0.772</td>
<td>13.552</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OC1</td>
<td>0.783</td>
<td>11.068</td>
<td></td>
<td>0.912</td>
<td>0.909</td>
<td>0.717</td>
</tr>
<tr>
<td>OC2</td>
<td>0.813</td>
<td>11.859</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OC3</td>
<td>0.748</td>
<td>10.832</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OC4</td>
<td>0.810</td>
<td>11.823</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IL1</td>
<td>0.886</td>
<td>14.261</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IL2</td>
<td>0.901</td>
<td>16.479</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IL3</td>
<td>0.640</td>
<td>10.112</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IL4</td>
<td>0.711</td>
<td>11.786</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: $\lambda$ is factor loading, CR is composite reliability, AVE is average variance extracted.

Next, as a result of verifying the validity, the average variance extracted (AVE) of the latent variables was 0.50 or higher with $p < 0.05$, confirming a convergent validity. In addition, the AVE square root values of the latent variables were found to be greater than the correlation coefficient with other constructs.
thus satisfying the discriminant validity criterion [Table 2]. Therefore, it was confirmed that all constructs have strong reliability and validity.

In this study, whether there is a systematic measurement error due to common method bias (CMB) that may occur in the self-report questionnaire was also validated. As a result, self-leadership—the variable with the highest explanatory power—accounts for 43.9% of the total variance and no dominant factor that explains 50% or more variance existed. Accordingly, it was confirmed that no CMB occurred.

### IV. Results

To verify the hypothesis, the authors sequentially performed mediation effect analysis, moderation effect analysis, and moderated-mediation effect analysis. For all analysis, hierarchical multiple regression was performed first, followed by a 95% bootstrapped confidence interval (95% CI, n = 1000) analysis to clarify the effect. Hayes[93] Process Macro Model No. 7 was also used to analyze the moderated-mediation effect. The analysis was performed by using the 'lavaan', 'semTools', 'semPlot', 'processR', and 'tidyverse' packages of the statistical analysis tool R.

**1. Mediation Effects of Organizational Commitment**

First, a hierarchical regression analysis was conducted to test the hypothesis that organizational commitment will mediate self-leadership’s influence on intention to leave[94]. Looking at the analysis results, all regression models are statistically significant and the explanatory power of the three-stage model, including independent, mediation, and control variables, is 24.8% [Table 3].

To verify the size and statistical significance of the mediating effect, the 95% confidence interval (CI) of the indirect effect was verified through the bootstrapping method. The direct effect of self-leadership on intention to leave is statistically significant because 0 is not included between the low (-0.481) and high (-0.196) values of the 95% bootstrapping confidence level. Therefore, hypothesis 1 is supported. The indirect effect through organizational commitment is also statistically significant (95% CI = [-0.206, -0.017], n = 1000) and the size is b = -0.098 [Table 4]. Therefore, hypothesis 2 is supported by confirming a partial mediating effect of organizational commitment.
2. Moderating Effect of Perceived Job Autonomy

Hypothesis 3 proposed that perceived job autonomy controls the effect of self-leadership on organizational commitment. Therefore, it was confirmed whether the influence of self-leadership on organizational commitment changes according to the level of perceived job autonomy. Hierarchical regression analysis shows that the effect of the interaction term of self-leadership and perceived job autonomy on organizational commitment is significant (b = -0.332, p < 0.01), so hypothesis 4 is supported (Table 5). In addition, the author performed a simple slope analysis (plotting simple slopes at +1SD of organizational commitment) to interpret the moderating effect. As a result, it is confirmed that as self-leadership increases, organizational commitment increases, and the width of the increase also increases when the perceived job autonomy is high [Figure 2].

Hypothesis 4 predicted that there would be a moderated-mediation effect in which self-leadership lowers intention to leave through organizational commitment, which is controlled by perceived job autonomy. Therefore, it was confirmed whether the indirect effect of self-leadership on lowering intention to leave when perceived job autonomy is high is significantly greater than when it is not. Analysis results show that the
indirect effect of self-leadership on intention to leave is significant ($b = -0.103, 95\% CI = [-0.291, -0.027]$). In addition, the indirect effect that self-leadership lowers intention to leave through organizational commitment is stronger and more significant when perceived job autonomy is high ($b = -0.143, 95\% CI = [-0.293, -0.034]$) (Table 6). Therefore, hypothesis 4 is supported.

Table 5. Moderating the Effect of Perceived Job Autonomy

<table>
<thead>
<tr>
<th>DV: Organizational Commitment</th>
<th>b</th>
<th>SE</th>
<th>t</th>
<th>DV: Intent to Leave</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control Variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.105</td>
<td>0.057</td>
<td>1.840</td>
<td>-0.025</td>
</tr>
<tr>
<td>Gender</td>
<td>0.031</td>
<td>0.101</td>
<td>0.302</td>
<td>-0.176</td>
</tr>
<tr>
<td>Education</td>
<td>0.026</td>
<td>0.057</td>
<td>0.449</td>
<td>-0.072</td>
</tr>
<tr>
<td>Experience</td>
<td>-0.032</td>
<td>0.038</td>
<td>-0.842</td>
<td>-0.033</td>
</tr>
<tr>
<td>(Independent Variables)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-Leadership</td>
<td>0.896</td>
<td>0.183</td>
<td>4.894</td>
<td>-0.337</td>
</tr>
<tr>
<td>Job Autonomy</td>
<td>0.958</td>
<td>0.384</td>
<td>2.498</td>
<td></td>
</tr>
<tr>
<td>Self-Leadership X Job Autonomy</td>
<td>-0.332</td>
<td>0.118</td>
<td>-2.815</td>
<td></td>
</tr>
<tr>
<td>Organizational Commitment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: *p < 0.05, ** p < 0.01, *** p < 0.001. b is unstandardized coefficient. SE is standard error. DV is dependent variable.

Table 6. Conditional Indirect Effect of Self-Leadership on Intention to Leave through Organizational Commitment with Different Levels of Perceived Job Autonomy

<table>
<thead>
<tr>
<th>Index of Moderated Mediation</th>
<th>Conditional Indirect Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Index</td>
<td>Levels</td>
</tr>
<tr>
<td>-0.103</td>
<td>+1SD</td>
</tr>
<tr>
<td></td>
<td>-1SD</td>
</tr>
</tbody>
</table>

Note: b is unstandardized coefficients. CI = 95% confidence interval.
V. Discussion

1. Theoretical Implication

This study has a theoretical contribution to prove that the influence of knowledge workers' self-leadership on work-related outcomes applies equally to data scientists. Data scientists' self-leadership lowers intention to leave and raises organizational commitment, which is consistent with previous self-leadership researches[13][39][95].

But, the findings in this study have distinctive meaning in that it focuses on the influence of knowledge workers' leadership style on intention to leave. According to a knowledge-based view, knowledge is the essential asset in determining a company's competitiveness[96]. Moreover, knowledge workers play a crucial role in creating and effectively utilizing corporate knowledge. Therefore, the turnover of knowledge workers such as data scientists directly leads to weakening corporate competitiveness. However, prior researches on knowledge workers' intention to leave have been done in part in jobs such as healthcare professions[97], research scientists[98], and university personnel[99][100]. As far as the author knows, this is the first study to empirically analyze a data scientists' intention to leave who emerged as a new knowledge worker in the knowledge economy era.

Of particular note is that the knowledge possessed by data scientists is tacit knowledge with low transferability. Unlike explicit knowledge, which is easy to express and convey, tacit knowledge is embodied in the knowledge workers[101]. Therefore, acquiring tacit knowledge essentially needs experience and takes a long time [102]. Scholars considered tacit knowledge as a key to sustainable competitive advantage[103]. In that data scientists require business acumen or domain knowledge[104], the negative impact of turnover of data scientists on a company's competitiveness is more significant. So, the findings are meaningful because it broadened the theoretical understanding of the management of knowledge workers, which has emerged as a major topic in research related to knowledge workers since the 1990s in terms of improving corporate competitiveness[105].

Second, the author introduced self-leadership as the antecedent influencing the intention to leave of knowledge workers and empirically showed its effects. There are several previous studies on the effect of knowledge workers' leadership styles on individual and team performance[105]. Also, many scholars demonstrated the influence of knowledge workers' self-leadership on empowerment, commitment, productivity, trust, and team potency[4][106-108]. However, prior studies' limitation is that it did not focus on the higher turnover rates that arise because knowledge workers seek professional qualifications and less dependent upon a single organization[19]. For this reason, researchers such as Issahaka and Lines[43] pointed out that there is still a lack of theoretical understanding of the leadership effects of knowledge workers.

As pointed out in the studies of the knowledge-based view[96], the retention of knowledge workers is a vital factor for an organization's competitiveness[109]. However, as far as the authors know, there has not been a study that directly confirmed the effects of knowledge-workers' self-leadership on turnover
intention, even though the high turnover intention is a prominent characteristic of knowledge workers. Therefore, this study theoretically contributed that verified the influence on knowledge workers' leadership, especially on turnover intention, by demonstrating the underlying mechanism by which knowledge workers' self-leadership lowers intention to leave. Above all, the author conducted empirical analysis on data scientists, knowledge workers who have high perceived alternative job opportunities (PAJO). High PAJO increases the intention to leave of employees [110]. The fact that self-leadership lowers the intention to leave of data scientists with high PAJO verifies that the effect of self-leadership is as strong.

Third, the importance of self-leadership was presented for data scientist management, thereby suggesting the theoretical background for subsequent research. Data scientists' essential abilities are to self-directedly solves non-standardized problems and engage and persuade stakeholders [111]. Accordingly, effects of data scientists' self-leadership on individuals and organizations is significant. Considering that even traditional companies are actively attempting to create value from data [112], this study suggests the types of leadership required for data scientists, which are the critical success factors for leveraging data.

Lastly, based on TAT [37], the author proved that data scientists' self-leadership is further strengthened when perceived job autonomy is high. As consistent with and emphasized in previous studies, job autonomy is essential because knowledge workers have a low level of work standardization, and autonomous judgment and execution are more important than instruction and control [12]. The results of this study showed that when a data scientist recognized that job autonomy was satisfied, a trait of self-leadership was activated, thereby increasing organizational commitment. It was also demonstrated that the mediation path by which self-leadership lowers intention to leave through organizational commitment is also significantly controlled by the level of perceived job autonomy. Therefore, the results not only contributed to the generalization of the theory but also broadened the theoretical understanding for knowledge worker related research by showing that TAT was applied to data scientists, a representative knowledge worker in the 21st century.

2. Practical Implication

This study provides practical implications for hiring and fostering strategies for data scientists. First, when hiring a data scientist, self-leadership should be considered an essential requirement.

As a professional knowledge worker, the data scientist requires soft skills such as proactive and cooperative attitude, and communication skills [3]. In other words, it is necessary to hire talented data scientists who have a strong motivation to solve ambiguous real-world problems and have the ability to lead the process of problem-solving while managing complex environment and conflicts of interests.

In the data scientist hiring process, a company could introduce a quantitative measurement tool such as the revised self-leadership questionnaire (RSLQ) [113] and reflect this evaluation score. Besides, it is necessary to provide training to improve current data scientists' self-leadership in the
organization. Self-leadership is learnable competence[114]. In fact, it was proven that only six weeks of self-leadership training has a positive effect on performance by increasing employees’ self-efficacy[115]. Other scholars also demonstrated that self-leadership training increases organizational innovation and creativity[116]. Therefore, it is necessary to have systematic training programs for data scientists’ self-leadership, which has not been relatively noticed so far. Prior researchers emphasized the following five strategies for self-leadership training[115]: 1) natural reward strategies (increasing intrinsic motivation), 2) self-cueing (reminding of the goal), 3) self-talk, 4) visualizing successful performance, and 5) self-reward.

It is also necessary to differentiate the roles among data science team members. Few data scientists have everything from analytical knowledge and skills to business acumen, interpersonal skills, and communication skills. Therefore, it is possible to distinguish roles that focus on the hard skills of analysis and those focusing on soft skills such as collaboration, persuasion, and communication. Data scientist team managers can assign roles appropriate to the level of individual self-leadership. In fact, many data science organizations have separate roles that require more soft skills expertise such as business analysts[25].

Next, the findings demonstrated that perceived job autonomy is needed to maximize the effects of data scientists’ self-leadership. Therefore, companies should have management practices and cultures that provide a high level of job autonomy to data scientists. First, it is necessary to provide sufficient information about the organization’s strategy, vision, and values. Data scientists can make optimal decision independently when they clearly understand the context of the problems[116].

As for organizational structure, it is desirable to adopt a structure in which job autonomy can be maximized. In a mechanistic structure based on the effectiveness of control, self-leadership advantages are difficult to exert. Besides the opportunities to develop self-leadership would be limited. The evaluation system should encourage data scientists to lead their work with autonomy without the burden of failure by assessing not only the results but also the process[117].

Finally, corporate culture must be supported, in which each data scientist independently research, discuss, present, and implement ideas as experts. By maximizing data scientists’ autonomy, they will be highly motivated to lead business problem-solving, commit more to the organization, stay in the organization for a longer period, and exert their capabilities.

3. Limitation and Recommendations for Future Research

Although this study has suggested theoretical and practical implications, there are some limitations. First, the sample of the study is limited to data scientists working for large Korean companies. In Korean companies, knowledge workers showed low organizational commitment and high intention to leave, there is also an insufficient supply of data science talent[118]. Therefore, the author judged that the sample used in this study was appropriate. Even though Korean employees are likely to show a relatively higher level of organizational commitment and low intention to leave due to
cultural characteristics[119]. The level of perceived job autonomy may vary due to cultural characteristics[120]. Therefore, in future research, it is necessary to generalize the research results by verifying this research model in a more diverse cultural setting.

Next, the author only considered self-leadership and organizational commitment as factors influencing intention to leave. Although the two factors are widely recognized as the essential antecedents of knowledge workers’ intention to leave, introducing other factors (e.g., perceived alternative job opportunities, providing educational opportunities) that can explain the characteristics of the data scientist will provide richer implications. In addition, if future research is expanded by adding variables that evaluate organizational effectiveness in subsequent research, the influence of data scientists’ self-leadership on the organization can be more accurately revealed.

Finally, diversifying a moderating variable that activates self-leadership will be a meaningful direction for further research. For example, task significance in the job characteristic theory[121], organizational support, and the quality of the leader-member exchange relationship in the leader-member exchange theory[122] may be considered. These follow-up studies will broaden the understanding of data scientist self-leadership and give practical implications.

참고문헌


[29] T. Agasisti, and A. J. Bowers, Data Analytics and Decision Making in Education: Towards the Educational Data Scientist as a Key Actor


[65] L. Gažiauskienė and Z. Tunčkienė, "Organizational Level Factors of Knowledge


[99] M. Koc, “Research on Correlations between Academicians’ Levels of Organisational Commitment and Their Intention to Quit Their


