



Are We Culturally Fit? Designing a Human-AI Collaborative Interview System for Culture Fit Evaluation Between Startups and Employees

Seungyeon Song¹, Jeongeun Lim¹, Yejin Kim¹, Youjeong Nam¹, Yoori Koo^{2*}

¹Department of Interaction Design, Graduate School, Student, Hongik University, Seoul, Korea

²Department of Service Design, Graduate school of Industrial Art, Professor, Hongik University, Seoul, Korea

Abstract

Background Hiring failures in early-stage startups often stem from misalignment between candidates and organizational culture. Since startups lack structured evaluation criteria and sufficient interview experience, their hiring processes may easily become subjective, and it is hard to maintain consistent judgment standards due to rapidly shifting cultural identities and talent needs. To address these challenges, this study proposes a human-artificial intelligence (AI) collaborative interview system that enables effective culture-fit assessment even in environments with limited evaluative resources. The system supports hiring decisions by analyzing candidates' values, behaviors, and organizational compatibility in relation to the dynamic environments of startups.

Methods First, in-depth interviews with startup hiring practitioners were conducted to identify key needs in evaluating cultural fit. Second, the hiring process was analyzed to structure task-level behaviors to determine opportunities for AI collaboration. Third, a user survey was conducted to define the appropriate AI involvement level, and these findings were then applied to the system prototype design. With the prototype, Study 1 evaluated usability of the system, and Study 2 was a retrospective quantitative assessment in which participants compared past hiring experiences with prototype use.

Results While the automatic question generation feature reduced users' cognitive and time burden, clarity of AI reasoning showed negative effect on user trust. However, collaborative features like how humans can always edit the AI outputs led to high user satisfaction and perceived efficiency. Compared to users' previous experiences with culture-fit interviews, the suggested system demonstrated improvements in time efficiency, task efficiency, objectivity, and fairness.

Conclusions The study presents design guidelines for integrating AI into culture-fit interviews in a human-centered manner. By structuring evaluation criteria and enabling transparent collaboration, the proposed system reduces uncertainty in hiring decisions and supports startups in selecting culturally aligned candidates. Future research may examine long-term hiring outcomes and explore scalability across diverse organizational contexts.

Keywords AI-Assisted Evaluation, Human-Centered AI(HCAI), Person-Organization Fit, Organizational Culture Profile(OCP), Startup Hiring

*Corresponding author : Yoori Koo (yrkoo@hongik.ac.kr)

Citation : Song, S., Lim, J., Kim, Y., Nam, Y., & Koo, Y. (2026). Are We Culturally Fit? Designing a Human-AI Collaborative Interview System for Culture Fit Evaluation Between Startups and Employees. *Archives of Design Research*, 39(1), 245-282.

<http://dx.doi.org/10.15187/adr.2026.02.39.1.245>

Received : Sep. 02. 2025 ; **Reviewed** : Jan. 04. 2026 ; **Accepted** : Jan. 19. 2026

pISSN 1226-8046 **eISSN** 2288-2987

Copyright : This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>), which permits unrestricted educational and non-commercial use, provided the original work is properly cited.

1. 서론

1. 1. 연구 배경 및 목적

최근 채용 과정에서 단순히 개인의 직무 역량이나 경력 중심 평가를 넘어, 기업의 조직문화와의 적합성을 고려하는 채용 패러다임의 변화는 어렵지 않게 찾아볼 수 있다. 이는 입사만큼이나 중요한 조직 내 적응, 기존 팀원과의 원활한 협업을 통한 상호간의 동반 성장, 그리고 장기적인 근속 유지를 위한 중요한 평가 요소이기 때문이다(Leadership IQ, 2020). 실제로 채용 이후 조기 퇴사나 반복되는 이직은 기업에게 상당한 리스크로 작용하며, 그 원인으로는 조직과의 ‘핏(fit)’ 문제에서 비롯되는 경우가 많다.

이에 따라 많은 기업은 채용 과정 중 조직문화 적합성(Culture Fit)을 평가하기 위해 별도의 단계 및 인성검사 등을 도입하고 있다. 고용노동부(2024)에 따르면, 국내 주요 기업의 61%가 채용 과정에서 조직문화 적합성을 평가하고 있으며, 이는 조직 적응 속도와 협업의 질 향상, 퇴사율 감소 등에 긍정적 효과로 연결된다고 보고되고 있다.

이러한 조직문화평가에는 상당한 경험과 전문성을 갖춘 채용 전문가의 개입이 필요한데, 최근에는 이를 보완하기 위한 기술적 시도로 비동시적 AI 면접(Asynchronous Video Interview-AI, 이하 AVI-AI)과 같은 비언어적 행동 분석 기반 면접 평가 기술이 주목받고 있다. AVI-AI는 지원자의 답변은 물론 음성 및 표정, 말투 등 비언어적 신호를 분석하여 조직적합성을 평가하는 주요 기술로 주목받고 있다(Lee & Kim, 2021). 그러나 AVI-AI 기술도 여전히 알고리즘의 편향성, 평가 결과의 신뢰도와 타당도 부족, 그리고 인간적인 상호작용 맥락의 부재 등의 한계로 인해 실제 채용 현장에서는 1차 스크리닝 단계로만 한정적으로 활용된다. 이에 따라 인간과 AI가 협력하여 AI가 인간의 판단을 보조하는 ‘인간-AI협업’ 기반 채용 시스템이 필요한 실정이다.

한편, 이러한 기술 발전과는 별개로, 스타트업과 같은 소규모 조직은 상대적으로 체계적인 인적자원관리(HRM) 시스템이 부족하고, 채용 경험이 제한된 경우가 많아, 조직문화 적합성을 타당한 기준으로 판단하거나 구조화하여 반영하기 어렵다(Cardon & Stevens, 2004). 현존하는 대부분의 컬처핏(Culture Fit) 평가 시스템은 대기업 중심으로 개발되고, 각 기업의 특징에 맞게 커스터머되어 있으므로, 스타트업에 바로 적용하는 데는 한계가 존재한다(Coad et al., 2014). 이로 인해 단기적 관점에서 직무 역량 평가 중심으로만 채용이 이루어지며, 그 결과 반복되는 퇴사, 내부 갈등 및 조직문화 와해 등 기업의 경쟁력과 생존력을 위협하게 된다(Kristof-Brown et al., 2005).

이에 본 연구는 스타트업이 조직문화 적합성 평가의 부재로 인해 겪는 문제를 해결하기 위해, 조직문화 적합성 평가 과정을 처음부터 구조화하고, 인사 전문가의 부재 속에서도 실질적인 판단을 도울 수 있는 AI 기반 협업 시스템을 설계하는 것을 목표로 한다. 특히 조직의 핵심 가치와 협업 방식이 명확히 정의되지 않은 상태에서도, 내부 구성원의 언어·행동 데이터를 기반으로 조직 문화를 정의하고, 이를 기준으로 지원자의 적합도를 판단할 수 있는 서비스 시나리오를 고안하였다. 즉, 인사 전문 인력이 별도로 존재하지 않는 상황에서도, 기업은 스스로 조직문화 진단부터 인재 평가까지 이어지는 채용 여정 전반에서 AI가 판단 보조의 파트너로 작동할 수 있도록 인간-AI 협업 접근을 통해 이를 효과적으로 개선할 수 있는 가능성을 제안하고자 한다.

본 연구는 구체적으로 다음과 같은 세 가지의 목적을 갖는다. 첫째, 스타트업 및 소규모 기업의 조직문화 적합성 평가 현황과 그 과정에서 발생하는 문제를 진단하고 분석한다. 둘째, 스타트업 환경에 적합한 인간-AI 협업 수준별 조직문화 적합성 평가 면접 시나리오를 정리하고, 실행 가능성과 사용자 인식 기반의 타당성을 평가한다. 셋째, 인간-AI 협업 기반의 조직문화 적합성 평가 서비스 프로토타입을 설계하고, 이를 실증적으로 검증하여 스타트업 채용에 최적화된 서비스 모델과 운영 전략을 제안한다.

1. 2. 연구 범위 및 절차

본 연구는 설립 5년 미만의 국내 초기 스타트업 경험자를 대상으로 하였으며, 주 사용자로는 스타트업의 임원직 혹은 주요 직책자와 기업 인사 업무 담당자를 선정하였다. 연구의 주요 범위는 조직문화 적합성 평가와 AVI-AI를 활용한 인간-AI 협업 기반의 면접 서비스 설계이다. 본 연구의 대상인 스타트업은 특정 산업 도메인에 한정하지 않고, 다양한 분야(예: IT, 콘텐츠, 제조, 교육 등)를 포함하므로, 산업별 특성에 따른 전문성이나 도메인 특화 질문은 연구 범위에서 제외하였다. 대신, 스타트업 전반에서 공통적으로 적용 가능한 가치관, 협업 태도, 조직문화 적합성 관련 질문 요소를 중심으로 면접 구조를 정의하였다. 본 연구는 이러한 면접 보조 서비스를 통해 스타트업의 채용 실패를 예방하고 조직문화에 적합한 인재를 선발할 수 있도록, 조직문화 진단부터 면접 시나리오 설계, 프로토타입 제작 및 검증에 이르는 전 과정을 포괄적으로 수행하였다. 이를 위한 연구의 구체적인 절차는 Table 1과 같다.

Table 1 Flowchart of the Research

Step 1. 이론적 고찰	연구 분야 이해	인간-AI 협업	
	스타트업 채용 환경 및 조직문화 적합성 이론 검토	인간중심의 인간-AI 협업	
	AVI-AI 기술 이해 및 사례 분석	인간-AI 협업 방식 및 단계별 차이점	
▼			
Step 2. 인간-AI 협업 환경 및 사용자 조사	심층 인터뷰	경험맵 작성	
	스타트업 대표 및 인사업무 담당자 인터뷰로 실제 조직문화 적합성 평가에 대한 이해 및 니즈 파악	면접 단계별 페인 포인트 분석 및 인간-AI 협업 중심 개선 방향 도출	
▼			
Step 3. 인간-AI 협업 기회요인 탐색	인간-AI 협업 기회요인 분석	인간-AI 협업 시나리오 도출	시나리오 선호도 조사
	상황별 AI 협업 기회요인 분석	기회요인에 따른 구체적 서비스 사용 시나리오 작성	사용자 관점에서 인간-AI 협업 수준 선호도 조사
▼			
Step 4. 프로토타입 제작 및 검증	① 시나리오 기반 프로토타입 제작	② 프로토타입 평가 기준 수립	
	다빈도 선호 시나리오 기반 프로토타입 제작	사용성, 인간-AI 협업 적합성, 조직문화 적합성 판단 결과 적합성 등의 기준 수립	
	③ 연구 1. 사용자 경험 평가	④ 연구 2. 프로토타입 효과 검증	
	프로토타입 사용 후 평가자가 지각한 HCAI 기반의 경험 품질 평가	연구 대상자의 조직문화 적합성 평가 방식에서 기존 방식 vs. 프로토타입 사용 시 업무 효율·정확·부담 감소 수준 평가	
▼			
Step 5. 인간-AI 협업 시스템화 및 최종 설계 제안	시스템 맵	AI 협업 수준 기반 서비스 블루프린트 제안	
	To Be 이해관계자 중심의 시스템 맵 도식화	인간-AI 협업 기반 블루프린트 프레임워크 제안	

2. 이론적 고찰

본 장에서는 연구의 주요 대상이 되는 스타트업의 채용 환경에 대한 이해와 인간과 AI의 협업 단계에 대한 이해를 바탕으로 AI 서비스가 해결해야 할 주요 문제 상황을 도출하고자 한다.

2. 1. 채용 산업에 대한 이해

2. 1. 1. 스타트업 채용 환경의 특징과 문제점

스타트업은 일반적으로 조직 규모가 작고, 설립 초기 단계의 불확실성을 안고 있는 경우가 많다. 이로 인해 조직문화나 인재상에 대한 정의가 명확하지 않으며, 채용 절차가 표준화되어 있지 않아 채용 과정 중 담당자의 경험이나 직관에 크게 의존하는 경향을 보인다(Cardon & Stevens, 2004; Coad et al., 2014). 특히 초기

스타트업의 경우 창업자 본인이 직접 인재 채용을 수행하는 경우가 많고, 전담 인사 시스템이나 명확한 채용 기준 없이 즉흥적으로 인력을 충원하는 사례가 일반적이다(Cardon & Stevens, 2004; Leung, 2003). 이로 인해 사업적 시점마다 조직에 적합한 인재를 객관적으로 판단하고 선별하는 데 어려움을 겪는다.

또한, 스타트업은 설립 초기 특성상 낮은 고용 안정성, 불확실한 수익 구조, 부족한 복지 제도 등으로 인해 대기업 대비 경쟁력 있는 인재 유치에 한계가 있으며(Chung & Parker, 2023), 조직에 대한 외부의 신뢰도나 평판 또한 낮아 우수 인재가 선호하지 않는 경우가 많다. 즉, 공급 측면에서도 인재가 스타트업을 기피하는 경향이 존재하고, 수요 측면에서도 적절한 인재 선별 기준이 없어 인사 실패가 반복되는 문제가 나타난다(Cardon & Stevens, 2004; Rocha & Grilli, 2023). 이러한 구조적인 제약은 스타트업 채용의 비효율성을 심화시키고, 결과적으로 조직문화와 맞지 않는 인재 채용으로 이어질 수 있다.

CB Insights(2014)는 스타트업 실패의 주요 원인 중 하나로 ‘잘못된 팀 구성’을 지적하며, 구성원 간 문화적 불일치가 갈등, 협업 저하, 성과 하락으로 연결된다고 설명한다. 또한 Leadership IQ(2020)의 연구에 따르면 신규 채용 인력의 46%가 18개월 이내에 퇴사하거나 부적응을 경험하며, 이 중 89%는 기술 부족이 아닌 동기, 감성지능, 태도 등의 문제에서 비롯된다고 밝혔다. 이는 스타트업 채용에서도 동일하게 적용되며, 기술적 역량 중심의 평가에서 벗어나 태도와 조직문화 적합성을 평가할 수 있는 정성적 기준의 부재가 문제로 지적된다.

본 연구에서는 이러한 문제 인식을 바탕으로, 스타트업 채용에서 조직문화 적합성 평가가 왜 중요한지, 그리고 이를 어떻게 구조화하고 설계할 수 있는지를 중심으로 논의를 전개하고자 한다. 궁극적으로 스타트업의 채용 실패는 인력 교체 비용 증가, 내부 갈등 확산, 조직 생산성 저하, 장기 생존 가능성의 약화 등 심각한 부작용을 초래할 수 있다. 따라서 본 연구에서는 조직문화 적합성이라는 비가시적 요소를 효과적으로 진단하기 위해, 정량·정성 데이터를 통합하고 인간-AI 협업 구조에 기반한 평가 체계를 제안하고자 한다. 이는 인사 담당자가 보다 일관성 있고 신뢰도 높은 기준을 마련할 수 있도록 지원하는 동시에, 스타트업의 인재 선발에서 실질적인 오류를 줄이는 데 기여할 수 있을 것이다.

본 연구에서 스타트업의 채용 환경에 집중하여 문제점을 식별하고 개선 방안을 제안하고자 하는 이유는 기업의 규모와 성장 단계에 따라 걸쳐핏이 작동하는 방식은 본질적으로 다르기 때문이다. Table 2와 같이, 조직의 규모와 성장 단계에 따라 걸쳐핏이 어떻게 다르게 작동하는지를 살펴보고자 한다. 대기업은 일반적으로 상시 근로자 수 1,000명 이상 또는 매출액 1조원 이상 규모의 조직으로, 체계적 인사제도와 안정된 문화가 확립되어 있다. 중견기업은 300명 이상, 1,000명 미만의 인력을 보유하며, 성장 안정성과 변화 수용성을 동시에 추구하는 조직으로 정의된다. 반면, 스타트업은 100명 미만의 소규모 조직이거나 설립 7년 이하의 초기 기업으로, 빠른 의사결정과 실험적 조직문화가 특징이다(중소기업기본법 제2조; 중견기업 성장촉진 및 경쟁력 강화에 관한 특별법). 이러한 규모 구분은 조직문화의 성숙도, 구조의 경직성, 인사 시스템의 체계에 영향을 미치며, 각 기업군별로 ‘걸쳐핏’의 개념이 다르게 작동되는 배경을 설명하는 기준으로 활용된다.

Table 2 Comparative Analysis of Culture-Fit Assessment by Organizational Size

하위 요인	대기업	중견기업	스타트업	참고문헌
조직문화 특성	안정적이고 이미 확립된 가치 및 규범 중심	전통과 변화가 공존하며 점진적 문화 변화	가치와 규범이 유동적이며 빠르게 변화	Cardon & Stevens, 2004; Coad et al., 2014
채용 목적	안정성 유지 및 기존 문화와의 일치	조직문화 유지와 동시에 문화 수용	성장 잠재력 확보 및 문화 공동 창출	Chung & Parker, 2023
걸쳐핏 요소	보충적 적합성 (기존 문화와의 유사성)	보충적 + 보완적 적합성 병행	보완적 적합성 (새로운 시각과 역량)	Kristof-Brown et al., 2023
핵심 평가 기준	조직 적응성, 동화 능력	조화 능력, 유연성, 적응력	혁신성, 학습 민첩성, 가치 공동 정의 능력	Schneider et al., 2013
인재상 예시	“조직 시스템 안에서 효율적으로 일할 수 있는 사람”	“조직과 조화를 이루며 변화를 이끌 수 있는 사람”	“새로운 문화와 방향성을 만들어낼 수 있는 사람”	Kristof, 1996; Chatman, 1989

대기업은 이미 확립된 절차와 체계 안에서 개인과 조직의 유사한 가치를 의미하는 보충적 적합성 (supplementary fit)을 통해 안정적 성과를 추구하고, 중견기업은 기존 문화와의 조화 및 변화 대응 능력을 동시에 고려한다. 반면, 스타트업은 급변하는 시장 환경과 유동적인 조직 구조 속에서, 기존 문화에 없는 새로운 가치와 시각을 더해 조직을 재정의하고 확장할 수 있는 보완적 적합성(complementary fit)을 특히 중요시한다. 이뿐만 아니라, Table 3에 정의된 바와 같은 특징을 보인다. 이러한 차이는 조직의 성장 단계, 전략적 목표, 경쟁 환경에서 기인하며, 특히 스타트업에서는 단순히 조직에 적용하는 인재를 넘어 문화의 공동 창출자로서 팀의 방향성에 기여할 수 있는 인재가 핵심이라는 점에서 걸쳐핏의 의미가 확장된다. 이와 같은 인재상의 예시는 대표적인 조직문화 적합성(Person-Organization Fit, 이하 P-O Fit)의 구성 요소이며, 다음 장에서 구체적으로 정의된다

Table 3 Supplementary & Complementary Fit (Kristof, 1996)

유형	정의	조직 내 기대효과
보충적 적합성	- 개인과 조직이 유사한 가치, 성향, 태도를 지닌 경우를 의미함 - 개인이 조직 구성원과 쉽게 동화되어 일체감을 빠르게 형성할 수 있음	- 조직 응집력 향상 - 내부 갈등 감소 - 조직 안정성 확보
보완적 적합성	- 개인이 조직이 보유하고 있지 않던 자질 혹은 관점을 새롭게 제공하는 경우 - 개인과 조직 간의 상호 보완성을 형성할 수 있음	- 조직원 다양성 증진 - 혁신성 강화 - 유연한 문제 해결

2. 1. 2. 조직문화 적합성에 대한 이해

조직문화 적합성은 개인이 속한 조직의 가치, 신념, 행동 양식과 얼마나 일치하는지를 의미하며, 이는 개인의 직무 만족도, 조직 몰입, 성과, 이직 의도 등 다양한 변수에 영향을 미치는 핵심 개념이다(Kristof, 1996). 오랜 연구에 따르면 P-O Fit이 높을수록 구성원은 조직 내에서 더 높은 심리적 안정감과 성취감을 경험하여, 연속적으로 조직은 지속 가능한 성과를 창출할 수 있다(Chatman, 1989).

과거에는 이러한 적합성을 주로 가치 차원에서 이해했으나, 최근 연구는 이를 보다 다차원적인 개념으로 확장하고 있다. Cable & DeRue(2002)는 개인의 성장 욕구와 조직이 기대하는 행동 간 상호작용을 통해 P-O Fit이 형성된다고 보았으며, Kristof-Brown et al.(2023)은 원격근무, 글로벌 협업, 디지털 전환 등의 변화 속에서 조직문화 적합성이 단순한 가치 일치를 의미하는 가치 적합성을 넘어 행동 규범에서의 적합성과 적응 역량까지 포괄하는 개념으로 진화하고 있다고 강조하였다. Table 4는 이러한 구성 요소별로 적합성 평가 포인트를 정리하고 있다.

Table 4 Components of Person-Organization Fit

구성 요소	설명	적합성 평가 포인트	참고 문헌
가치 적합성	개인의 가치관과 조직의 핵심 가치-철학의 일치 정도	조직 미션·비전에 대한 공감 및 개인 목표와의 연계성	Chatman, 1989; Kristof, 1996; Kristof-Brown et al., 2023
행동 규범 적합성	조직의 규칙, 업무 수행 방식, 의사소통 스타일과의 부합성	업무 스타일, 협업 방식, 피드백 수용성	Cable & DeRue, 2002; Schneider et al., 2013
적응 및 성장 적합성	환경 변화 속에서 개인이 성장·적응하며 조직과 함께 발전할 수 있는 역량	학습 민첩성, 성장 동기, 자기 주도성	Tang et al., 2021; Kristof-Brown et al., 2023

또한, 최근 HR·조직 연구에 따르면 AI 및 데이터 기반 채용 환경에서 P-O Fit의 중요성이 더욱 부각되고 있음을 알 수 있다. 특히 AI기술이 조직문화와 개인 특성 간의 복잡한 패턴을 정량적으로 분석할 수 있게 됨에 따라(Lee & Kim, 2021), P-O Fit은 과거의 단발적 일치 여부를 측정했던 정적 상태를 넘어, 환경 변화에 대응하여 개인과 조직이 상호 조율해 동적인 과정의 적응적 상태의 개념으로 이해되고 있다(Kristof-Brown et al., 2023; Chen et al., 2024). 즉, 조직 가치에 당장 완벽히 부합하는 사람만을 찾는 것이 아니라, 향후 변화에 맞춰 성장·학습하며 협업할 수 있는 잠재력을 가진 인재를 평가하는 방향으로 이동하고 있음을 의미한다. 이러한 적합성은 단순 성향의 유사성을 넘어서 직무 만족도, 조직 몰입도, 이직률 등 구성원의 태도, 정서, 행동

양식 등에 영향을 미치는 조직행동 변수로 작용되며, Kristof-Brown et al.(2005)의 메타 분석 결과에 따르면, P-O Fit이 직무 만족과 조직 몰입의 증가, 이직 의도 및 이직률 감소와 유의한 상관관계를 보임은 이미 확인된 바가 있다.

이에 따라 전통적인 P-O Fit 측정 도구인 Organizational Culture Profile(OCP)(O'Reilly et al., 1991)도 진화하고 있다. 정성적 가치 평가를 넘어, 구조화된 인터뷰, 행동 기반 질문, 정량적 평가, 시뮬레이션 기반 모델 등이 활용되며, 지원자의 가치관뿐만 아니라 협업 역량, 행동 패턴, 학습 민첩성까지 종합적으로 고려하는 평가 방식이 강조된다(Schneider et al., 2013).

Table 5 Structured Interview Questions and Evaluation Criteria Based on P-O Fit

질문 유형	평가 목적	평가 방식
가치관 질문	지원자의 내면적 가치관 및 조직철학과의 적합성 판단	신념, 태도, 선택 기준에 대한 답변
행동 기반 질문 (BEI)	과거 상황에서의 실제 행동 및 문제 해결 방식 파악	구체적인 사례 기반 진술로, 진술 내 행동 맥락과 결과 분석
상황 판단 질문 (SJT)	가상의 상황 속 조직 문화 맥락에 맞는 대응을 전략적으로 탐색하며, 개인의 의사결정 기준, 우선순위 선정 기준 평가	특정 시나리오 제시 후 우선 대응 선택 이유 설명 분석

이러한 구조화된 평가 프레임워크는 채용 과정을 단순히 '현재의 적합성'을 근거하여 판단하는 절차에서, '미래 협업 가능성과 성장성'을 진단하는 전략적 판단 과정으로 전환한다는 점에서 의의가 있다. 특히 AI 기반 채용 시스템에서는 Table 5의 질문 유형과 같은 평가 항목이 데이터 입력 요소로 활용되어, 지원자의 가치관, 행동 패턴, 학습 가능성 등을 종합적으로 분석할 수 있다.

결국 P-O Fit은 조직과 개인 간 가치 일치 여부만이 아니라, 조직문화 이해도, 행동 규범의 적응력, 미래 성장 가능성을 통합적으로 판단하는 개념이다. 따라서, 본 연구를 통해 제안되는 AI 면접 시스템의 프로토타입 설계 역시 이 세 가지 요소를 기반으로 하여, 조직의 전략적 목표와 개인의 성장 궤적이 장기적으로 일치할 수 있도록 하는 평가 모델을 구축하고자 한다.

2. 2. AI기술 활용 사례와 한계점 분석

최근 채용 분야에서는 비동시적 AI 면접(Asynchronous Video Interview-AI, 이하 AVI-AI)기술이 활발히 사용되고 있다. AVI-AI는 지원자가 사전에 녹화된 면접 질문에 영상으로 응답하고, 이를 AI가 분석하여 평가하는 방식이다. 이 방식은 시간과 장소의 제약 없이 일관적이고 효율적인 평가가 가능하며, 면접관의 피로도나 응답자의 긴장도 같은 외부 변수의 영향을 최소화하여 공정한 평가 환경을 조성하는 데 기여한다는 장점이 있다(Moon & Nam, 2022).

AVI-AI에 활용되는 대표적인 기술로는 자연어 처리(NLP), 표정 분석, 음성 분석, 생리적 신호 분석 등이 있으며, 이는 지원자의 언어적, 비언어적 표현을 종합적으로 분석하여 정서, 인지, 성격, 문화적합도 등의 다양한 요소를 추론하는 데 사용된다. 예를 들어, 자연어 처리는 지원자의 문장 구조, 어휘 사용, 논리성을 분석하여 의사소통 역량이나 태도, 성격 특성을 간접적으로 파악하고(Agarwal et al., 2020), 표정 분석은 미세 근육 움직임을 통해 감정 상태나 스트레스 수준을 실시간으로 추정한다(Liu et al., 2024). 또한, 음성 분석은 말의 속도, 억양, 떨림 등을 기반으로 자신감이나 감정 안정성 등을 평가한다. 이외에도 영상에서 파악되는 생리적 신호를 통해 신체적 긴장도 같은 생리적 반응성을 객관적으로 측정하는 등 AI는 이러한 복합 데이터를 정량화하여 평가 지표로 활용할 수 있게끔 설계된다(Lee & Kim, 2021).

이처럼 AVI-AI는 다차원적 평가 지표 구성을 가능하게 하며, 기존 대면 면접에서 간과되기 쉬운 정서적·비언어적 정보를 분석에 포함함으로써 보다 정밀하고 종합적인 지원자 이해를 가능하게 한다. 감정 표현, 제스처, 시선 처리 같은 정서적 행동 신호, 어휘 다양성과 문장 명료성 같은 언어적 신호, 응답의 논리성과

문제 해결력 등 인지적 신호, 그리고 외향성·성실성 같은 성격 요소를 기반으로 지원자의 직무 역량과 조직문화 적합성을 간접적으로 평가할 수 있다(Lee & Kim, 2021). 본 연구는 이러한 AI 기반 면접 기술이 실제로 조직문화 적합성 판단에 어떤 방식으로 기여할 수 있는지를 분석하고자 한다.

실제 사례로는 글로벌 기업 유니레버(Unilever)가 HireVue 플랫폼을 활용해 AVI-AI를 채용 전형에 적용한 것이 대표적이며, 해당 기업은 지원자의 비언어적 신호 분석을 통해 커뮤니케이션 태도, 자기주도성 같은 평가 요소를 정량화하여 일관된 기준으로 활용하고 있다(Hu, 2023). 국내에서는 마이다스아이티(inAIR)나 제네시스랩(뷰인터) 등이 상용 AVI-AI 솔루션을 도입하여 채용의 효율성과 객관성을 높이고 있으며, 특히 다수의 지원자를 동시에 처리해야 하는 대규모 공채 전형에서 높은 활용 가치를 보인다.

그러나 AVI-AI에는 기술적, 윤리적 한계도 존재한다. 대표적으로 알고리즘의 학습 데이터가 편향될 경우 특정 인종, 성별, 언어 사용 스타일에 불리하게 작용할 수 있으며, 이는 인공지능이 의도치 않게 차별적 결과를 초래하는 위험성으로 이어질 수 있다(Köchling & Wehner, 2020). 또한, 지원자 입장에서는 인간이 아닌 AI가 자신의 표정, 말투, 시선 등을 평가하는 상황에 심리적 위화감을 느끼거나, 공정한 평가에 대한 신뢰 부족으로 이어질 가능성도 존재한다. 무엇보다도 현재의 AVI-AI 기술은 인간의 말과 행동에 담긴 복잡한 맥락, 의도, 문화적 배경을 완전히 이해하거나 해석하는 데 있어 기계적 한계를 가진다(Langer et al., 2020).

이러한 점에서 AVI-AI는 완전한 자동화 수단이 아닌, 인간 평가자와 협력하는 보완적 도구로서의 활용이 바람직하다는 인식이 확산되고 있다. 즉, 인간-AI 협업 관점에서 인공지능은 방대한 데이터를 기반으로 한 분석과 정량화 도구로 기능하고, 인간은 정성적 판단과 맥락적 해석을 담당함으로써 보다 신뢰도 높은 의사결정을 가능하게 하는 것이다(Kolbjørnsrud, 2024). 본 연구에서는 특히 조직문화 적합성과 같은 정성적 평가 항목에 대해, 인간 평가자의 판단을 중심에 두면서 인공지능의 분석 결과를 어떻게 보조적 도구로 통합할 수 있을지를 중점적으로 살펴보고자 한다.

2. 3. 인간중심 AI와 인간-AI협업 수준의 개념

인공지능의 기술이 다양한 산업과 사회 전반에 확산됨에 따라, 기존의 성능과 효율 중심의 기술 패러다임은 점차 한계를 드러내고 있다(Crawford, 2021). 이러한 변화 속에서 주목받는 인간중심 AI와 인간-AI협업 수준에 대한 이해에 대한 기본적 이해는 다음과 같다.

2. 3. 1. 인간중심 AI에 대한 이해

인간중심 AI(Human-Centered AI, 이하 HCAI)는 기존의 기술 효율성 중심 AI 개발 방향에서 벗어나, 인간의 권리와 가치를 중심에 둔 설계를 추구하는 패러다임이다. 이는 단순히 인간을 대체하는 기술이 아니라, 인간의 능력과 책임감을 강화하고 인간 중심의 경험을 증진시키는 기술로서 기능해야 한다고 강조한다(Shneiderman, 2020).

HCAI는 크게 두 가지 관점에서 설명된다. 첫째, 프로세스 관점에서는 사용자 관찰, 반복적 개선, 사용성 평가 등의 사용자 중심 설계 원칙을 통해 AI 시스템을 개발한다(Shneiderman, 2020). 둘째, 제품 관점에서는 자동화 기능을 갖추되, 사용자가 통제권을 유지할 수 있도록 설계된 ‘슈퍼툴(super tool)’로서 AI를 지향한다. 디지털 카메라나 내비게이션 시스템처럼 사용자의 창의성과 효율을 동시에 지원하는 기술이 그 예시이다.

이러한 HCAI를 바라보는 관점을 토대로, 핵심 목표는 세 가지로 요약된다. 첫째, 인간 중심 가치 실현을 통해 자기효능감, 창의성, 책임감, 사회적 연결성을 증진하는 기술을 설계하는 것이며, 둘째, 기술을 통한 사회적 선(善) 추구로서 AI4Good, 몬트리올 선언과 같은 사회적 실천과의 연계를 지향한다. 마지막으로, 셋째, 자동화와 인간 통제의 균형을 통해 높은 수준의 자동화와 높은 수준의 인간 주도성을 동시에 실현하는 설계를 추구한다.

2. 3. 2. 인간-AI협업 수준의 개념

인공지능의 기술 발전이 가속화되면서 다양한 산업과 사회 전반에 걸쳐 AI 기술이 적용되고 있다. 이에 따라 기존의 성능 및 효율성 중심의 기술 지향적 접근은 점차 한계를 드러내고 있으며(Crawford, 2021), AI 기술이 인간의 가치와 사회적 책임을 고려한 방향으로 설계되어야 한다는 목소리가 커지고 있다. 이러한 맥락에서 등장한 것이 인간중심 AI 개념이다.

대표적인 사례로는 자율주행 분야의 인간과 기술 간의 책임 및 역할 분담을 단계적으로 구분한 표준화된 프레임워크가 될 수 있다. SAE International(2021)에서 제시한 자율주행 레벨 분류는 시스템 자동화 수준과 인간 개입 정도를 기준으로, 인간과 기술 간의 역할 및 책임 범위를 명확히 정의한다. 구체적으로는, Level 0(No Automation)부터 Level 5(Full Automation)까지 총 6단계로 나누어 운전이라는 과업에서 인간과 시스템의 역할 정도를 구분한다. Level 0의 경우 모든 운전 작업을 인간이 수행하고 시스템은 단순한 경고나 정보 제공 기능만 수행하는 반면, Level 5에서는 모든 조건과 환경에서 시스템이 완전한 자율주행을 수행하며 인간의 개입이 전혀 필요하지 않다. 중간 단계에 해당되는 Level 1과 2에서는 조향 또는 가감속과 같은 특정 물리적 제어를 부분적으로 자동화 하되, 운전자의 지속적인 모니터링과 개입 대기 상태는 필수이다. Level 3과 4에서는 특정 조건에서는 시스템이 주행을 전담하되 상황에 따라 인간의 제어권 인수가 요구되거나 제한된 조건에서만 완전 자율이 가능하다.

본 연구에서 제안하는 AI 협업 단계 또한 이러한 자율주행 프레임워크의 핵심 원리를 참고하여 채용의 맥락에 적용한 것이다. 자율주행 레벨이 “운전 상황에서의 통제권 이동”을 설명한다면, 본 연구의 6단계는 “채용 의사결정 과정에서의 통제권과 책임의 이동”을 설명하는 개념적 대응 관계를 갖는다. 그러나 자율주행 단계가 물리적 제어 중심으로 설계된 반면, 채용 과정은 평가 판단, 윤리적 개입, 결과 책임 등 비물리적이고 사회적 요소를 포함한다는 점에서 차이가 존재한다. 따라서 본 연구는 SAE J3016의 단계적 자동화 개념을 참조하되, 인간의 판단과 윤리적 책임이 중요한 채용 맥락에 맞추어 확장·재구성한 인간-AI 협업 프레임워크를 제안하게 된다.

본 연구의 주요 연구 주제가 되는 HCAI는 AI를 인간의 판단을 대체하는 주체가 아닌, 인간의 역량 증진 및 협업을 지원하는 도구로 간주한다. 이는 인간의 자율성과 책임을 중심에 두고 기술이 이를 보조하도록 설계하는 접근이며, HCAI는 ‘신뢰할 수 있고 책임 있는 방식으로 인간의 역량을 강화하는 AI 시스템’으로 정의된다(Shneiderman, 2020).

구체적인 개념은 단순히 기술을 인간 친화적으로 사용하는 수준을 넘어, AI 설계 및 운용 전반에 인간 가치, 사회적 맥락, 윤리적 책임을 반영하는 프레임워크로 발전하고 있다. 예를 들어, HCAI의 핵심을 ‘설명 가능한 AI(Explainable AI)’로 규정하며 사용자 신뢰와 투명한 소통 구조를 필수 조건으로 제시한 연구가 있다(Ehsan & Riedl, 2020). 또한 인간-AI 협업에서 인터페이스 설계와 피드백 루프의 중요성을 강조하며, 인간 참여를 적극 유도하는 설계 패턴의 필요성을 지적한 바 있다(Amershi et al., 2019). 이러한 연구는 전면적 자동화가 아닌, 인간-AI 간 협업 구조 설계를 중심으로 기술 발전의 방향을 ‘정확성과 속도’에서 ‘신뢰성과 협력’ 중심으로 전환하려는 시도로 이해할 수 있다. 본 연구에서도 이러한 HCAI의 관점을 바탕으로, AI기반 채용 시스템에서 인간과 AI간 협력 구조가 어떻게 설계되어야 하는지를 탐색하고자 한다.

2. 3. 3. HCAI 기반 인간-AI 협업 접근법

AI 시스템을 인간 중심으로 설계하려는 최근 흐름은 단순한 도구를 넘어, 인간과 AI가 상호 보완적으로 협업하는 구조에 대한 논의로 확장되고 있다. 다양한 연구자들은 HCAI 설계의 핵심 가치로 인간의 자율성 보장, 투명성, 윤리성, 안정성, 협업 중심성을 제시하며, 이러한 요소들은 기술이 인간의 판단을 대체하기보다 지원하고 강화하는 방향으로 기능해야 함을 강조한다(Shneiderman, 2020; Ehsan & Riedl, 2020; Amershi et al., 2019). 본 연구에서는 이러한 선행연구에서 제시된 원칙을 종합하여, 인간-AI 협업의 실질적 구현

방안을 모색하고자 한다. 특히 인간의 주도성과 AI 자동화 수준의 균형을 중심으로 협업 구조를 네 가지 단계로 구분한다. 이는 향상(Enhance), 증폭(Amplify), 보완(Augment), 강화(Empower) 유형으로 정리되는데, 각 유형은 인간과 AI의 상호작용 방식과 역할 분담 정도에 따라 구분된다.

먼저, ‘향상’ 유형은 인간의 주도성이 낮고 AI 자동화 수준이 높은 단계로, 주로 반복적인 업무를 AI가 자동화하여 처리함으로써 인간의 경험과 작업 효율성을 향상시키는 데 초점을 둔다. 둘째, ‘증폭’ 유형은 인간의 주도성이 중간 수준이고 AI 자동화 수준이 중간에서 높은 단계에 해당한다. AI가 인간의 인지 능력을 증폭시키고, 방대한 정보에서 필요한 내용을 추출 및 분석하여 인간의 의사결정을 돕는 역할을 수행하게 된다. 셋째, ‘보완’ 유형은 인간의 주도성이 높고 AI 자동화 수준이 중간 정도인 단계로, AI는 인간에게 필요한 정보 구조를 제공하거나 다양한 분석 결과를 제시하여 인간의 판단을 보완하여 최종적인 결정은 인간이 주도적으로 내리게 된다. 마지막으로 ‘강화’ 유형은 인간의 주도성이 매우 높고 AI 자동화 수준이 가장 낮은 단계이다. 인간이 전체적인 전략적 판단을 주도하고, AI는 그 과정에서 필요한 보조적인 정보나 실행 도구를 제공함으로써 인간의 역량을 최대한 강화할 수 있는 역할을 수행한다.

2. 2. 4. 인간-AI 협업 수준(Level 1~6)의 개념과 적용

AI 시스템 설계에서 인간과의 관계를 단순히 보조 또는 자동화 기능으로만 구분하는 것은 기술의 복잡성과 사회적 맥락을 충분히 반영하지 못한다. 이에 따라 Shneiderman(2020)을 비롯한 여러 연구에서는 AI의 자율성 수준과 인간 개입의 강도를 기준으로 양자 간 관계를 세분화하여 정의할 필요성을 제기해왔다. 특히 인간 중심 AI의 구현을 위한 설계 프레임워크에서는, 인간과 AI가 어떠한 방식으로 의사결정 주도권과 통제권을 공유하는가가 핵심 논점으로 제시된다. 이러한 관점에서, 본 연구는 앞서 사례로 언급된 자율주행의 자동화 레벨 분류 체계를 참고하여, 인간-AI 협업 수준을 AI의 자율성과 인간 개입 정도에 따라 여섯 단계로 체계화하여 Table 6에 정리했다.

본 연구에서 제안하는 Level 1은 인간이 전적으로 판단과 실행을 담당하며, AI가 전혀 개입하지 않는 단계로 자율주행 단계 중 ‘수동 운전’에 해당한다면, Level 2는 AI가 다양한 대안 세트를 제시하여 인간의 선택을 지원하는 수준으로, 자율주행의 ‘Driver Assistance’로 조향 또는 가감속을 지원하는 정도로 부합한다. 즉, AI는 방향 제시자로서 가능성을 넓히지만 최종 판단은 전적으로 인간에게 있다. Level 3은 AI가 여러 옵션 중 일부를 선별하고, 인간이 이를 승인하거나 거부하는 단계로 ‘조건부 자동화’에 해당한다. 즉, AI는 제한된 상황에서 판단을 주도하지만, 인간이 여전히 제어권을 회수할 수 있는 구조다. Level 4는 AI가 일정 수준의 자율성을 바탕으로 실제 실행을 수행하되, 사전적으로 인간의 승인(human-in-the-loop)이 필수인 단계로, 의료나 국방 등 고위험 분야에 적합하다. Level 5는 AI가 자율적으로 실행하되, 인간은 사후적으로 개입하여 필요 시 중단 혹은 결과를 수정할 수 있는 단계로, 자율주행의 ‘완전 자동화’ 직전 단계에 해당한다. 이 단계는 효율성과 위험이 공존하며, 인간은 AI의 결정과 실행을 취소하거나 중단시킬 수 있는 권한으로 통제권을 간접적으로 유지한다. 마지막 Level 6은 인간의 승인이나 개입 없이 AI가 독립적으로 판단과 실행을 수행하는 완전 자율화 단계로, 모든 책임과 결과는 시스템 내부 로직에 의해서만 결정된다.

이와 같은 6단계 모델은 AI 기술의 발전 수준뿐 아니라, 인간 개입의 시점(사전·사후), 통제 방식(승인·감시·거부), 그리고 책임 구조의 명확화를 함께 고려하여 인간 AI 협업 설계의 기준점을 제공한다. 특히 채용, 의료, 교육, 금융 등과 같이 인간적 판단과 신뢰가 핵심적인 영역에서는, 각 업무의 리스크와 책임 수준에 따라 적절한 협업 레벨을 사전에 정의하고 이에 맞는 인터페이스·운영 정책·감독 체계를 설계하는 것이 필수적이다.

Table 6 Summary of AI-Human Collaboration Levels

단계	AI 자율성 수준	AI 역할	인간-AI 관계	참고 문헌
Level 1	없음	Assistant	인간이 모든 의사결정 수행	Stryker (2025); Forbes (2019)
Level 2	낮음	Advisor	AI가 다양한 옵션을 우선순위 없이 제시하되, 선택과 실행은 인간이 결정	Sheridan & Verplank (1978); Stryker (2025)
Level 3	중간	Partner	AI가 대안의 폭을 좁혀 우선순위가 있는 추천을 제시하되, 인간이 최종 승인 혹은 거부로 판단	Berretta et al. (2023); Festor et al. (2021)
Level 4	중간~높음	Operator with Human-in-the-loop	인간의 승인 및 감시하에 AI가 특정 작업 제안 및 실행까지 진행	Sheridan & Verplank (1978); Stryker (2025)
Level 5	높음	Operator with Human-on-the-loop	AI가 주도적으로 작업을 전체적으로 수행하며, 인간은 제한된 개입 혹은 사후 거부 가능	Jeon (2024)
Level 6	매우 높음	Autonomous Agent	AI가 모든 의사결정과 실행을 독립적으로 수행하며, 인간의 개입은 거의 없거나 불가능	Stryker (2025); Forbes (2019)

2. 4. 연구 프레임

본 연구는 앞선 문헌조사를 통해 도출한 스타트업 채용의 주요 문제점과 인간-AI 협업 설계 원리를 바탕으로, HCAI(Shneiderman, 2020) 이론과 인간-AI 협업 유형 분류 체계를 적용하여 단계적 연구 프레임을 수립하였다. 먼저, 스타트업 채용 실패 요인과 조직문화 적합성 문제를 분석하여 연구의 문제의식과 이론적 기반을 확립하였고, 이후 인덱스 인터뷰를 통해 조직문화 적합성 평가 과정의 페인 포인트와 요구를 탐색하고, 이를 토대로 면접 전·중·후 단계에 해당하는 핵심 업무 과업 8가지를 도출하였다. 이후 도출된 과업을 HCAI 관점에서 네 가지 서비스 기회 영역으로 재구조화하고, AI의 개입 목적에 따른 네 가지 협업 방식(향상, 증폭, 보완, 강화)을 각 과업별로 연결함으로써 인간-AI 협업 모델의 개념적 가능성을 제시하였다. 이어서 협업 방식과 AI 개입 수준(Level 1~6)을 조합하여 다양한 시나리오를 설계하고, 설문조사를 통해 사용자가 선호하는 AI 개입 정도를 검증하여 최종 시나리오를 확정하였다. 마지막으로, 확정된 시나리오를 기반으로 프로토타입 제작 후 두 번의 연구를 거쳐 사용자 평가와 효과성 평가를 통해 제안된 협업 모델의 효과성과 실현 가능성을 검증하였다. 이러한 연구 프레임은 1장에서 제시한 Table 1의 연구 절차를 이론적 기반 위에서 구체화한 것으로, 이론적 고찰 → 기회요인 도출 → 프로토타입 검증으로 이어지는 연구의 전체 흐름을 체계적으로 설명한다. 이에 따라 다음은 먼저 사용자 인터뷰를 통해 서비스 제안을 위한 기회 요소를 도출하였다.

3. 인간-AI 협업 기회 요소 도출

앞선 문헌조사를 통해 확인한 스타트업 채용의 문제점과 인간-AI협업의 설계 원리를 토대로 실제 본 연구의 대상자인 스타트업에서의 채용 경험이 있는 임원직, 주요 직책자, 인사 관련 업무 담당자를 대상으로 인덱스 인터뷰 형식의 사용자 조사를 실시하였다. 조사를 통해 기존 채용 과정에서 특히 컬처핏 목적의 면접을 진행하며 겪은 조직문화 적합성 평가 과정의 어려움을 토대로 니즈를 분석하고, 이를 바탕으로 면접 컬처핏 평가 과정의 관점에서 면접 전·중·후 단계에서의 핵심 과업을 분류하였다. 이후 분류된 과업을 중심으로 새로 제안할 서비스의 기회 영역을 구조화하고, 각 서비스 기회 영역별로 AI의 협업 방식 역할을 앞서 언급한 향상, 증폭, 보완, 강화로 구분한다. 구분된 협업 방식별로 AI개입 수준을 의미하는 Level1~6까지 동일 시나리오를 모두 작성한 후, 사용자 선호도 조사를 진행하며 사용자가 선호하는 AI의 개입 정도를 검증하고 최종적으로 시나리오를 확정하였다. 이는 1장에 제시된 Table 1의 연구 절차를 이론적 기반 위에서 구체화한 것이며 본 장에서는 기회요인 도출을 위한 사용자 조사에 대해 먼저 알아본다.

3. 1. 사용자 조사 계획 수립

3. 1. 1. 사용자 조사 목표 및 방법

본 연구는 인간-AI 협업 기반 조직문화 적합성 면접 서비스를 설계함에 있어, 실질적인 서비스 기회 요소를 발굴하고 서비스 방향성을 구체화하기 위한 기초 작업으로 사용자 조사를 실시하였다. 사용자 조사의 주요 목적은 스타트업 채용 과정에서 조직문화 적합성 평가와 관련하여 실제로 경험되는 문제와 요구를 심층적으로 파악하고, 이에 대한 사용자들의 인식 및 AI 기반 채용 서비스에 대한 기대를 탐색하는 데 있다. 이를 통해 서비스 초기 설계 단계에서 사용자 중심의 요구를 반영하고, 인간-AI 협업이 효과적으로 작동할 수 있는 접점을 발견하고자 하였다.

조사 방법으로는 질적 탐색 연구의 일환으로 반구조화된 인덱스 인터뷰를 활용하였다. 이 방식은 스타트업의 채용 업무 담당자와 실무자의 실제 경험, 문제 인식, 니즈를 유연하게 탐색하기에 적합하다 판단되었다. 인터뷰는 대면 또는 온라인 화상회의(Zoom)를 통해 진행되었으며, 회당 약 40~60분 내외로 이루어졌다. 모든 인터뷰는 참여자의 사전 동의를 얻어 녹음되었다. 수집된 데이터는 주제 분석(thematic analysis)과 어퍼니티 다이어그램 기법으로 분석하였으며, 인터뷰에서 도출된 주요 의미 단위를 정리하고, 이를 기반으로 조직문화 적합성 면접에서 AI가 개입할 수 있는 실질적 협업 기회 요소를 정리하였다.

3. 1. 2. 조사 대상자 선정 및 진행

본 연구에서는 인간-AI 협업 기반 조직문화 적합성 면접 서비스의 기회 요소를 도출하기 위해, 스타트업 채용 과정에서 핵심적인 역할을 수행하는 이해관계자를 사용자 조사 대상으로 선정하였다. 특히 본 서비스의 주요 사용자군으로 정의되는 스타트업 내 HR 담당자를 중심으로 조사 대상을 설정하였다. 이는 연구 목적이 스타트업의 채용 실패를 예방하고자 하는 데 있으며, 실질적인 채용 판단과 조직문화 적합성 평가에 관여하는 실무자의 경험을 바탕으로 보다 현실적인 요구를 파악하고자 하였기 때문이다. 여기서 스타트업 HR 담당자는 인재 채용을 주도하고 면접 서비스에 직접적으로 관여하는 실무자 전반을 포함하는 개념으로, 인사 담당자뿐 아니라 팀 리더, 면접관, 채용 결정권자 등 채용 전 과정에 실질적으로 참여하는 인물을 모두 아우른다.

이러한 기준에 따라 인터뷰는 2025년 4월 14일부터 4월 24일까지 11일간 진행되었으며, 조직문화 적합성 평가 경험이 있는 스타트업 대표 6명과 실무자 2명을 대상으로 인당 45~60분가량의 인덱스 인터뷰를 실시하였다. 인터뷰 참여자의 상세 프로파일은 Table 7과 같다.

Table 7 Interview Participants

사용자	직책	기업 규모	창업 기간	업종
U1	스타트업 공동대표	30인 미만	7년 미만	라이브 커머스
U2	스타트업 대표	10인 미만	3년 미만	코스메틱
U3	스타트업 대표	10인 미만	7년 미만	IT
U4	스타트업 대표	10인 미만	3년 미만	의료기기
U5	스타트업 대표	20명 미만	7년 미만	시니어 사업
U6	스타트업 대표	10인 미만	5년 미만	광고대행업
U7	실무자	10인 미만	6년 미만	제조업
U8	실무자	30인 미만	7년 미만	연구개발

인터뷰 질문은 5가지 카테고리 내에서 구성된 30개 문항으로 진행되었으며, 각 문항은 다음과 같은 목적에 따라 설계되었다. (1)조직문화 적합성에 대한 인식과 실제 채용 평가 기준 파악, (2)AI 기반 채용 시스템에 대한 인식과 도입 조건 확인, (3)지원자-조직 간 적합성 평가에서 요구되는 소프트스킬 탐색, (4)감성지능, 리더십, 책임감 등 비인지적 역량의 평가 방식과 AI 보완 가능성 확인, (5)행동 기반 면접 및 비언어적 신호에 대한 판단 기준과 AI 적용 가능성 탐색.

3. 1. 3. 인터뷰 분석

수집된 인터뷰 내용은 질적 데이터를 구조화하여 의미 단위 간 상관관계를 파악하기 위해 주제 분석 기법을 활용하였다. 인터뷰 전자본에서 의미 있는 발화 중심으로 핵심 데이터 단위(Meaning Unit)를 도출하고, 이를 응축된 의미 단위(Condensed Meaning Unit), 하위 범주(Sub-category), 범주(Category)로 체계화하여 유사성 기반으로 개념을 정리하였다. 이 과정은 어피니티 다이어그램 기법과 병행하여 진행되었으며, 두 명의 연구자가 독립적으로 코딩을 수행한 뒤 교차 검토를 통해 최종 주제를 도출하였다.

Table 8에 정리된 분석 결과, 총 4개의 주요 주제와 11개의 하위 주제가 도출되었으며, 이는 추후 4가지 서비스 기회 요소 도출 시 모두 활용되었다. 특히 이 중 일부 범주는 스타트업 채용 과정에서 P-O Fit을 실천적으로 구현하기 위한 과정을 구체적으로 보여준다.

첫 번째 범주인 ‘채용 시 선호되는 인재상 정의’는 P-O Fit의 ‘Person’ 요소와 직결되며, 조직이 어떤 사람을 ‘맞는 사람’으로 판단하는지를 보여준다. 인터뷰에 따르면, 정서 기반 커뮤니케이션 능력, 책임감 있는 태도, 팀워크 중심의 협업 태도, 환경 적응력 등이 반복적으로 강조되었다. 이들은 감정에 휘둘리지 않으면서도 솔직하고 유연하게 소통하며, 조직 내 관계에서 조화롭게 행동할 수 있는 특성으로서, 조직문화에 자연스럽게 스며들 수 있는 사람에 대한 조직의 기대치를 구체적으로 보여준다. “못하는 건 못한다고, 안 되는 건 안 된다고 솔직히 말해야 한다,” “커뮤니케이션 안 되는 직원은 필요 없다” 등의 인터뷰 내용은 감정에 휘둘리지 않되, 솔직하면서도 유연한 소통을 중요시함을 보여준다. 또한 “태도나 비언어적 표현은 지식보다 더 중요할 수 있다”는 언급처럼, 조직은 단순 역량보다 태도와 협업 적응력을 더 중요하게 평가한다. 즉, 앞서 문헌조사를 통해 파악한 바와 같이 스타트업에서는 보완적 적합성이 가능한 직원을 선호함을 확인할 수 있었다.

두 번째 범주인 ‘조직문화 적합성 판단 기준’은 P-O Fit의 ‘Organization’ 측면을 구성하며, ‘맞는 사람’을 판단하기 위해선 조직 스스로가 먼저 ‘우리는 어떤 조직인가’를 명확히 정의할 수 있어야 한다는 전제를 반영한다. 인터뷰에서는 단순한 직무 적합성보다는 팀과 조직의 문화적 맥락에 잘 어울릴 수 있는 태도와 성향을 중시하며, 특히 팀워크와 소통을 주도적으로 수행할 수 있는 적응력 역량이 강조되었다. “직무 역량은 어느 정도는 보고 배우면 되는데, 태도는 그게 안 되면 조직에 오래 버티기 힘들다”라거나, “조직문화에 맞지 않으면 실력 있어도 팀워크가 계속 깨지더라”와 같은 응답은, 이러한 주장을 뒷받침한다. 또한, 이러한 적합성을 평가하기 위해서는 조직문화 진단이 선행되어야 하며, 평가 기준 역시 조직 특성과 연동되어 구조화되어야 한다는 필요성이 제기되었다. 또한, 태도에 대한 키워드가 다수 수집되었는데, 이는 P-O Fit의 근거가 되는 구조화 질문 중 가치관 질문에 대한 답변을 통해 확인해 볼 수 있는 내용임을 알 수 있다.

세 번째 범주인 ‘AI 평가에 대한 인식’은 P-O Fit 평가 과정에서 AI의 보완적 역할 가능성과 한계를 동시에 조명한다. 응답자들은 AI가 조직의 감정적 맥락이나 문화적 분위기를 정밀하게 해석하긴 어렵다는 점에 “지원자가 진짜 그런 사람이 맞는지 AI가 맥락 없이 판단하면 위험하다고 생각해요”와 같은 우려를 표하면서도, “AI가 표정, 말투 같은 비언어 정보까지 함께 분석해주면 오히려 사람보다 더 공정할 수도 있을 것 같아요”와 같이, AI가 정량화된 근거를 제공하는 보조 도구로 활용되기를 기대하는 의견도 다수 존재했다. 이는 AI가 전면적 평가자가 되기보다는, 사람의 판단을 보완하는 정량 분석 도구로 기능하는 방향이 현실적이라는 시사점을 제공한다.

마지막으로, ‘채용 평가 방식의 현황과 한계’ 범주는 현재 평가 방식이 직관에 크게 의존하고 있음을 드러낸다. “면접과 자기소개서에서는 열정을 보였지만, 입사 후에는 경력이 과장됐다고 느꼈다”, “포트폴리오가 본인 것인지 알기 어렵다”는 응답은 서류 기반 평가의 한계를 보여주는 대표적인 사례다. 이에 따라 “과거 행동을 실제로 드러내는 행동기반 질문에는 연기가 통하지 않는다”는 언급처럼, 행동 기반 질문을 통한 경험 검증의 필요성이 강조되었다. 한편 일부는 “비언어적 신호는 중요하지 않다”고 응답했지만, 전반적으로는 표현의

진위를 가리기 위한 다각적 평가 방식의 필요성에 공감이 형성되었다. 전반적으로 앞서 Table 5에 언급된 구조화 질문을 통해 Table 4에 언급된 P-O Fit의 주요 구성 요소인 가치 적합성, 행동 규범 적합성, 적응 및 성장 적합성에 대한 확신이 반드시 필요함을 시사한다.

Table 8 Qualitative Content Analysis

Category	Sub-Category	Meaning Unit
채용 시 선호하는 인재상 정의	정서 기반 커뮤니케이션 역량	조직은 상대의 정서를 인식하고 상황에 맞게 조율하며, 솔직하고 유연하게 소통하는 역량을 매우 중요시함
	조직에 적합한 태도	조직에 적합한 태도는 책임감 있는 역할 수행, 피드백 수용에 유연한 협업능력 등을 포함함
조직문화 적합성 판단 기준	우리 조직의 현재 문화는 어떤가	조직문화 평가 시 평가 기준이자 목표가 되는 자사의 문화 진단이 선행되어야 함
	지원자는 어떤 성향의 사람인가	행동 기반 질문과 비언어적 신호를 함께 분석하면 지원자의 태도와 역량을 더 입체적으로 평가할 수 있음
	조직문화에 잘 맞는 사람인지 평가	팀마다 요구하는 커뮤니케이션 방식과 집중의 정도가 다르기 때문에, 환경에 맞는 소통 태도와 몰입 역량이 중요함
AI 평가에 대한 인식	AI 평가에 대한 우려와 의심	AI가 사람의 감정이나 조직 맥락을 충분히 이해하기 어려워, 적합성 판단에 대한 신뢰와 공정성에 의문이 제기됨
	면접 중 AI의 역할에 대한 기대	AI가 표정, 말투, 반응 등을 분석하여 지원자의 진정성과 감정, 문제 해결능력 및 다양한 정성적 평가를 신뢰도 높게 서포트 해줄 것을 기대함
채용 평가 방식의 현황	서류 기반 평가의 한계	서류나 면접에 드러난 표현만으로는 과장 여부나 실무 적합성을 정확히 판단하기 어려운 한계가 존재함
	행동기반 평가와 비언어적 신호	행동 기반 질문과 비언어적 신호를 함께 분석하면 지원자의 태도와 역량을 더 입체적으로 평가할 수 있음

이렇게 도출된 4개의 주제와 그 하위 내용은, 인터뷰 분석을 통해 정리된 총 4가지 인사이트를 기반으로 다음과 같은 1차 기회 요소로 정리되었다. 1) 조직 구성원 기반 조직문화 진단 및 기업의 인재상 정의, 2) 지원자별 맞춤 질문 생성 및 면접 전략 자동 배포 시스템, 3) 실시간 답변 및 행동 분석을 통한 인재상과의 핏 정밀 검증, 4) 기업-지원자 간 최종 핏 통합 평가 및 피드백 자동화. 이 네 가지 기회 요소는 조직문화 적합성 평가가 단발성 판단이 아닌 면접 전-중-후 단계 전체에 걸쳐 연속적으로 작동해야 한다는 전제하에 설정되었으며, 전체 채용 프로세스 흐름은 여정 지도로 Figure 1과 같이 시각화되었다. 여정 지도는 사용자(채용 담당자)의 경험을 기반으로, 채용 과정의 상호작용과 문제 지점을 구조화하여 실질적인 서비스 개선 기회를 발견하기 위한 도구로 활용되었으며, 각 면접 단계별로 위의 기회 요소를 구현할 수 있는 핵심 과업 8개가 함께 도출되어 실질적인 서비스 설계의 출발점으로 정리되었다.

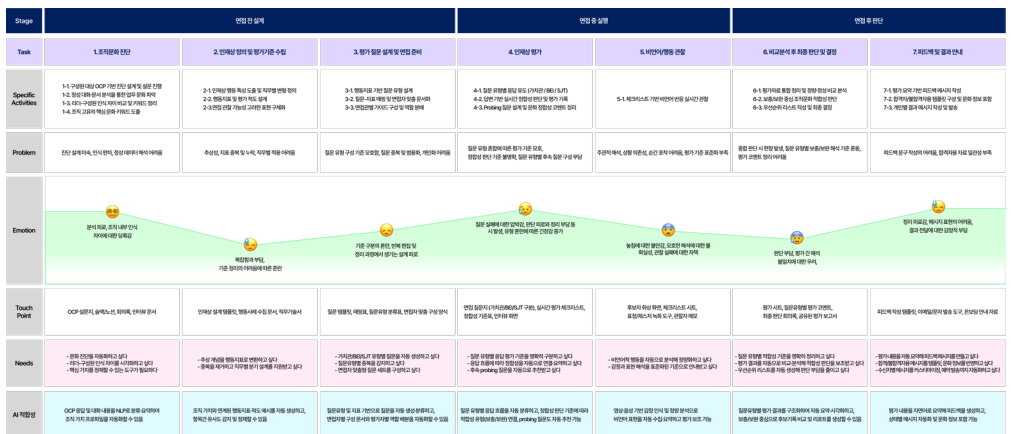


Figure 1 Experience map

3. 2. 인간-AI협업 기획요인 도출

3. 2. 1. AI 협업 적합성 분석

본 장에서는 조직문화 기반 인재 채용의 면접 과정에서 도출된 8가지 과업을 중심으로, 각 과업의 AI 협업 적합성과 수행 과정에서 나타난 페인포인트 및 니즈를 종합하여 Table 9에 정리하였다. 면접 전 설계 단계에서는 조직문화 진단, 인재상 정의 및 평가 기준 수립, 평가 질문 설계 등의 업무에서 AI가 효과적으로 활용될 수 있다. 특히 OCP(Organizational Culture Profile) 응답 데이터를 기반으로 조직 고유의 가치와 협업 성향을 분석하고, 이를 통해 조직의 핵심 가치를 정량적으로 도출할 수 있다. 이렇게 도출된 가치에 따라, 이를 평가할 수 있는 적합한 행동지표 기반의 면접 질문을 자동으로 생성하고, 질문 유형별로 평가 적합성을 판단하는 기준을 설정하는 데에도 AI의 역할이 가능하다.

면접 중 실행 단계에서는 AI와의 협업을 통해 면접자 답변을 실시간으로 분석하여 질문 유형에 따라 응답을 자동 분류하고, 후속 질문 자동 생성 및 추천하는 기능이 가능하다. 또한, 영상 및 음성 데이터를 활용한 비언어 표현 분석 등을 통해 평가의 객관성과 일관성을 높일 수 있다.

면접 후 최종 결정 단계에서는 사전에 정의된 인재상과의 적합도를 평가하거나, 복수의 면접자를 비교하여 보완적, 보충적 관점에서의 조직문화 적합성을 분석할 수 있다. 이러한 분석을 통해 인재 채용 과정에서 인간의 전문적 판단과 AI의 정량적 분석 능력이 상호 보완적으로 작용할 수 있는 협업 모델의 가능성을 확인하였다.

Table 9 Human-in-the-loop for AI system

Step	Task	페인포인트 (니즈)	AI 적합성 판단이유
면접 전 설계	1. 기업의 현 조직문화 진단	조직 핵심가치가 명확히 드러나지 않음. 주관적 감에 의존하거나 문서화 어려움.	OCP 응답 및 대화 내용을 NLP로 분류·요약하여 조직 가치 프로파일을 자동화할 수 있음.
	2. 인재상 정의 및 평가 기준 수립	조직마다 다르게 정의된 인재상 정리 어려움. 항목화/행동지표 구성이 어려움.	조직 가치와 연계된 행동지표·척도 예시를 자동 생성하고, 항목 간 유사도 감지 및 정제할 수 있음.
	3. 평가 질문 설계 및 면접 준비	평가자마다 질문 구성 방식이 상이하여, 구조화된 문서 없이 경험에 의존하여 평가하게 됨.	가치관, 행동기반, 상황판단의 질문 유형별로 자동 생성·매핑하고, 개인화 구성 문서 및 평가자별 역할 분배를 자동화 가능.
면접 중 실행	4. 인재상 기반 면접자 답변 평가	답변을 종합적으로 이해하고 판단하기 어려움. 면접 중 필요한 질문 놓치는 경우가 자주 발생함.	응답을 유형별로 자동 분류하고, 적합성 판단 기준에 따라 평가 요약 및 probing 질문 추천 가능.
	5. 비언어적 행동관찰	표정/말투/눈빛 등 비언어 단서 해석 시 개인의 주관 개입 많음. 실시간 기록이 어려움.	영상·음성 기반 감정 인식 및 정량 분석으로 비언어 표현을 자동 수집·요약하고 평가 보조 가능.
면접 후 최종 결정 및 팔로업	6. 비교 분석 후 최종 판단 및 결정	평가 기준이 불명확하여 후보자 비교가 불투명하고 평가 근거 부족함.	질문 유형별 평가를 자동 요약·시각화하고, 보완적/보충적 적합성 기준별로 후보 비교 및 리포트 자동 생성 가능.
	7. 피드백 및 결과 안내	피드백 제공이 어려움. 구조화된 피드백 없이 단순 결과 통보에 그침.	평가 내용을 자연어로 요약해 피드백을 생성하고, 상태별 메시지 자동화 및 문화 정보 포함 가능.
	8. 입사 후 기대 행동 검증 (수습기간)	입사 이후 기대했던 인재상과 역할에 부응하는지에 대한 객관적 평가가 다소 어려움.	채용 시 기준이 된 인재상 평가 지표 중심으로, 입사 후 주요 협업자들의 평가가 필요함.

3. 2. 2. HCAI 기반 인간-AI 협업 기획 영역 도출

본 장에서는 앞서 Table 8에서 정의된 네 가지 서비스 기회 요소였던 1) 조직 구성원 기반 조직문화 진단 및 기업의 인재상 정의, 2) 지원자별 맞춤 질문 생성 및 면접 전략 자동 배포 시스템, 3) 실시간 답변 및 행동 분석을 통한 인재상과의 핏 정밀 검증, 4) 기업-지원자 간 최종 핏 통합 평가 및 피드백 자동화를 중심 플로우로, 실제 면접 경험 기반으로 정리된 8개 핵심 과업을 재구성하고, 각 기회요소로 파악되었던 영역별로 HCAI관점에 따라 인간-AI 협업의 적합 방식과 역할을 정리하였다. 이는 단순히 기술 적용 가능성을 나열하는 것이 아니라, 조직문화 기반 채용이라는 맥락에서 인간의 개입과 AI의 처리 역량이 어떻게 유기적으로 작동할

수 있는지를 구조적으로 정리한 결과이다. 본 분석은 면접 전-중-후 전반의 흐름에 따라 앞서 정의된 구체적인 과업을 HCAI 협업 중심 관점에서의 정의되었던 네 가지 기회 요소로 다시 구조화한다. 이후, 각 과업의 구체적인 행동 요인의 성격에 따라 AI 협업 방식의 4가지 유형인 향상, 증폭, 보완, 강화로 연결하여 Table 10과 같이 정리하였다.

먼저, A 영역인 ‘조직문화 구조화 및 인재상 정의 지원’에서는 조직 내부의 언어 데이터(OCP 응답, 협업 대화 등)를 NLP 기반으로 분석하여 조직 고유의 핵심 가치를 자동 도출하고 클러스터링하는 작업이 수행된다. 이는 인간이 수행하기 어려운 방대한 언어 데이터를 해석하는 작업으로, AI의 분석 능력을 활용하는 증폭 방식이 적합하다. 이후 도출된 조직 가치에 따라 인재상의 행동 특성과 평가 척도를 자동 제안하고, 중복되거나 모호한 항목을 정제하는 과정에서는 인간의 도메인 전문성과 AI의 구조화 능력이 결합된 보완 방식이 적용된다.

B 영역인 ‘구조화된 면접 질문 설계 및 배포 자동화’는 질문 유형(가치관/행동기반/상황판단)에 따른 행동지표 기반 질문 생성을 중심으로, 질문-지표 간 매핑 문서화, 지원자 맞춤형 질문 세트 구성 등 면접 설계 과정 전반을 자동화하는 것을 목표로 한다. 이 역시 규칙성과 반복성이 강한 구조화된 작업으로서 보완 방식의 AI 개입이 효과적이다.

C 영역인 ‘비정형 정보 분석 및 실시간 평가 지원’에서는 면접 중 실시간으로 수집되는 정성적 응답 데이터를 AI가 유형별로 구조 분석하고, 적합성 판단 및 후속 질문을 자동 추천함으로써 평가자의 인지 부담을 경감시킨다. 이는 면접자의 사고 맥락을 이해해야 하는 작업으로 보완 방식의 협업이 요구된다. 또한 응답자의 표정, 시선, 말투, 미세떨림 등 비언어적 표현을 영상 및 음성 기반으로 자동 감지하고 정량화함으로써, 인간 평가자의 해석 편차를 보완하는 기능은 증폭 방식의 대표적인 예시로 볼 수 있다.

마지막으로 D 영역인 ‘평가 결과 통합 및 피드백 자동화’는 질문 유형별 응답 데이터를 기반으로 평가 리포트를 자동 생성하고, 지원자 간 적합성을 정량적으로 비교 분석하는 기능을 포함한다. 이를 통해 지원자와 기업 사이의 보완적 보충적 적합성은 물론, 가치 적합성, 행동 규범과 적응 및 성장 적합성도 함께 수치화하여 파악할 수 있다. 또한, 유사 구성원 프로파일과의 비교 예측 등을 통해 최종 우선순위 결정을 지원하는 강화 방식이 적용된다. 또한 평가 내용을 기반으로 합격·불합격 결과를 커스터마이징된 피드백 메시지로 자동 생성하고, 조직문화 관련 정보를 포함한 의사소통 품질을 향상시키는 기능은 반복성과 정성 판단이 결합된 증폭 방식의 협업이다. 마지막으로, 입사 이후에는 면접에서 평가한 조직문화 적합성과 실제 업무 수행 간의 일치 여부를 검증하는 과정이 필요하다. 구성원 피드백이나 업무 지표를 통해 입사자의 태도와 팀 적응을 평가하며, 이는 면접 기준의 유효성을 점검하는 역할을 한다. 반복 평가 구조의 정비와 자동화를 지원하는 보완 방식의 AI 협업이 적합하다.

이상의 네 가지 서비스 기회 영역은 인간의 전문성과 AI의 처리 능력이 상호 보완적으로 작용할 수 있는 구조를 제시함으로써, 기존 채용 프로세스의 객관성, 적합성, 효율성 향상에 기여할 수 있도록 설계되었다.

Table 10 Insight for Human-in-the-loop

서비스 기회 영역	과업	AI기능 인사이트	AI 권장 처리 방식
A. 조직문화 구조화 및 인재상 정의 지원	1. 기업의 현 조직문화 진단	조직 언어 분석 및 가치 클러스터링: OCP 응답 및 대화 데이터를 NLP 기반으로 분석해 조직 핵심 가치 자동 도출	증폭 (Amplify)
	2. 인재상 정의 및 평가기준 수립	행동지표 자동 제안 및 중복 제거: 조직 가치 입력 시 행동지표, 평가척도 자동 생성 및 중복/모호 항목 정제	보완 (Augment)
B. 구조화된 면접 질문 설계	3. 평가 질문 설계 및 면접 준비	질문 자동 생성 및 매핑 문서화: 행동지표 기반으로 질문 생성, 질문-지표 매핑표 자동 작성, 지원자 맞춤형 질문 세트 구성 면접 준비 자동화: 지원자 별 면접 준비 세트 내 면접관 역할 분담 및 자료 배포	보완 (Augment)
	4. 인재상 기반 면접자 답변 평가	답변 구조 분석 및 적합성 판단: 응답을 유형별로 분해·요약하고 적합성 판단 및 후속 질문 자동 추천	보완 (Augment)
C. 비정형 정보 분석 및 실시간 평가 지원	5. 비언어적 행동 관찰	비언어 정보 자동 감지 및 분석: 영상·음성 기반 표정·톤·리듬 감지, 감정 정량화, 언어-비언어 일치 여부 분석	증폭 (Amplify)
	6. 비교분석 후 최종 판단 및 결정	평가 리포트 자동 생성 및 후보 비교: 정량·정성 평가 자동 요약, 유형별 적합성 비교, 유사 구성원 행동 프로파일 기반 예측	강화 (Empower)
D. 평가 결과 통합 및 피드백	7. 피드백 및 결과 안내	피드백 자동 요약 및 커스터마이징: 평가 내용 자동 요약, 합볼 템플릿 추천, 문화 정보 포함 맞춤 메시지 자동 생성업무 지표 및 구성원 피드백 기반 평가 자동화: 입사자의 실제 업무 행동이 면접 평가 지표와 일치하는지 확인하고 평가 결과와 연동	증폭 (Amplify)
	8. 입사 후 기대 행동 검증	업무 지표 및 구성원 피드백 기반 평가 자동화: 입사자의 실제 업무 행동이 면접 평가 지표와 일치하는지 확인하고 평가 결과와 연동	보완 (Augment)

4. 서비스 경험 프로토타이핑 및 검증

4. 1. 협업 수준별 서비스 시나리오 설계

앞서 도출된 기회요인을 바탕으로, 스타트업 채용 과정 중 조직문화 적합성을 평가를 위해 면접 전, 중, 후 단계에서 AI와 인간의 협업 상황을 가정한 시나리오를 통해 개입의 적절성을 검증하였다. 각 단계별 구체적인 과업 수행 과정에 있어 AI의 개입도에 따라 시나리오 작성 후 적절한 협업 수준을 검증한 뒤, 적정 협업 수준의 협업 시나리오를 최종 프로토타입 구현 시 반영하였다.

시나리오 설계는 각 기회요인별 다음 두 가지 사항을 고려하여 작성하였다. 먼저, 각 과업별로 사전에 정의된 권장 처리 방식의 특징을 고려하고(AI의 개입 정도가 높은 수준부터 인간의 개입 수준이 낮은 순서로 향상, 증폭, 보완, 강화 중 하나로 지정), 각 권장 처리 방식 내에서 AI 개입 6개의 단계(Level)로 세분화하였다. Level 1은 AI의 개입 없이 인간이 모든 결정을 내리는 단계이며, 단계가 높아질수록 AI의 개입 수준이 점진적으로 증가하여, Level 6에서는 AI가 인간의 승인 없이 모든 결정을 내리고 시스템을 실행한다. 사용자 이용 플로우 중심에 설계된 4가지의 기회요인에 포함된 과업별로 6가지 Level의 협업 시나리오를 작성하였다.

Table 11 Example scenarios by level of Human-AI collaboration

시나리오 4: 면접 중 면접자의 답변에 따른 후속 연계질문 설계 및 문화 적합성 코멘트 정리 협업 방식: 보완(Augment)	
단계	시개입 시나리오
Level 1	면접자의 답변을 들으면서 실시간으로 면접 중 면접관이 직접 추가 질문을 생각하고 질의 후 코멘트 작성
Level 2	AI가 실시간으로 면접자의 답변을 분석하고, 답변의 모호성을 판단 및 안내하면 면접관이 직접 추가 질문을 생각하고 질의 후 코멘트 작성
Level 3	AI가 실시간으로 면접자의 답변을 분석하고, 답변의 모호성 판단과 후속 질문 후보 생성하면 면접관이 선택하여 질의 후 코멘트 작성
Level 4	AI가 실시간으로 면접자의 답변을 분석하고, 답변의 모호성 판단과 후속 질문 생성 후 우선순위를 통해 추천 시 질의 후 코멘트 작성
Level 5	AI가 실시간으로 면접자의 답변을 분석하고, 답변의 모호성 판단과 후속 추천 및 평가 기준 자동 제안
Level 6	AI가 실시간으로 면접자의 답변을 분석하고, 답변의 모호성 판단과 후속 질문 추천 및 추가 답변에 대한 코멘트 자동화

예를 들어, Table 11과 같이 면접 중 면접자의 답변에 따른 후속(Probing) 질문 설계 및 문화 적합성 코멘트 정리(기회요인 3)의 경우 권장 처리 방식은 보완으로 설정하였다. 가장 낮은 레벨인 Level 1은 AI의 개입 없이 면접관이 스스로 면접자의 답변을 실시간으로 판단 분석하여 추가 질문을 수행하는 것이며, 단계가 높아질수록 AI 면접자 응답의 모호성 판단, 후속 질문 후보 생성, 질문 자동 추천, 문화 기반 후속 질문 및 평가 문장 자동 제안, 전체 대화 흐름 기반 코멘트 자동화까지 점진적으로 AI의 역할이 확대되도록 시나리오를 설계하였다.

다만, 기회요인 D 평가 결과 통합 및 피드백의 마지막 과업인 입사 후 기대 행동 검증은, 실제로 입사 후에 행동 분석을 통해 검증 가능한 부분이므로, 설문 참여자의 객관적 판단이 어려운 항목이라 판단되어 유효성 유지를 위해 시나리오 선호도 조사 범위에서 제외하였다.

4. 2. 시나리오 선호도 조사

앞선 사용자 인터뷰를 통해 도출된 4가지 기회요인 내에서 총 7개(8번 과업 제외) 과업 중심으로, 6가지 Level로 총 42개의 시나리오를 대상으로 연구 참여자들이 적절하다고 생각하는 AI협업 수준을 탐색하기 위한 선호도 조사를 진행하였다. 조사는 스타트업 채용 과정 중 ‘면접’ 과정에서 조직문화 적합성을 판단하는 맥락 중심으로 AI와 인간의 협업 수준을 Level 1~6 중 선택하도록 설계되었다. 조사는 실제 스타트업에서 경험 가능한 현실적인 면접 상황을 몰입형 상황으로 설문에 함께 제공하여 참여자들이 조직문화 적합성 평가 면접 전형 과정에 몰입하도록 하였으며, 기술적 가능성보다는 각 과업별 인간-AI의 협업 균형을 중심으로 평가하도록 유도하였다.

조사 참여자는 실제 스타트업 구성원 혹은 최소 3년 이상의 실무 경력을 가진 직장인 32명을 대상으로 온라인 비대면 설문 형태로 진행하였으며, 소요 시간은 약 10~15분이었다. 모든 응답은 익명 처리되었고, 참여자들에게는 소정의 보상이 제공되었다.

4. 3. 시나리오 선호도 조사 결과

선호도 조사는 총 이틀간 진행되었다. 참여자의 성별은 여성 20명(62.5%), 남성 12명(37.5%)이었으며, 연령 분포는 30대 19명(59.4%), 20대 5명, 40대 4명, 50대 이상 4명으로 30대가 가장 많았다. 참여자는 금융, 교육, 공공행정 등 다양한 산업군에 종사하고 있으며, IT/소프트웨어 관련 업무를 담당하는 경우가 가장 많았고, 주로 중소기업이나 일반 기업에 재직 중이었다. 또한 참여자들의 현재 직급은 중간 관리자, 대리/주니어 순으로 많았다.

협업 시나리오 검증은 정성적, 정량적 분석으로 진행되었다. 먼저 인간의 통제와 AI 자동화를 축으로 4분면 모델 기반으로 인간-AI협업 개념을 매핑했을 때, 대부분의 참여자가 인간의 판단을 보조하고 의사결정을 돕는 역할로서 AI를 활용하는 보완(Augment) 영역에 가장 많이 매핑되었다. 또한, 평균적으로 AI가 인간의 승인을 전제로 작업하며 제한적 실행 자율성을 갖는 Level 4로 많이 조정되었다.

4. 4. 최종 AI 협업 서비스 시나리오 도출

시나리오 선호도 조사를 통해 대상 사용자들이 생각하는 AI의 적절한 개입 정도를 최종적으로 선정하였다. 이를 바탕으로 프로토타입을 설계하였으며, 확정된 최종 AI협업 서비스 시나리오는 아래 Table 12와 같이 정리되었다.

Table 12 Human-AI Collaboration Service Scenarios

단계	과업	서비스 시나리오	AI Level 1순위	AI Level 2순위
면접 전	조직 구성원 기반 조직문화 진단 및 기업의 인재상 정의 (기회요인 1)			
	1. 기업의 현 조직문화 진단	- 기업 구성원 대상으로 OCP 설문 진단 - 정성 대화/문서 등 자료 분석을 통한 업무 문화 파악 - 리더-구성원간의 회사 문화의 인식 차이 비교 - 현재 조직의 핵심 문화 키워드 도출	Level 4 (34.4%)	Level 3 (18.8%)
	2. 인재상 정의 및 평가기준 수립	- 핵심 문화 키워드 기반 채용 포지션/직무별 변형 적용 (채용 포지션별 목표 인재상 구체적 정의) - 평가 가능한 행동 지표 척도 설계	Level 4 (31.1%)	Level 2 (18.8%)
면접 중	지원자별 맞춤 질문 생성 및 면접 전략 자동 배포 시스템 (기회요인2)			
	3. 평가 질문 설계 및 면접 준비	- 행동지표 기반 질문 유형 설계 - 면접자 이력서 기반 상세 면접 질문 변형 및 목표 평가 항목과 매핑 - 면접관 별 면접 역할 목표 분배 및 면접 준비 가이드 배포	Level 3 (28.0%)	Level 2 (21.6%)
	4. 인재상 기반 면접자 답변 평가	- 질문 유형별 응답 분석 (가치관/행동기반/상황판단) - 면접자 답변 기반 실시간 정합성 판단 및 기록 - 답변 부족 시 추가 후속(probing) 질문 설계 및 문화 정합성 자동 코멘트	Level 3 (31.1%)	Level 5 (21.6%)
	5. 비언어적 행동 관찰	- 주요 비언어 평가 항목 (눈맞춤, 손 사용, 목소리 떨림 등) 실시간 분석	Level 3 (34.4%)	Level 5 (21.6%)
면접 후	기업-지원자 간 최종 피드백 통합 평가 및 피드백 자동화 (기회요인 4)			
	6. 비교분석 후 최종 판단 및 결정	- 평가 자료 통합 정리 및 정량, 정성 비교 분석 - 사전 정의된 채용 포지션 인재상과의 부합성 평가 - 기업-면접자 간의 보충, 보완 중심 조직문화 적합성 평가 - 지원자간 비교 후 우선순위 제안 및 최종 결정	Level 3 (40.0%)	Level 2 (18.8%)
	7. 피드백 및 결과 안내	- 면접에 대한 피드백 메시지 작성 - 최종 합격, 불합격 결과별 템플릿 자동 구성 및 발송	Level 4 (34.4%)	Level 3 (18.8%)

조사 결과, AI권장 처리 방식이 증폭인 경우 Level 4와 같이 AI의 개입도가 높은 방식을 더 선호하는 반면, 보완 및 강화의 경우에는 AI가 인간의 업무를 효율화하는 도구 역할로는 긍정적으로 평가하지만, AI가 스스로 판단을 내리지 않는 Level 3에서 가장 높은 선호도를 보였다. 과업별로 분석한 결과, 자료 수집, 평가 기준 설정, 반복적 분석 업무와 같이 정보 처리 중심의 과업에서는 AI의 높은 개입도를 선호했다. 반면, 실제 면접자를 대면하고, 면접자 평가, 최종 의사결정과 같이 인간의 판단이 핵심인 과업에서는 명확하게 인간 주도 방식이 선호되었다.

4. 5. 최종 서비스 콘셉트 제안

앞서 조사된 시나리오 선호도 분석 결과에 기반하여, 채용 과정에서 조직문화 적합성을 정량적·정성적으로 평가하고자 하는 니즈를 반영해 서비스 콘셉트를 도출했다. 조직문화가 명확히 정의되어 있지 않은 스타트업 및 소규모 기업을 주요 대상으로 설정하고, 채용 시 지원자와 조직 간의 문화적 적합도를 구조적이고 신뢰성 있게 평가할 수 있는 AI 기반 면접 솔루션을 제안한다.

본 서비스는 ‘Culture Bridge AI’라는 가칭으로, 채용 담당자가 현재 조직의 조직문화와 지향하는 인재상을 명확히 정의하고, 그에 부합하는 지원자를 효과적으로 식별할 수 있도록 돕는다. 구체적으로 AI 기반의 현 조직문화 진단, 채용 포지션별 인재상 정의, 맞춤형 면접 질문 설계, 면접 중 비정형 데이터 분석 등의 기능을 통해 궁극적으로 문화적 적합성 검증을 지원한다.

본 서비스의 주 타겟은 초기 스타트업 및 성장기 소규모 기업의 경영진을 비롯한 인사 및 채용 업무를 수행하는 실무자이다. 이들이 공통적으로 겪는 어려움은 기업 내부에서 조직문화나 인재상에 대한 정의가 부족하여 채용 시 명확한 기준 없이 판단이 이루어진다는 점이다. 이로 인해 조직과 지원자 간의 문화적 부적합이 채용 후에 발견되어 양측 모두 어려움을 겪는 문제가 발생한다. 이러한 문제의식을 바탕으로 본 서비스는 조직 내부 데이터를 분석하여 현재의 조직문화를 진단하고, 이를 기준으로 채용 포지션에 적합한 인재상을 구체적으로 정의한 뒤, 체계화된 면접을 통해 적합성을 사전 검증하는 것을 목표로 한다.

본 서비스의 차별점은 보편적인 “조직문화” 키워드가 아닌, 현 조직의 특성과 채용 포지션이 주로 협업할 직원들 간의 적합성을 보완적, 보충적 관점에서 다각적으로 평가한다는 점이다. 즉, 모든 기업에 대표적인 동일한 기준을 적용하는 것이 아니라, 기업별·포지션별로 최적화된 평가 기준을 사전에 수립하고, 이를 위한 맞춤 설계된 면접 프로세스를 통해 최적의 인재를 선별한다. 예를 들어, 본 서비스의 기회요인 2에 해당되는 지원자별 맞춤 질문 생성 영역에서는, 기회요인 1을 통해 구체화된 해당 기업의 문화적 특성을 반영하여, 지원자에게 조직문화 적합성과 개인 가치관을 파악하기 위한 질의나, 해당 기업의 실제 직무 상황 사례를 활용하여 상황 기반 질문을 생성하여 지원자의 판단력 및 의사 결정력을 파악하는 질의가 이루어질 수 있다. 최종적으로는 조직과 지원자의 적합성 평가를 통해 합격자 선정 후, 합격 안내와 함께 채용 기준과 평가 결과, 그리고 기대사항을 지원자에게 명확히 전달한다. 이를 통해 입사 전 단계에서 상호 이해도를 높이고 지속가능성을 확인하여 잠재적 시간낭비를 최소화하는 데 집중한다.

중요한 것은 본 서비스가 AI 단독 평가에 의존하는 것이 아닌 인간-AI협업 중심으로 설계되었다는 점이다. AI는 대량의 다양한 타입의 데이터 수집과 분석 과정에서 효율성을 제공하면서도 전 과정에서 아이디어가 필요할 시 다양한 옵션을 제안하지만, 각 단계별 핵심 의사결정은 모두 사용자(채용 담당자)가 수행하도록 하여 인간-AI 협업 밸런스를 최적으로 유지하는 것을 목표로 한다.

4. 6. 프로토타이핑 제작

앞서 정의된 서비스 콘셉트를 바탕으로, 채용 과정에서의 주요 문제를 해결하기 위한 네 가지 기회요인을 중심으로 프로토타입을 제작하였다. 각 기회요인은 실제 면접 프로세스상에서 발생하는 구체적인 니즈를 반영하였고, 서비스 흐름에 따라 기능적 구조로 구현되었으며, 그 결과는 Table 13과 같다.

첫째, 조직문화 및 인재상 정의 단계에서는 조직 문화 프로필 설정을 위한 구성원의 설문 응답(OCP)과 내부 커뮤니케이션 데이터(예: 슬랙 채널, 노션 페이지 등)를 AI가 분석하여 조직 고유의 핵심 가치와 문화를 도출하고, 이를 기반으로 채용 포지션의 이상적 인재상과 구체적인 행동지표를 정의하는 기능이 구현되었다. 둘째, 면접 질문 설계 및 자동화 단계에서는 정의된 행동지표에 기반하여 AI가 면접 질문을 자동으로 생성하고, 질문-지표 간 매핑표를 작성한 후, 면접관별 질문 분배 및 면접 가이드 문서를 자동으로 배포하는 기능을 포함하였다. 이를 통해 면접관 간 질문 중복을 방지하고, 지원자별 지원 직무에 따른 맞춤형 행동기반 및 상황판단 질문 및 평가 기준 설계가 가능하도록 하였다. 셋째, 실시간 평가 및 비정형 정보 분석 단계에서는 면접 중 AI가 지원자의 언어 표현, 표정, 말투, 응답 리듬 등을 실시간으로 분석하고 시각화된 데이터로 제공하여, 면접관이 즉시 적합성을 판단하거나 추가 질문을 설계할 수 있도록 지원하였다. 마지막으로, 평가 결과 통합 및 피드백 자동화 단계에서는 면접 종료 후 AI가 면접관의 정량적 평가 및 서술형 코멘트를 자동으로 정리하고, 채용 포지션에 정의된 인재상과의 적합도를 다각도로 분석하였다. 본 연구에서는 사전 분석 결과를 바탕으로, 기업이 추구하는 성향에 비해 보충적 적합성이 높은 경우와 보완적 적합성이 높은 경우를 구분하여 비교·평가하였다. 이를 위해 두 유형의 지원자 간 평가 지표의 차이가 뚜렷하게 나타나는 상황을 시나리오로 구성하고, 본 프로토타입에서 제시하는 다면적 분석 결과 화면을 기반으로 한 평가 과정을 가정하였다. 또한, 최종 합격 여부에 따른 피드백 메시지를 자동으로 작성하여, 합격자에게는 기대 역할을 구체적으로 전달하고, 불합격자에게는 개선점을 안내하는 등 일관된 피드백 경험을 제공하도록 설계되었다.

해당 프로토타입은 위의 네 가지 기회요인으로 정리된 구체적 기능을 중심으로 구성되었으며, 특히 조직문화 진단과 인재상 정의 기능, 그리고 정의된 기준에 따른 면접 평가 및 피드백 기능이 유기적으로 연결되도록 설계되었다. 이는 인사 담당자의 직관이나 감각에 의존하던 기존 면접 평가를 체계화함으로써, 보다 신뢰성 있는 채용 의사결정을 가능하게 하기 위한 구조적 기반을 제공한다.

Table 13 Examples of Screen UI

기회요인 1. 조직 구성원 기반 조직문화 진단 및 기업의 인재상 정의		
세부과업 #	협업 방식/ Level	서비스 화면 및 주요 기능
과업 #1 인재상 정의	Amplify/ Lv.4	<p><서비스 메인화면></p> <ul style="list-style-type: none"> - 현재 조직의 조직 문화 진단 결과 히스토리 - 인재상에 따른 현 구성원 분포도 - 조직 문화 진단 정의 히스토리 
과업 #1, 2 기업의 현 조직문화 진단, 인재상 정의 및 평가기준 수립	Amplify, Augment/ Lv.4	<p><조직 문화 진단 상세화면></p> <ul style="list-style-type: none"> - 조직 구성원의 진단 근거 및 신규 진단 연결 - 조직원간의 보완적/보충적 적합성 평가 - 신규 채용 포지션의 목표 문화 적합성 진단 지표 

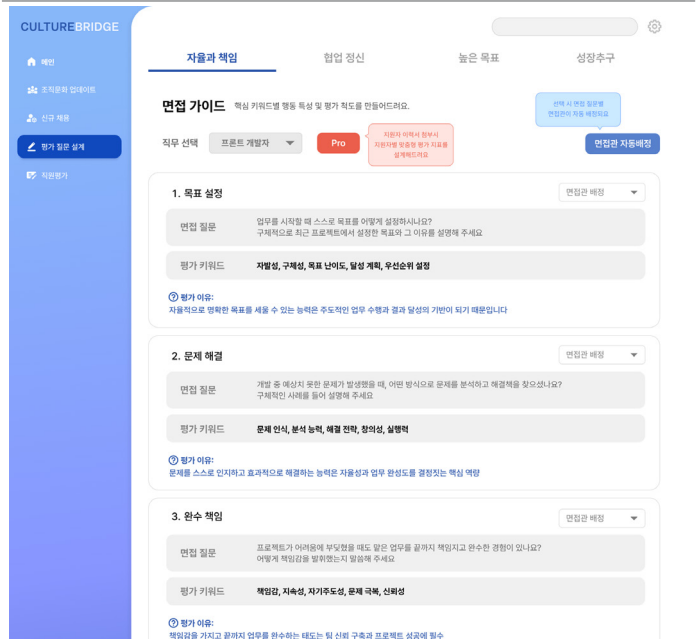
기획요인 2. 지원자별 맞춤 질문 생성 및 면접 전략 자동 배포 시스템

<면접 준비 화면>

- 신규 포지션의 목표 문화 적합성에 따른 평가 항목 자동 제공
- 지원자 이력서/자기소개서 연결 시 면접 질문 자동 커스텀
- 지원 포지션의 특징에 따른(예: 기획, 개발, 디자인 등) 상황판단, 행동기반 질문 자동 생성
- 질문별 면접과 자동 배정 및 면접 준비 자료 자동 배포

과업 #3
평가질문 설계
및 면접 준비

Augment/
Lv.3



기획요인 3. 실시간 답변 및 행동 분석을 통한 인재상과의 피트 검증

<면접 진행 화면>

- 면접 전체 흐름 텍스트 기록 및 주요 키워드 추출
- 준비된 평가 항목을 평가하기에 면접자 답변이 충분했는지 시각화 제공
- 부족할 경우 후속 질문 생성 및 추천
- 실시간 비언어적 행동 기록 및 평가

과업 #4, 5
인재상 기반
면접자 답변
평가 및
비언어적
행동 관찰

Augment,
Amplify/
Lv.3



기획요인 4. 평가 결과 통합 및 기업-면접자 적합성 평가

- <개별 평가 화면>
- 개인별 다각도 평가 결과 리포트 제공
- 다른 후보군과의 비교 평가를 위한 시각화 제공

과업 #6
최종 판단 및
결정

Empower/
Lv.3

- <최종 결과 메시지 발송 화면>
- 합격/불합격자 템플릿 설정
- 개인별 결과 피드백 자동 생성 후 수정 기능 제공
- 예약 발송 기능 제공

과업 #7
최종 결과
피드백 및
안내

Amplify/
Lv.4

본 프로토타입에 대한 평가는 연구 1과 2로 나누어 진행되었다. 연구 1은 본 연구에서 제시한 프로토타입에 대한 사용성 평가를 진행하기 위해 총 다섯 가지 관점의 정량 평가와 인덱스 정성 인터뷰로 진행되었다. 연구 1의 목적은 본 서비스가 타겟으로 하는 이용자를 대상으로 실제 문화 적합성을 평가하기 위한 면접 여정에서 시나리오에 따른 서비스 이용 시 사용성을 평가하는 것이다. 이후 사용자의 피드백이 일부 적용된 프로토타입으로 해당 서비스에 대한 효과성 검증을 위한 정량 평가가 진행되었다. 연구 2는 참여자 인식 기반 향상 문항으로 프로토타입 사용 전후의 변화를 동일한 시점에서 회상하여 비교하도록 설계되었다.

4. 7. 연구 1: 프로토타입 사용성 평가

4. 7. 1. 평가 항목 및 질문 도출

먼저, 연구 1의 프로토타입 사용성 평가는 Table 14와 같이 총 다섯 가지 관점으로 구성하였다. 먼저 Morville(2004)의 허니콤 모델과 Nielsen의 사용성 정의를 기반으로 한 사용성 평가 항목을 설정하였으며, 이를 통해 사용자가 각 기능을 학습하거나 조작하는 과정에서 어려움을 느끼지 않는지 확인하였다. 다음으로는, 설명 가능한 AI(XAI), 인간-AI 상호작용 설계, 인간중심 AI 원칙, 그리고 조직문화 적합도를 포함하여 본 연구에서 제시하는 프로토타입의 주요 평가 항목을 다섯 가지로 구성하였다. 이러한 항목들은 단순한 기능 검증이 아니라, 인간과 AI의 협업 경험이 어떻게 인식되고 수용되는지를 다층적으로 측정하기 위함이다.

Table 14 AI Service Experience Evaluation Criteria

속성	세부 속성	평가 기준
사용성 시스템 작동 방식을 이해하고 무리 없이 사용 가능한지	정보 이해 가능성	시스템이 제시하는 정보가 직관적이고 쉽게 해석될 수 있는가?
	단계별 학습 용이성	서비스 흐름이 복잡하지 않으며, 사용법을 쉽게 익힐 수 있는가?
	인터페이스	화면 배치나 시각 구성 요소가 자연스럽고 기능과 연결되는가?
신뢰성 AI의 판단 결과를 믿고 사용할 수 있는가	결과 설명 가능성	AI의 판단 및 제안 기준이나 과정을 이해할 수 있는가?
	자동화 신뢰도	자동 생성된 질문, 피드백, 결과에 대해 신뢰할 수 있는가?
상호작용성 AI의 개입이 사용자 흐름을 방해하지 않고 자연스럽게 작동하는가	행동유도 명확성	사용자의 다음 행동을 자연스럽게 명확히 안내하는가?
	맥락 적절성	AI의 추천, 피드백 등이 시점과 맥락상 적절한가?
	실시간 피드백 유용성	응답 흐름, 표정 감지 등 실시간 피드백이 판단에 도움이 되는가?
협업 적합성 AI가 판단을 대신하지 않고 인간의 결정을 보조하는 방식으로 사용되는가	AI 개입 수준 적절성	AI가 추천한 인제상/질문/결과가 조직의 핵심 가치와 일치하는가?
	인간의 통제권 유지	사용자가 판단을 수용/거부/수정할 수 있는 구조가 보장되는가?
	협업 유용성	실무적 효율 향상에 실제 도움이 되며, 반복 사용 의지가 생기는가?
결과 적합성 AI의 결과가 조직문화, 인제상 정의와 매칭되는가	결과의 문화 정합성	AI가 추천한 인제상/질문/결과가 조직의 핵심 가치와 일치하는가?
	피드백 정합성	자동 생성된 피드백이 면접 맥락 및 조직 기대와 부합하는가?
	질문-지표 가치 일치성	질문이 평가 지표와 조직 인제상의 행동 지표를 반영하는가?

4. 7. 2. 테스트 결과 및 피드백 분석

테스트 분석 결과 인간과 AI의 협업 경험의 인식과 수용과정을 평가한 5가지 항목인 사용성, 신뢰성, 상호작용성, 협업 적합성, 결과 정합성은 Table 15에 평균과 정성 분석의 주요 포인트가 정리되었다. 사용성 속성에서는 평균 4.42점, 신뢰성 속성에서는 평균 3.96점, 상호작용성 속성에서는 평균 4.15점, 협업 적합성 속성에서는 세부 항목 평균 4.61점, 결과 정합성 속성에서는 평균 4.02점으로, 협업 적합성이 가장 높게 평가되었다. 전체 평균은 4.23점으로, 대부분의 연구 대상자들이 5개의 기회요인과 각 시나리오에 대하여 협업 적합성, 사용성, 상호작용성 측면에서 고르게 높은 점수를 준 것으로 이해할 수 있다. 결과적으로, AI 개입이 인간의 판단 및 통제권과 조화를 이루는지에 대한 평가인 협업 적합성이 4.61점으로 가장 높게 평가되었다. 이는 즉, 사용자는 언제든지 AI의 판단을 사용자의 의사에 따라 수용, 거부, 수정할 수 있었다고 해석된다. 하지만 신뢰성 속성은 평균 3.96점으로, 다른 속성들에 비해 다소 낮은 평가를 받은 것으로 나타났다.

Table 15 AI Service Experience Evaluation Results

평가 항목	평균	주요 포인트 (정성분석)
사용성	4.42	<p>긍정</p> <p>“지금 현재 조직이 갖고 있는 인재상이 무엇인지 쉽게 알 수 있음 + 어떤 식으로 변화가는지 히스토리 파악이 용이함” (기회요인1/과업1)</p> <p>“질문 완료 여부가 시각적으로 표시되는 점이 면접 흐름을 파악하는 데 유용했고...” (기회요인3/과업4)</p>
		<p>개선</p> <p>“평가 결과가 표 형식으로 정리되어 있고, 후보자별로 어떤 점에서 강점/보완점이 있는지 시각적으로 비교되는 점이 인상적...” (기회요인4/과업6)</p>
신뢰성	3.96	<p>긍정</p> <p>“질문의 수나 표현이 조금 더 다양했으면 좋겠다는 아쉬움이 있었고...” (기회요인1/과업2)</p> <p>“자동으로 생성된 피드백이 조금은 일반적임... 조직이나 팀 특유의 말투나 정서가 들어가야 진정성 있게 느껴질 것 같음...” (기회요인4/과업7)</p>
		<p>개선</p> <p>“조직도에 있는 프로필(사진 그대로 매핑에 활용하므로, 우리 회사에 대한 진짜 데이터 기반임을 쉽게 알 수 있음)” (기회요인1/과업1)</p> <p>“결과적으로 시가 사람에 대한 적합도 등의 평가가 올바른 결과였는지는 결국 채용 후 실제로 업무를 하면서 알게 될 테니.. 시간이 흐르면서 신뢰도가 점점 더 쌓이는 구조일 듯 함” (서비스 전반)</p>
상호작용성	4.15	<p>긍정</p> <p>“당장은 얼마나 내가 이 시스템을 신뢰할 수 있는지 말할 수 없지만, 신뢰가 바탕에 깔린다면 사람 뽑을 때 저게 마치 복극성 지표처럼 따르면 될 것 같음” (기회요인1/과업1)</p> <p>“...디자인 같은 직무는 팀에 따라 요구하는 특성이 미묘하게 다를 수 있기 때문에... 커스터마이징 흐름이 중요할 것 같음”(기회요인1/과업1)</p> <p>“표정이나 말투, 리듬 등의 비언어 분석 결과는 있었지만, 왜 그런 상태로 해석되었는지 설명이 부족함” (기회요인3/과업5)</p>
		<p>개선</p> <p>“Team DNA 맵을 통해 어떤 팀과 잘 어울릴지까지 예측하고, 서로간의 적합도를 자유롭게 확인 가능...” (기회요인1/과업1, 2)</p> <p>“면접 답변에 따른 후속 Probing 질문이 자동으로 추천되어 면접자의 생각을 더 깊이, 주어진 시간을 최대한 활용하여 파악 가능”</p> <p>“실시간으로 시가 상황 판단력과 행동 기반 질문을 추천해줘서 면접관이 놓치기 쉬운 부분까지 짚어줄 수 있음” (기회요인3/과업4)</p>
협업적합성	4.61	<p>긍정</p> <p>“실시간 분석 시각화와 질문 추천이, 실제 면접 상황에서 약간 정신없을 수도 있을 듯함.”</p> <p>“면접관이 오히려 더 면접자 내용에 집중하지 않을까 우려” (기회요인3/과업4, 5)</p>
		<p>개선</p> <p>“질문을 사람이 직접 수정 가능한 구조도 유연함” (기회요인1/과업2)</p> <p>“최종 메시지는 사람이 세밀하게 조정할 수 있는 구조가 반드시 유지되어야 할 것 같다고 느낌” (기회요인4/과업7)</p> <p>“여러 면접관의 평가 내용이 자동으로 정리되는 과정이 특히 인상 깊고... 모든 최종 결정은 인간 주도가 반드시 필요” (기회요인4/과업6, 7)</p>
결과정합성	4.02	<p>긍정</p> <p>“AI 해석을 그대로 판단에 반영하는 것은 위험할 수 있다는 인상을 받았습니다.”</p> <p>“조직 특수성 반영을 위해선 사용자 개입 여지가 더 열려야 한다는 의견 존재.” (모든 과업 전반에 해당)</p>
		<p>개선</p> <p>“가치관-행동 특성-관찰 표현” 구조가 실제 인재상과 평가 지표를 연결하는 데 있어 명확한 기준으로 작용” (기회요인1/과업2)</p> <p>“지원자와의 적합성을 충분히 파악하기 위한, 질문 유형별 충분한 판단 응답을 받았는지를 면접 중 가이드해주는 게 유용” (기회요인3/과업4)</p> <p>“지원자의 조직 적합도를 수치화하고, 협업자별 pros/cons까지 시각화해주시 단순한 ‘좋다/나쁘다’ 판단이 아닌 근거 기반의 채용 결정을 도와주는 점이 납득 가능함” (기회요인4/과업6)</p>
		<p>개선</p> <p>“면접 질문의 문맥과 표현이 일반적인 인상이 있음” (기회요인2/과업3)</p> <p>“피드백 문장이 조직의 고유한 표현이나 맥락을 충분히 반영하지 못해 피상적으로 보일 수 있음.” (기회요인4/과업7)</p>

프로토타입 테스트 중 함께 진행한 정성 인터뷰를 통해 주요 평가 항목(사용성, 신뢰성, 상호작용성, 협업적합성, 결과 정합성) 내용을 정성 분석했을 때 주요 포인트 또한 Table 13에 함께 정리되었고, 상세 내용은 다음과 같다.

먼저, 사용성 측면에서 면접 질문 자동 생성, 응답 정합성 시각화, 질문 분배 및 면접 결과 피드백 초안 작성 기능은 실무자의 인지적·시간적 부담을 효과적으로 감소시키는 것으로 평가되었다. 특히 AI를 활용한 초안 작성 및 템플릿 추출 기능은 업무 효율성 향상에 기여하는 것으로 나타났다. 다만, 기업 특성을 반영한 커스터마이징 범위 확대의 필요성이 지적되었다. 다시 말해, 모든 기업에 동일하게 템플릿화해서 제공될 초안 수준에서 나아가, 본 기업만의 문화와 특성이 담긴 초안의 구조로 구성되기를 기대한다는 의견이 존재했다.

신뢰성 측면에서는 몰입형 시나리오 기반 테스트의 한계로 인해 전반적으로 다소 낮은 평가를 받았다. 사용자들은 AI 기반 기능들이 다소 일반적인 내용을 제공한다는 피드백을 반복적으로 제시하였으며, 자동화된 기능이 실질적 신뢰를 얻기 위해서는 결과 도출 과정에 대한 명확한 설명이 필요하다고 강조하였다. 다만, 실사용 상황을 고려하였을 때, 과업4와 같은 면접 중 실시간 후속 질문 생성 시에는 키워드 중심으로, 내용을 파악하기 편리한 방안으로 화면 구성이 풀이내 지기를 제안하기도 하였다. 또한, 비언어 정보 분석, 정합성 점수, 우선순위 제안 등의 결과가 어떤 판단 기준과 가중치를 기반으로 생성되었는지에 대한 설명 부족이 결과 수용의 주요 한계로 나타났지만, 본 서비스의 신뢰 여부는 결국 서비스 이용을 통한 채용 후, 직원과 회사의 핏이 적합함을 점진적으로 체감되면, 서비스에 대한 신뢰도 자체는 자연스럽게 올라갈 것이란 의견도 공통적으로 보였다.

한편, 상호 작용성 측면에서는 면접 중 실시간으로 이루어지는 면접자 답변 분석 후 후속 질문 안내, 비언어적 행동 분석 기능은 면접 시간을 효율적으로 활용하는 긍정적 효과로 평가되었다. 그러나 해당 항목과 관련해서도, 일부 사용자는 시스템의 실시간 개입이 면접관의 집중력을 저해할 수 있다는 우려를 표하며, 인지적 흐름을 방해하지 않는 인터페이스 설계의 중요성이 다시 한 번 언급되었다.

협업 적합성 측면에서는 AI가 반복적이고 구조화 가능한 업무를 자동화하면서도 사용자가 언제든지 개입하여 수정 및 조정할 수 있는 구조가 효과적으로 작용하였다. 면접 질문이나 피드백 메시지 작성에서 자동화된 초안을 바탕으로 사용자가 직접 맥락에 따라 문장을 다듬는 구조는 인간중심 AI의 실천적 설계 원칙을 구현한 예시로 볼 수 있다. 또한, 면접 전후의 과업에서, 본래는 면접 준비 파일, 이력서 파일 등 다양한 관련 문서를 모두 별도로 열어놓고 봐야하는 어려움이 존재했지만, 하나의 시스템 내에서 AI의 가이드에 따라 관련 문서를 쉽게 접근할 수 있다는 점도 협업 적합성의 일부 긍정 포인트로 언급되었다. 다만, 기업 고유 문화를 더욱 잘 드러낼 수 있는 표현 방법에 대한 LLM 학습을 통해 보완이 필요할 사항으로 예상된다.

마지막으로, 결과 정합성에 관해서는 시스템이 제시하는 결과가 조직의 인재상이나 문화적 맥락과의 부합도는 사용자의 수용 여부를 결정짓는 핵심 기준으로 작용하였다. 질문이나 피드백이 조직 고유의 언어, 톤, 가치 기준과 어긋날 경우, 표현의 부자연스러움이 결과 전반에 대한 신뢰 저하로 직결되었다. 이는 AI 기반 결과물이 기능적 완성도와 함께 문화적 일관성 및 정서적 정합성을 동시에 갖추어야 실제 업무 맥락에서 효과적으로 수용될 수 있음을 시사한다.

이 외 추가적인 의견으로는, 특히 면접관들 간 의견 차이가 존재할 시 본 서비스의 유용성을 더욱 긍정적으로 평가한다는 의견도 다수 분석되었다. 특히나, 두 명 이상의 후보자 중 최종 합격 의견이 일치하지 않을 때, 시스템이 제시하는 지표가 가장 객관적인 판단 근거가 될 수 있다는 점에 전반적으로 동의하였다. 일부 사용자는 이를 단순한 평가 도구를 넘어, 향후 채용 과정에서 보다 명확한 판단 기준을 구축하는 데 기여할 수 있는 데이터 축적의 측면에서 긍정적으로 인식하였다. 이러한 결과는 앞서 문헌조사를 통해 언급한 스타트업의 특성 과도 맞닿아 있다. 인사 업무에 대한 전공 배경이 없는 채용 담당자들이 기존 조직문화를 유지하면서도 보완적 적합성을 지닌 인재를 선발하는 과정에서, 본 시스템이 기업 입장에서 유익한 참고 데이터로 활용될 수 있음을 확인하였다.

4. 8. 연구 2: 업무 개선도 기반 정량적 검증

연구 1에서 수집된 정성적 사용자 경험 기반 피드백을 일부 반영한 뒤, 본 연구에서는 최종 프로토타입 화면을 기반으로 실제 업무 개선도에 대한 기대효과를 정량적으로 검증하기 위한 추가 연구를 진행하였다. 본 연구가 제안하는 프로토타입 화면은 연구 참여자별로 개인의 조직 환경이나 특정 기업의 업무 흐름을 반영한 것이 아니며, 연구 참여자가 기존에 경험한 채용 경험 또한 하나의 조직이나 하나의 직무에 한정되지 않는다는 점에서, 특정 기업의 고유 채용 프로세스와 프로토타입을 직접적으로 비교하는 것은 적절하지 않다고 판단했다. 따라서 연구 2에서는 참여자가 과거에 경험한 채용 관련 실무 경험 중 본 연구에서 요구하는 특정

목적과 상황에 따라 종합적으로 회상하게 한 뒤, 동일한 시점에서 프로토타입 기반 시나리오의 활용성을 함께 평가하도록 구성하여, 참여자 인식 기반 항상 문항을 활용한 정량적 비교 평가 방식을 적용하였다.

또한, 연구 1에서 수집된 정성적 사용자 경험 기반의 개선 피드백을 반영하여, 프로토타입 화면 내 예시 문구를 보다 구체적이고 현실적으로 구성함으로써, 참여자가 실제 면접 준비, 질문 구성, 평가, 피드백 과정에서의 개선 가능성을 명확하게 상상할 수 있도록 하였다. 본 연구 2는 총 30명의 참여자를 대상으로 설문문을 통해 진행되었으며, 이는 HCI 분야에서 사용자 인식 기반 정량 평가를 수행할 때 일반적으로 20~40명 규모의 표본을 활용하는 연구 관행(Farzand et al., 2024)에 부합한다. 따라서 본 연구에서 확보한 30명의 표본을 기술통계 기반의 개선도 분석을 수행하는 데 적절한 수준으로 판단하였다.

4. 8. 1. 평가 항목 및 연구 진행 방식

연구 2에서는 참여자 인식 기반 항상 문항을 활용하여, 본 연구에서 제안하는 서비스 사용 시 조직문화 적합성 관점에서 보다 적합한 인재를 채용하는 과정에서 실제 업무가 개선될 수 있는지를 평가하였다. 참여자 인식 기반 항상 문항은 프로토타입 사용 전후의 변화를 동일한 시점에서 회상하여 비교하도록 설계되었다. 이는 실제 채용 담당자들이 속한 조직마다 사용하는 채용 관련 도구의 종류와 활용 수준이 크게 상이하며, 경우에 따라 별도의 도구를 사용하지 않는 사례도 존재하기 때문이다. 이러한 환경적 차이로 인해 본 프로토타입 수준의 서비스를 기존 방식과 동일 기준에서 직접 비교하는 것은 타당하지 않다고 판단되었으며, 이에 본 연구에서는 참여자가 과거에 경험한 채용 프로세스를 회상하고 이를 바탕으로 본 연구에서 제안된 시스템 도입 후의 기대 변화를 종합적으로 평가하도록 하였다. 설계는 이와 같은 회상 기반 자기보고 방식에 근거한 것으로, 특정 솔루션 도입 전후를 동일 시점에서 회상하여 비교할 경우 응답 변동 편향을 최소화하고, 인지적 변화나 자기보고식 평가에서 보다 타당한 결과를 얻을 수 있다는 장점이 있다(Lam & Bengo, 2003).

이와 같이 인식 기반으로 평가한 세부 항목은 AI 채용 시스템 도입 기대 성과로 거론되는 여섯 가지 항목 중 네 가지를 중심으로 구성하였다(Son & Oh, 2023). 선행연구에 따르면, AI 채용 시스템의 주요 도입 성과는 시간 효율성 증가, 업무 효율성 증가, 커뮤니케이션 용이성 향상, 객관성 향상, 더 넓은 지원자 풀 확보, 공정성 향상의 여섯 가지로 정리된다. 그러나, 본 연구의 프로토타입은 채용 전 과정보다는 면접 준비 과정에 초점을 맞춘 서비스이므로, 기업의 아웃바운드 채용 활동에 해당하는 더 넓은 지원자 풀 확보와 인사 담당자와 지원자 간 조율 업무와 관련된 커뮤니케이션 용이성 향상은 제외하였다. 결론적으로 Table 16에 정의된 시간 효율성, 업무 효율성, 객관성 향상, 공정성 향상의 네 가지 항목을 중심으로, 제안된 서비스의 도입 전후 인식 변화를 7점 리커트 척도(1=매우 그렇지 않음, 7= 매우 그러함)로 평가하도록 설계하였다.

Table 16 Perception-Based Performance Improvement Evaluation Criteria

속성	평가 기준
시간 효율성 증가	- AI 시스템 도입 이전에는 채용 면접 전/중/후 단계별 업무에 시간이 많이 소요되었는가? - AI 시스템 도입 후 채용 면접 전/중/후 단계별 업무에 소요되는 시간이 보다 효율적으로 관리되었는가?
업무 효율성 증가	- AI 시스템 도입 이전에는 채용 면접 관련 업무에 반복적인 작업으로 인한 업무의 비효율성이 있었는가? - AI 시스템 도입 후에는 채용 면접 관련 업무가 보다 체계적이고 효율적으로 진행된다고 느끼는가?
객관성 향상	- AI 시스템 도입 이전에는 주관적 판단이 업무에 영향을 많이 주었는가? - AI 시스템 도입 후에 데이터와 기준에 근거한 일관된 판단이 이루어진다고 느끼는가?
공정성 향상	- AI 시스템 도입 이전에는 지원자 평가에 완전한 공정성이 보장된다고 느꼈는가? - AI 시스템 도입 이후에는 지원자 평가에 완전한 공정성이 보장된다고 느끼는가?

4. 8. 2. 연구 절차 및 대상

연구 2는 온라인 설문 형태로 진행되었으며, 최소 3회 이상 면접관으로 면접 참여 경험이 있는 사람을 대상으로 진행되었다. 설문 시작 시 면접 참여 경험 여부를 체크하는 스크리닝 질의로, 3회 이상 면접관 경험이 없는 경우 설문 참여를 제한하였다.

구체적인 설문 진행 방법은 앞서 Table 12에서 정리되었던 면접 전-중-후 단계에 따른 4가지의 기회요인별로 포함된 업무 테스크를, 설문 참여자는 기존에 어떻게 진행해 왔는지를 회상하도록 하는 주관식 질문에 답변하고, 본 서비스에서 제안하는 새로운 방식을 충분한 설명과 함께 이미지로 확인한 뒤 Table 15에 정의된 시간 효율성, 업무 효율성, 객관성, 공정성에 대한 평가를 기존 방식과, 본 연구에서 제안하는 새로운 방식에 대해 각각 수행하게 하였다. 그리고 마지막으로, 이전 방식에 비해, 본 서비스에서 제안하는 방식이 각 네 가지 지표에서 얼마만큼 개선되었다고 생각하는지를 평가하는 질의로 마무리하였다.

4. 8. 3. 분석 결과

연구 2는 총 30명을 대상의 온라인 설문으로 진행되었다. 참여자 성별은 남성과 여성이 각각 15명으로 동일했으며, 연령은 30대가 16명으로 가장 큰 비중을 차지했고, 50대 5명, 40대 4명, 20대 3명 순으로 집계되었다. 설문 응답자는 다양한 조직 규모와 직급에서 면접관 경험을 보유하고 있었으며, 회사 규모는 중소기업·스타트업(48.6%)이 가장 높았고, 중견기업(18.9%)과 대기업(16.2%)도 고르게 분포하였다. 직급은 대리·중간 관리자(34.2%)가 가장 많았으며, 팀장·매니저(21.1%)와 임원·대표(21.1%) 등 실제 채용 의사 결정권을 가진 비율도 높았다. 직무는 인사·채용(27.9%), 기획(20.9%), 디자인(18.6%), 개발(14.0%) 순으로 구성되어, 다양한 실무 관점에서 면접 경험을 반영하고 있음을 확인할 수 있다.

또한 설문 시작에 앞서 면접 과정에서 느꼈던 어려움과 불편함을 자유롭게 기술하도록 한 결과, 참여자들이 경험한 핵심 문제는 크게 세 가지로 나타났다. 첫째, 면접관 간 질문 기준·방향이 서로 달라 중복 질문, 질문 누락, 면접 순서 혼선, 지원자 정보 혼재 등 평가 일관성 부족 문제가 반복된다는 점이였다. 둘째, 제한된 시간과 정보 속에서 지원자의 실제 업무 역량, 협업 스타일, 조직문화 적합성을 판단하는 것이 어렵다는 평가 정확성의 한계가 지적되었다. 셋째, 이력서 관리, 일정 조율, 면접 노쇼 대응, 면접 역량 부족 등 운영 리소스 부담이 크다는 점도 공통적으로 언급되었다. 이러한 문제는 면접의 신뢰성·효율성·객관성 전반에 영향을 미치는 것으로 나타났으며, 주요 사례는 Table 17에 정리하였다.

Table 17 Summary of Reported Difficulties in Past Interview Procedures

주요 요인	특징	예시
평가 일관성 부족	<ul style="list-style-type: none"> - 면접관마다 질문 방향과 관심 포인트가 제각각 - 중복 질문, 질문 누락, 순서 혼선 - 연속 면접이 진행될 시, 면접자 정보 혼재 및 기억 혼란 발생 - 평가 기준이 없어, 주관적 판단에 의존 	<p>“면접이 많아지면서 점점 초반 면접과 후반 면접자 평가 일관성이 부족했다” P26</p> <p>“같은 날 여러 명을 보면 비슷한 경험이 혼재되어, 한 명 한 명이 기억이 안 난다” P7</p> <p>“면접관 간에 질문 공유가 전혀 안 돼서, 민망했다” P2</p>
지원자 역량 및 적합성 판단의 어려움	<ul style="list-style-type: none"> - 실제 업무 투입 시 역량이 어떤지 판단이 어려움 - 문화 적합성 판단 기준이 모호함 - 짧은 시간으로 인해 파악이 어려움 	<p>“실제로 같이 업무를 할 때 어떤지 확신이 안 들었다” P14</p> <p>“짧은 시간 안에 사람을 파악하기 어려웠다” P4</p> <p>“우리 회사랑 잘 맞을지 판단이 안 선다” P15</p>
프로세스 및 리소스 부담	<ul style="list-style-type: none"> - 여러 포지션을 동시에 진행하는 경우, 이력서 관리 및 일정 진척사항 관리의 어려움 - 면접 노쇼 스트레스 - 스케줄 조율의 부담 - 면접관의 면접 역량 부족 	<p>“이력서가 몰리는 경우 관리가 어려움” P8</p> <p>“서류 작업이 힘들었다” P12</p> <p>“당시 디자이너가 저뿐인데, 저는 면접관 경험이 없어서 난감했다” P16</p> <p>“사업을 처음 시작하고, 누굴 어떻게 뽑아야 할지 지식이 부족했다” P30</p>

이어 본 연구에서는 네 가지 기회요인별로 기존 방식과 신규 서비스 간 체감 효과 차이를 검증하기 위해 대응표본 t-검정을 실시하였다(Table 18). 분석 결과, 시간 효율성, 업무 효율성, 객관성, 공정성 평가에 있어, 모든 기회요인에서 신규 서비스가 기존 방식보다 유의미하게 더 높은 점수를 보였다($p < .001$). 특히 시간 효율성과 업무 효율성에서 큰 차이가 나타났는데, 이는 신규 서비스의 질문 생성·추천, 분석 자동화, 평가 정리 기능이 실제 면접 준비 및 진행 과정에서 실질적인 업무 부담 감소와 시간 절약 효과로 이어짐을 의미한다. 또한 주관성과 공정성에서도 신규 서비스가 일관된 개선을 보이며, 면접관 간 편차를 줄이고 체계적인 기준을 제공하는 데 효과적이라는 점이 확인되었다.

기회요인별 구체적인 차이를 살펴보면 다음과 같다. 첫째, 조직 구성원 기반 조직문화 진단 및 기업의 인재상 정의 과정에서 기존 방식은 대부분 문서화가 이루어지지 않고, 구성원의 경험이나 감각에 의존하는 경우가 많았다. 일부는 경영진 인터뷰나 회의를 통해 비공식적으로 수행되었으나, 공식적인 HR 프로세스의 부재로 인해 부서별/개인별 기준 차이가 발생하는 것을 확인했다. 결과적으로 기존 방식은 네 가지 평가 항목 모두 낮은 점수(3.37~3.73)를 보였다. 반면 신규 서비스는 자동 진단·구조화·시각화 기능을 통해 모든 항목에서 5.57~5.70점으로 크게 향상되었으며, 이는 모두 통계적으로 유의미함($p<.001$)을 확인하였다.

둘째, 지원자별 맞춤 질문 생성 및 면접 전략 수립 과정에서 기존 방식도 사전 이력서 검토는 이루어졌지만, 질문 구조가 일관되지 않고 면접관 간 역할 분배나 평가 기준 공유가 부족해 중복 질문 혹은 질문 누락이 발생하는 경우가 많았다. 조직문화 적합성 평가 및 행동 기반 질문 역시 체계적으로 준비되기 어려웠다. 이에 비해 신규 서비스는 지원자별 맞춤 질문 생성, 면접 전략 안내, 질문 분배 자동화 등을 통해 평가 항목 전체에서 유의미한 개선을 보였다.

셋째, 실시간 답변 및 행동 분석을 통한 컬처핏 검증 과정에서는 기존 면접관들이 실시간 기록 부담, 후속 질문 생성의 어려움, 행동 기반 평가 기준의 모호성 등으로 어려움을 겪었다. 면접에 집중하기 어려운 상황이 반복되며 평가 품질 또한 저하되는 경험으로 인해 이번에도 역시 모든 평가 항목에서 신규 서비스가 더 높게 나타났다. 특히, 실시간 분석, 후속 질문 추천, 자동 기록 기능을 제공하여 모든 평가 항목이 유의하게 상승했으며, 특히 업무 효율성에서 가장 높은 개선 폭이 나타났다.

마지막으로, 기업과 지원자 간 최종 핏을 통합적으로 평가한 뒤 합격자를 결정하는 최종 의사결정 단계에서 기존에는 기준이 명확하지 않고, 상황과 사람에 따라 결과가 크게 달라지는 비구조적 프로세스였다. 또한, 면접관끼리의 의견 조율은 대부분 주관적·정성적 의견 취합 수준이었으므로, 신규 서비스가 모든 평가 항목에서 유의미하게 높게 평가되었다.

종합하면, 연구 2의 결과는 신규 서비스가 기존 면접 방식의 구조적 한계를 명확히 해소하며, 시간 효율성, 업무 효율성, 객관성, 공정성 측면에서 전반적인 면접 경험을 유의미하게 향상시키는 서비스임을 실증적으로 입증하였다.

Table 18 Paired T-Test Results

기회요인	평가 항목	기존 방식의 Mean (S.D)	신규 서비스의 Mean (S.D)	t	p
1. 조직 구성원 기반 조직문화 진단 및 기업의 인재상 정의	시간 효율성	3.57 (1.17)	5.70 (0.79)	-7.66	0.00***
	업무 효율성	3.73 (1.08)	5.60 (0.86)	-7.39	0.00***
	객관성	3.37 (1.27)	5.57 (1.04)	-6.97	0.00***
	공정성	3.47 (1.00)	5.57 (1.10)	-8.70	0.00***
2. 지원자별 맞춤 질문 생성 및 면접 전략 자동 배포 시스템	시간 효율성	3.10 (1.09)	5.70 (0.87)	-10.69	0.00***
	업무 효율성	3.17 (1.05)	5.63 (0.82)	-10.82	0.00***
	객관성	3.03 (1.12)	5.63 (0.93)	-11.25	0.00***
	공정성	3.13 (0.97)	5.67 (1.01)	-9.79	0.00***
3. 실시간 답변 및 행동 분석을 통한 인재상과의 핏 정밀 검증	시간 효율성	3.08 (1.09)	5.73 (0.88)	-8.73	0.00***
	업무 효율성	3.08 (0.95)	5.73 (0.87)	-8.60	0.00***
	객관성	3.08 (1.17)	5.50 (1.03)	-7.94	0.00***
	공정성	3.10 (1.13)	5.60 (1.05)	-6.99	0.00***
4. 기업-지원자 간 최종 핏 통합 평가 및 피드백 자동화	시간 효율성	3.57 (1.41)	5.53 (0.98)	-6.61	0.00***
	업무 효율성	3.70 (1.36)	5.43 (1.01)	-5.52	0.00***
	객관성	3.43 (1.41)	5.50 (1.20)	-7.19	0.00***
	공정성	3.57 (1.41)	5.60 (1.07)	-5.64	0.00***

$p<.05^*$, $p<.01^{**}$, $p<.001^{***}$

결론적으로 앞서 진행된 연구 1에서의 사용성 평가와 연구 2의 업무 개선도 평가 결과를 종합하면, 신뢰성 영역에서 추가적인 보완이 필요함에도 불구하고, 제안된 인간-AI 협업 기반 시스템은 기존 면접 프로세스의 핵심 문제점(과도한 주관성, 정보 누락, 일관성 부족, 기록 부담, 의사결정 지연)을 실질적으로 해결하며, 면접관의 업무 부담을 크게 덜어주는 것으로 나타났다. 특히 연구 1에서 확인된 인간 중심적 개입 가능성과 연구 2에서 검증된 정량적 효율성 향상 측정은 상호보완적 관계로 작용하여, AI가 인간의 판단을 강화하는 방식으로 설계될 때 면접의 품질과 효율성을 동시에 개선할 수 있음을 보여준다. 따라서 본 연구는 조직문화 기반 채용이라는 복잡한 의사결정 상황에서 AI의 역할은 대체가 아닌 협업적 보조자이며, 이러한 설계 원칙이 사용자 수용성뿐 아니라 실제 효율성 측면에서도 높은 효과를 발휘할 수 있음을 실증적으로 제시한다.

5. HCAI 기반 조직문화 적합성 면접 시스템 개발을 위한 청사진 제시

본 연구는 AI 기술을 활용하여 스타트업 및 성장기 소규모 기업이 조직문화에 적합한 인재를 효과적으로 채용할 수 있도록 지원하는 서비스를 제안한다. 제안하는 서비스는 단순히 조직문화 적합성만을 판단하는 것이 아니라, 조직문화 진단부터 인재상 정의, 그리고 적합성 평가까지의 전 과정을 체계화한다.

구체적으로는 현재 조직의 문화를 진단하고, 채용 포지션에 필요한 인재상을 정의하며, 조직원들 간의 보충적·보완적 적합성을 종합적으로 평가하여 조직문화 부적합으로 인한 채용 실패를 방지하고자 한다. 이 과정에서 조직문화 정의를 경영진 혹은 인사팀 등 조직원 중 일부의 개인 시각에서 주관적으로 판단하거나 정의하는 것이 아닌, 전체 조직원의 평가와 내부 협업 데이터를 기반으로 객관적으로 진단한다. 이를 토대로 신규 채용 포지션의 이상적 지표를 설정하고, 구조화된 면접 질문과 다면적 평가를 통해 적합성을 검증하는 체계를 구축한다. 이러한 서비스 구조를 시스템 맵과 서비스 블루프린트를 통해 이해관계자 간의 관계와 역할, 그리고 서비스의 작동 방식을 명확히 시각화하여 제시한다.

5. 1. 서비스 시스템 맵

기존의 면접관과 지원자 간 소통 중심의 면접 프로세스는 AI 기술 도입을 통해 보다 체계적인 시스템으로 발전할 수 있다. 본 서비스의 시스템 맵은 면접 전·중·후 단계에서 일관된 조직문화 지표를 바탕으로 적합성을 판단하는 구조를 보여준다. 주요 특징은 조직문화 정의 과정에서의 집단적 참여이다. 경영진 혹은 인사 담당자 등 채용 업무를 진행하는 실무자만이 지원자와 접점을 갖는 것이 아니라, 전체 조직 구성원이 조직문화 정의 과정에 참여한다. 이를 통해 실제 조직 현실을 반영한 문화 지표를 수립할 수 있다. 채용 완료 후에는 신규 입사자가 조직문화 형성에 기여하는 구성원이 되어 선순환 구조를 만든다. 이러한 사이클이 반복되면서 스타트업의 모호했던 인재상이 점차 구체화되고, 성공 사례 축적을 통해 시스템과 사용자 모두 채용 역량을 자연스럽게 향상시키는 학습 구조를 형성한다.

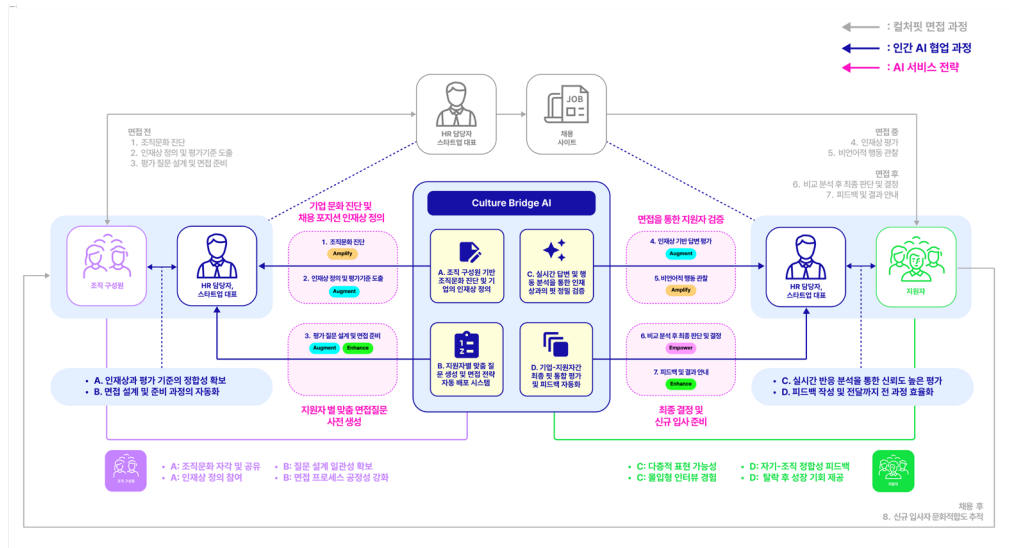


Figure 2 System map

5. 2. HCAI 기반 조직문화 적합성 평가 중심 면접 시스템 서비스 블루프린트

본 연구는 기존의 컬처핏 면접 과정을 기반으로 고객의 행동과 요구를 세분화하고, 각 단계별 터치포인트와 니즈 분석을 통해 잠재적인 서비스 기회요인을 도출하였다. 고객 여정을 조직문화 적합성 기반의 채용 담당자 관점에서 ‘면접 전-중-후’ 세 단계로 구분하고, 앞서 정의한 7가지 주요 과업을 중심으로 각 단계에서 인간과 AI의 상호작용 과정을 체계적으로 분석하였다. 특히, 프론트페이지에서는 채용 담당자의 명시적 행동 흐름을, 백스테이지에서는 AI가 자동 수행하는 분석, 추천, 생성 등의 작업 흐름을 함께 제시하였으며, AI 협업 유형(향상, 증폭, 보완, 강화)에 따라 시스템 내 작동 방식도 구분하여 설계하였다. 사용자가 실제로 접하는 화면과 내부 시스템이 지원하는 데이터 흐름 등은 line of visibility 기준으로 구분하여 정리하였다. 이와 같은 블루프린트 구조는 단순한 면접 지원 기능을 넘어, 조직문화 적합성 평가라는 맥락적이고 복잡한 과업을 AI가 어떻게 인간과 협력하며 실행할 수 있는지를 드러내는 기반으로 기능한다.

먼저 면접 전 단계의 [기회요인1/ 과업1] ‘기업의 현 조직문화 진단’은 조직문화 기반 면접을 위한 출발점으로, 조직원 대상 설문과 내부 커뮤니케이션 톨(슬랙), 협업 문서(노션)를 분석하여 조직의 핵심 가치와 협업 방식을 도출하여 이상적인 인재상을 정의하는 과업이다. 이 과정에서 AI는 조직 내 비정형 데이터를 종합적으로 분석하며, 인간은 분석 개시와 결과 승인 권한을 갖는다. 프론트페이지에서는 담당자가 분석 요청을 시작하고, 필요에 따라 조직원은 설문을 수행하게 된다. 백스테이지에서는 AI가 자동으로 데이터를 수집·분석하여 구조화된 인재상 초안을 생성한다. 이는 AI가 인간의 작업을 증폭(Amplify)하는 협업 방식으로 수행되며, 이후 승인된 결과는 다음 과업으로 연계된다.

[과업2] ‘인재상 정의 및 평가기준 수립’은 [과업1]에서 도출된 인재상에 기반하여 AI가 이를 평가하기 위한 적절한 행동 지표와 평가 항목을 추천하고 자동 구성하는 단계이다. 사용자는 이 정보를 검토하고 필요시 보완하여 확정한다. 프론트페이지에서는 평가 기준의 수용·수정이 이루어지며, 백스테이지에서는 AI가 기존 데이터베이스 및 유사 기업 사례를 기반으로 평가 지표를 자동 추천한다. 이 과정은 AI가 주도적으로 인간의 업무를 보완(Augment)하는 협업 방식으로 정의되며, 인간은 AI의 도움을 받지만 최종적인 의사 결정은 사용자 본인이 직접 하는 방식으로 진행된다.

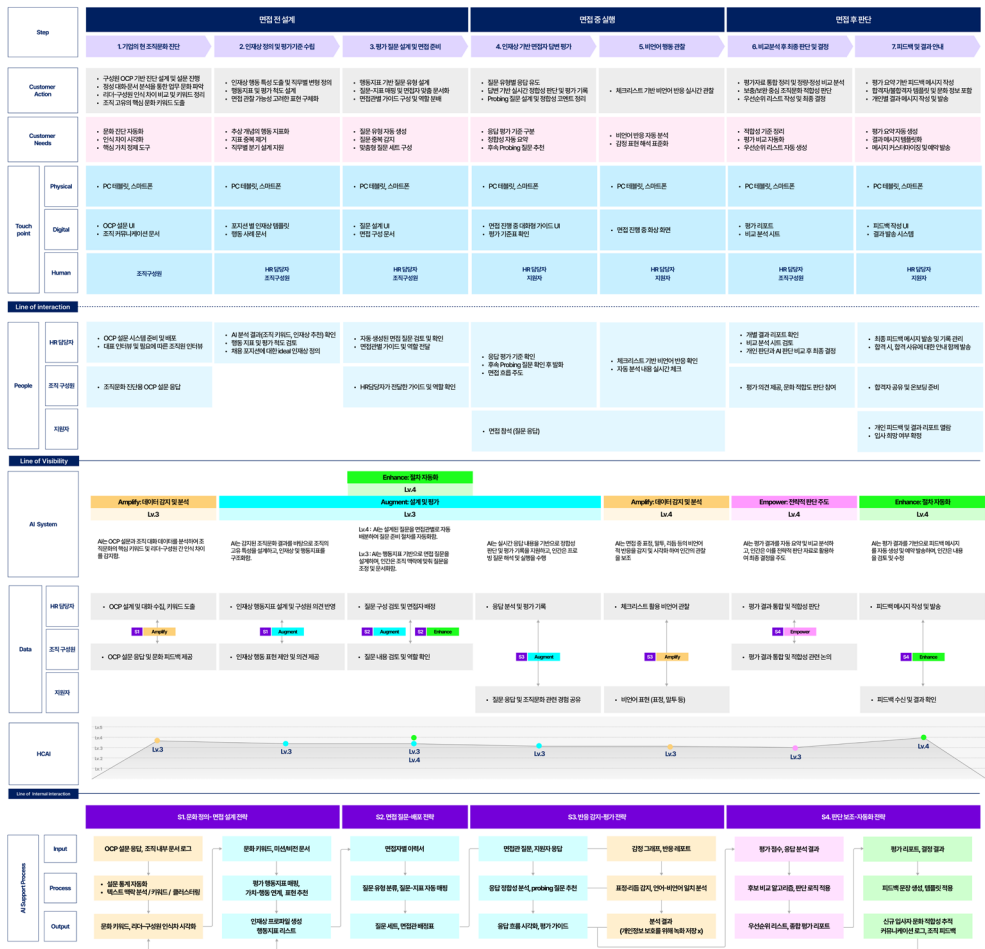


Figure 3 Service blueprint

[기회요인2/ 과업3]인 ‘평가 질문 설계 및 면접 준비’ 단계에서는 AI가 행동 지표 기반 질문을 면접자별 이력서를 참고하여 맞춤형으로 자동 생성하고, 면접관별 질문 배분 및 일정 공유 문서를 구성한다. AI는 백스테이지에서 질문 생성을 주도하되, 프론트스테이지에서는 담당자가 질문을 검토·수정하거나 새로운 질문을 제안할 수 있다. 이 과정은 사용자의 판단에 따라 유연하게 조정되며, 인간의 업무를 보완(Augment)하는 방식의 협업이 이루어진다. 해당 단계부터는 본격적으로 면접자별 맞춤화가 필요하므로, AI의 다양한 대안과 제시를 기반으로 사용자가 직접 판단하는, 인간 중심 의사결정이 중요한 단계이다.

면접 중 단계에서는 [기회요인3/ 과업4, 5]가 동시에 이루어진다. [과업4] ‘인재상 기반 면접자 답변 평가’는 실시간으로 AI가 면접자의 응답을 분석하여 적합성을 판단하고, 부족한 부분에 대해 후속 질문을 추천하는 기능이다. 면접관은 제안된 질문을 수용하거나 거절할 수 있으며, 프론트스테이지에서는 실시간 면접 진행이, 백스테이지에서는 응답 텍스트 분석과 후속 질문 추천이 동시에 이루어진다. 또한 응답 텍스트 분석을 통해 사전에 준비된 면접 중 검증하고자 하는 지표가 모두 확인되었는지 시각적으로 프론트스테이지에 노출된다. 해당 지표를 통해 면접관은 면접 진척률을 실시간으로 확인할 수 있다. 해당 과업은 인간 중심의 판단을 유지하면서 AI의 분석 역량을 활용하는 보완(Augment)방식의 협업이다.

[과업5] '비언어적 행동 관찰'은 면접자의 음성 안정감, 목소리 떨림, 시선 처리 등 비언어적 행동 데이터를 기반으로 AI가 실시간 분석을 수행하고, 면접 종료 시 요약 리포트를 생성한다. 영상 자체는 저장되지 않으며, 분석 결과만 저장되어 개인정보를 보호한다. 프론트페이지에서는 면접관이 이를 참고해 종합 평가를 진행하며, 백스페이지에서는 센서 기반 AI 분석이 자동으로 실행된다. 이는 AI가 인간이 감지하기 어려운 정보를 대량으로 분석하는 증폭(Amplify)방식으로, 판단에 필요한 정보의 양을 대폭 확장시킨다.

면접 종료 후에는 [기회요인4/ 과업6]을 통해 면접자별 '비교 분석 후 최종 판단 및 결정'이 이루어진다. AI는 정량·정성 평가, 비언어적 행동 분석, 이상적 인재상 일치도 등 다양한 정보를 통합하여 후보자별 리포트를 자동 생성하고, 후보 간 비교 분석 결과를 제공한다. 프론트페이지에서는 채용 담당자가 이를 기반으로 최종 결정을 내리며, 백스페이지에서는 AI가 후보별 장단점을 구조화하여 제안한다. 또한, 내부 데이터베이스에 저장된 현재 조직원의 성향 프로필과 면접자 간의 보충적·보완적 적합성 리포트도 포함되며, 프론트페이지에서 사용자가 원한다면 현 조직 모든 구성원들과의 1:1 적합도 분석도 받아볼 수 있다. 즉, 내부 데이터베이스에 저장된 조직원 개별 성향과 면접을 통해 확보된 면접자의 평가 결과는 모두 데이터로 존재하므로, 필요에 따라 다양한 분석이 가능하다. 이 단계는 AI가 다양한 분석 방향을 통해 인간의 최종 결정의 근거를 강화(Empower)하는 협업 방식으로 정의된다.

[과업7] '피드백 및 결과 안내'에서는 AI가 합격자와 불합격자 대상 이메일 초안을 자동 작성하며, 사용자가 이를 확인하고 수정 후 발송한다. 프론트페이지에서는 메일 발송 승인, 백스페이지에서는 면접 기록 기반 자동 초안 생성이 이루어지며, 이는 AI가 사용자 업무 효율을 증폭(Amplify)하는 방식으로 협업이 이루어진다.

전반적으로 본 서비스 블루프린트는 면접 전-중-후에 걸친 일련의 과업에서 인간과 AI의 협업 수준을 명확히 구분하고, 프론트/백스페이지 간 역할 구조를 기반으로 상호작용을 체계적으로 제시하였다. 특히, AI 수집·분석·추천 등 반복적이거나 복잡한 연산 작업을 자동화함으로써 인간의 인지적 부담을 줄이고, 인간은 맥락 기반의 판단과 의사결정에서 주도권을 유지한다. 이와 같은 협업 구조는 HCAI의 원칙에 기반한 시스템으로, 일방적 자동화를 지양하고 인간의 주도성과 AI 기능을 균형 있게 조율한 사례라 할 수 있다.

또한, 본 과정에서 생성된 모든 데이터는 일회성으로 소모되지 않고, 조직 고유의 인재상과 가치 기준을 지속적으로 학습·정교화하는 피드백 루프 구조를 형성한다. 이는 조직문화 진단-인재상 정의-면접 평가-결과 피드백의 전 과정이 연결되어, AI 시스템이 단순 자동화 툴이 아니라 조직문화 관리와 인재 판단을 위한 협력자로 기능할 수 있음을 보여준다. 따라서 본 연구의 서비스 블루프린트는 HCAI 관점에서 AI의 실질적 역할과 조직 적합성 평가에 대한 새로운 구조를 제안했다는 점에서 학문적·실무적 의의도 제공하게 된다.

6. 결론

6. 1. 연구 요약

본 연구는 초기 스타트업 및 소규모 기업이 조직문화 적합성 판단의 어려움으로 인해 겪는 채용 실패 문제를 해결하기 위한 인간-AI 협업 기반의 면접 서비스 모델을 설계하고 검증하는 것을 목표로 하였다. 스타트업의 제한된 자원 환경과 조직문화의 비정형성을 고려하여, 기존 면접 시스템과 절차를 보완하면서도 효율적이고 신뢰 가능한 걸쳐핏 평가 솔루션을 도출하고자 하였다. 연구 과정에서는 실제 채용 시장에서 조직문화 진단을 위해 사용하는 다양한 진단 방법 중 OCP 진단, 구조화 면접 질문 유형, AVI-AI 기술 분석, HCAI 협업 모델에 관한 이론적 고찰과, 현재 스타트업 대표 및 인사업무 담당자들의 사용자 인터뷰를 통해 페르소나와 고객 여정 맵을 도출하였다. 타깃 사용자들의 리얼 보이스를 통해 면접 단계 별 조직문화 적합도를 평가하는 데에서 발생하는 어려움을 파악하고, 각 단계별 적절한 인간-AI 협업 방안을 설계하고 검증하여 최종적인 프로토타입

테스트를 진행하였다. 프로토타입 테스트는 두 차례의 연구에 걸쳐 사용자 경험 관점의 사용성 평가 및 만족도와, 제안된 서비스에 대한 기대되는 업무 효과 개선도를 평가하였다.

연구 1을 통해 도출된 핵심 결과로는 각 채용 단계별 인간-AI 협업 방식에 따라 AI 시스템의 개입 수준과 역할을 구조화함으로써, 실제 스타트업 사용자들이 AI를 보조 도구로 수용하면서도 판단 권한을 유지할 수 있는 방식을 제시할 수 있었다. 또한 조직문화 진단, 질문설계, 면접자 반응 분석, 면접자 평가 및 최종 합격 결정까지의 전 면접 단계에서 AI는 데이터 기반 인사이트를 제공하여, 지원자에 대한 맥락적 판단을 지원한다. 결국 사용자는 AI 시스템을 단순 자동화 도구가 아닌 조직문화 적합성 판단을 지원하는 결정 보조자로 기능할 때 가장 신뢰하고 수용할 가능성이 높다는 결과를 얻을 수 있었다.

이어 수행된 연구 2에서는 보다 실증적인 검증을 위해 30명을 대상으로 한 정량적 설문 평가로 제안된 서비스가 실제 면접 프로세스에서 기대되는 업무 개선 효과성을 평가하였다. 그 결과, 연구 참여자들의 기존 면접 경험에서 반복적으로 나타났던 비효율성·비일관성·주관성 등의 문제들이 본 연구에서 반복적으로 이야기하는 네 가지의 기회요인 모두에서 신규 서비스에 의해 유의미하게 개선될 것으로 기대됨을 확인했다. 대응표본 t-검정 결과 신규 서비스는 시간 효율성, 업무 효율성, 객관성, 공정성 모두 기존 방식 대비 유의미하게 높은 점수를 보였으며, 이는 제안된 시스템이 실제 실무 환경에서 면접 전·중·후 전반의 품질과 효율성을 확실하게 향상시킬 수 있음을 실증적으로 입증한다.

종합적으로 본 연구는 스타트업이 직면하는 채용 의사결정의 복잡성과 조직문화 적합성 판단의 어려움을 해결하기 위해, 인간-AI 협업 구조를 기반으로 한 실질적인 면접 지원 모델을 제시하였으며, 정성적·정량적 검증을 모두 통해 서비스의 실효성과 업무개선 효과를 확인하였다. 본 연구는 AI가 단순 자동화 도구가 아닌, 판단의 근거와 맥락을 강화하는 결정 보조자로 기능할 때 사용자의 신뢰와 수용성이 가장 높아진다는 점을 밝혔다는 점에서 의의가 있다.

6. 2. 연구의 시사점 및 향후 연구 제안

본 연구는 인간-AI 협업을 통해 지원자와 기업 간의 조직문화 적합성을 평가하는 새로운 프로토타입 시스템을 설계하고 검증하는 과정 중 학문적, 실무적 시사점을 도출하였으나, 동시에 명확한 한계점도 존재한다.

6. 2. 1. 연구의 학문적 시사점

본 연구는 AI가 우리 삶에 점차 깊숙이 개입하고 있는 현실 속에서, 무조건적인 AI 활용만이 최선이 아니라는 관점을 전제로 시작하였다. 이에 따라 기술의 전면적인 자동화보다는 인간과의 협업을 전제로 하는 인간중심 AI(HCAI) 관점을 면접 시스템에 실천적으로 적용하였다는 점에서 학문적 의의를 지닌다.

기존 AI 기반 채용 시스템이 기술적 정확성과 효율성에만 집중해온 것과 달리, 본 연구는 조직문화 적합성이라는 정성적이고 맥락 의존적인 평가 영역에서 AI를 도입하여, 인간의 판단 주도성과 개입 권한을 유지한 상태에서 AI가 실질적으로 기여할 수 있는 역할과 방식을 탐색했다. 이 과정에서 Shneiderman (2020)을 비롯한 여러 학자들이 제안한 인간중심 AI 프레임워크를 기반으로, 인간-AI 협업의 자율성과 개입 수준에 따라 향상(Enhance), 증폭(Amplify), 보완(Augment), 강화(Empower)의 네 가지 협업 유형을 구체화하였다. 각 유형은 AI 개입의 깊이와 역할에 따른 단계별 차이를 포함하며, 면접 시스템 설계에서 상황별로 적용 가능한 협업 방식을 이론적으로 체계화하였다는 점에서 HCAI 이론의 실천적 확장에 기여할 수 있다.

또한, 본 연구는 조직문화 적합성 기반의 면접 절차를 구조화하고, 각 단계에 적합한 인간-AI 협업 강도를 설계 및 검증했다는 점에서도 학문적 의의가 있다. 협업 유형 내에서도 AI의 개입 깊이와 역할에 따른 단계별 차이를 명확히 설정하여, 조직문화 진단, 인재상 정의, 질문 설계, 응답 분석 등 면접의 전 과정에 적용하였고, 시나리오

기반 사용자 실험을 통해 유형별 활용 가능성과 효과성을 실증적으로 검토하였다. 이를 통해 사용자들이 AI의 맥락적 개입에 대해 높은 신뢰성과 만족도를 보였으며, HCAI 이론이 실제 면접 시스템 설계에서 어떻게 구현될 수 있는지를 보여주는 실증적 근거를 제시했다. 이는 기존 AI 기반 채용 시스템의 한계를 넘어, AI를 단순 보조 도구가 아닌 인간과 함께 협업하는 주체로서 사용자 목적과 경험(UX)에 기반한 방향으로 재정의하려는 이론적 확장을 시도했다는 점에서 의의가 있다.

6. 2. 2. 연구의 실무적 시사점

실무적으로는 본 연구는 자원과 인사 경험이 부족한 초기 스타트업 및 소규모 기업에 적합한 실질적인 조직문화 적합성 평가 대안을 제시하였다는 점에서 의미가 있다. 기존 대기업 중심의 표준화된 평가 도구는 스타트업 고유의 문화와 맥락을 반영하기 어렵고, 일률적인 기준을 강제한다는 한계가 있었다. 본 연구에서 제안한 시스템은 각 조직이 자체적으로 인재상과 조직문화를 정의할 수 있도록 설계되어, 조직의 상황과 니즈에 맞는 맞춤형 평가 환경을 제공한다. 이는 실제 도입 시 사용자 신뢰와 수용성을 높이는 전략적 기반이 될 수 있다. 특히, 조직문화 적합성이라는 영역은 본질적으로 정성적이며, 객관적인 수치로 환산하기 어려운 분야임을 고려할 때, 본 시스템은 채용 업무 담당자가 실질적으로 체감할 수 있는 유연한 평가 구조를 지원한다. 또한, AI를 단순히 판단을 대체하는 도구가 아닌, 인간의 인식과 판단을 보완하는 협업적 파트너로 작동하도록 설계함으로써, 실무자가 의사결정 과정에서 주도권을 유지할 수 있다는 점 역시 실질적 강점으로 작용한다.

더 나아가, 연구 2의 정량적 분석 결과는 본 시스템이 실무 운영 측면에서도 유의미한 개선 효과를 제공할 수 있음을 보여준다. 구체적으로, 참여자들은 기존의 문화 적합성 기반 면접 방식과 비교했을 때 본 시스템이 시간 효율성, 업무 효율성, 객관성, 공정성 측면에서 모두 개선 가능성이 높다고 평가하였다. 이는 스타트업 채용 과정에서 반복적으로 지적되어 온 인력 부족, 준비 과정의 비효율, 평가 기준의 불일치 등 실무적 문제들을 체계적으로 완화할 수 있음을 시사한다. 즉, 본 시스템은 문화 적합성이라는 본질적 난제를 다루면서도, 실무자가 실제 업무 흐름에서 체감할 수 있는 개선 효과를 제공한다는 점에서 높은 실용성을 가진다. 또한 연구 2에서는 동일한 방식으로 네 개의 핵심 기회요인을 평가한 결과, 프로토타입이 면접 전·중·후 단계의 업무 부담을 분산시키고 판단의 일관성을 높이는 방향으로 작동할 수 있음을 정량적 근거로 확인하였다.

무엇보다 중요한 점은, 본 시스템이 조직문화 평가에 있어 표준화된 하나의 정답지를 제시하려는 시도가 아니라, 각 기업이 스스로 정답지를 정의하고 설계할 수 있도록 지원하는 방식이라는 점이다. 이는 고정된 평가 프레임워크를 일괄 적용하는 것이 아니라, 기업마다 서로 다른 가치관·문화적 특성·인재상에 맞춘 평가 기준을 구성할 수 있는 실질적인 자율성을 보장한다는 데에서 핵심적인 실무적 시사점이 있다. 결과적으로, 이 시스템은 조직의 고유한 인사 전략과 채용 철학을 존중하면서도 데이터 기반의 신뢰도 높은 채용 결정을 가능하게 하며, 장기적으로는 각 기업이 자체적인 채용 성공 기준 데이터를 축적해 나가는 데 실질적인 기여를 할 수 있을 것으로 기대된다.

6. 2. 3. 연구의 한계점 및 향후 연구 제안

본 연구는 다음과 같은 세 가지 한계점을 바탕으로 향후 연구 방향을 제안한다. 첫째, 본 연구의 사용성 평가는 국내 스타트업 중사자의 소수만 대상으로 한 정성 인터뷰 데이터를 기반으로 진행되었으므로, 조직문화 관점에서 스타트업 전반에 공통적으로 적용 가능한 가치관, 협업 태도, 조직문화 적합성과 관련된 질문 중심으로 면접 구조를 정의하였다. 따라서, 특정 산업을 아우르는 맞춤형화를 통해 직무 적합성의 판단을 포함하기에는 한계가 있다. 따라서 향후 연구에서는 다양한 산업군과 조직 구조를 반영한 확장된 실험 설계를 통해 포괄적인 타당성을 확보할 필요가 있다. 둘째, 본 연구는 단편적인 시점에서 시나리오 기반의 사용성 평가를 진행하였기 때문에, 실제 서비스 이용에 대한 판단으로 보기에 제한적일 수 있다. 참여자의 인식 기반으로 AI 채용 시스템의 전 후의 성과로 효율성, 객관성, 공정성 등의 항목을 평가하였으며, 업무의 개선 효과가 검증되었으나, 앞으로는 본 시스템을 실제 도입한 뒤, 이를 통해 채용된 인재가 조직문화에 얼마나

잘 적용하는지를 추적하는 종단적 연구(longitudinal study)가 필요하다. 마지막으로, 본 연구는 실제 작동 시스템이 아닌 프로토타입 수준의 기능을 제한적으로 구현하고 평가하였으며, 실제 기업 환경에서의 활용 가능성을 충분히 검증하지 못했다. 향후 연구에서는 MVP 모델 개발을 통해 기술적 고도화와 함께 실험적 파일럿 테스트를 수행하여, 채용 평가부터 입사 후 후속 관리까지 이어지는 실제 업무 맥락에서 시스템의 효과성과 신뢰성을 검증할 필요가 있다.

본 연구의 목적은 조직문화 적합성을 기반으로 한 스타트업 채용 과정에서, AI와 업무 담당자의 협업 수준을 인간중심 AI 관점에서 탐색하고 검증하는 데 있었다. 이러한 연구 결과는 향후 실제 서비스 운영의 토대가 될 수 있으며, 서비스가 현장에서 활용되는 과정에서 사용성과 효과성을 축적해 나감으로써 시스템의 완성도와 고도화를 달성할 수 있을 것이다.

References

1. Agarwal A., George, R. A., Ravi, S. S., Kamath, S., & Kumar, A. (2020). Leveraging multimodal behavioral analytics for automated job interview performance assessment and feedback. *Proceedings of the Second Grand-Challenge and Workshop on Multimodal Language(Challenge-HML)*, 46-54. <https://doi.org/10.18653/v1/2020.challengehml-1.6>
2. Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., ... & Horvitz, E. (2019). Guidelines for human-AI interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1-13. <https://doi.org/10.1145/3290605.3300233>
3. Berretta, S., Tausch, A., Ontrup, G., Gilles, B., Peifer, C., & Kluge, A. (2023). Defining human-AI teaming the human-centered way: A scoping review and network analysis. *Frontiers in Artificial Intelligence*, 6, 1250725. <https://doi.org/10.3389/frai.2023.1250725>
4. Cable, D. M., & DeRue, D. S. (2002). The convergent and discriminant validity of subjective fit perceptions. *Journal of Applied Psychology*, 87(5), 875-884. <https://doi.org/10.1037//0021-9010.87.5.875>
5. Cable, D. M., & Edwards, J. R. (2004). Complementary and supplementary fit: A theoretical and empirical integration. *Journal of Applied Psychology*, 89(5), 822-834. <https://doi.org/10.1037/0021-9010.89.5.822>
6. Cardon, M. S., & Stevens, C. E. (2004). Managing human resources in small organizations: What do we know? *Human Resource Management Review*, 14(3), 295-323. <https://doi.org/10.1016/j.hrmr.2004.06.001>
7. CB Insights. (2014). *The top 20 reasons startups fail*.
8. Chatman, J. A. (1989). Improving interactional organizational research: A model of person-organization fit. *Academy of Management Review*, 14(3), 333-347. <https://doi.org/10.5465/amr.1989.4279063>
9. Chen Q., Li, J., Shen, R., Wang, R., Xu, J., & Zhou, J. (2024). Career adaptability and work engagement: The roles of person-organization fit and job insecurity. *BMC Psychology*, 12(1), Article 562. <https://doi.org/10.1186/s40359-024-01907-2>
10. Chung, S. H. D., & Parker, S. C. (2023). Founders affiliations: Jobseeker reactions and impact on employee recruitment by start-up ventures. *Small Business Economics*, 61(1), 259-283. <https://doi.org/10.1007/s11187-022-00694-7>
11. Coad, A., Daunfeldt, S.-O., Hözl, W., Johansson, D., & Nightingale, P. (2014). High-growth firms: Introduction to the special section. *Industrial and Corporate Change*, 23(1), 91-112. <https://doi.org/10.1093/icc/dtt052>
12. Crawford, K. (2021). *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press.
13. Ehsan, U., & Riedl, M. O. (2020, July). 1. Human-centered explainable ai: Towards a reflective sociotechnical approach. In *International Conference on Human-Computer Interaction* (pp. 449-466). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-60117-1_33

14. Farzand, H., Al Baiaty Suarez, D., Goodge, T., Alexander Macdonald, S., Marky, K., Khamis, M., & Cairns, P. (2024, May). Beyond Aesthetics: Evaluating Response Widgets for Reliability & Construct Validity of Scale Questionnaires. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* (pp. 1–7). <https://doi.org/10.1145/3613905.3650751>
15. Festor, P., Habli, I., Jia, Y., Gordon, A., Faisal, A. A., & Komorowski, M. (2021). Levels of autonomy and safety assurance for AI-based clinical decision systems. In *Computer Safety, Reliability, and Security. SAFECOMP 2021 Workshops: DECSoS, MAPSOD, DepDevOps, USDAI, and WAISE* (pp. 291–296). Springer. https://doi.org/10.1007/978-3-030-83906-2_24
16. Forbes Technology Council. (2019). Why AI needs human oversight. Forbes. Retrieved from <https://www.forbes.com/councils/forbestechcouncil/2019/10/30/why-ai-needs-human-input-and-always-will/>
17. Hu, Q. (2023). Unilever's practice on AI-based recruitment. *Highlights in Business, Economics and Management*, 16, 256–263. <https://doi.org/10.54097/hbem.v16i.10565>
18. Jeon, J. H. (2024). 첨단 인공지능 안전 및 신뢰성 기술 표준 동향 [Standardization trends on safety and trustworthiness technology for advanced AI]. *ETRI Electronics and Telecommunications Trends*, 39(5), 108–122. DOI:10.22648/ETRI.2024.J.390511
19. Köchling, A., & Wehner, M. C. (2020). Discriminated by an algorithm: A systematic review of discrimination and fairness by algorithmic decision-making in the context of HR recruitment and HR development. *Business Research*, 13, 795–848. <https://doi.org/10.1007/s40685-020-00134-w>
20. Kolbjørnsrud, V. (2024). Designing the intelligent organization: Six principles for human-AI collaboration. *California Management Review*, 66(2), 44–64. DOI:10.1177/00081256231211020
21. Korea Ministry of Employment and Labor. (2024). 기업 채용 트렌드 조사 보고서 [Corporate recruitment trend survey report]. Ministry of Employment and Labor.
22. Kristof, A. L. (1996). Person-Organization Fit: An Integrative Review of Its Conceptualizations, Measurement, and Implications. *Personnel Psychology*, 49(1), 1–49. <https://doi.org/10.1111/j.1744-6570.1996.tb01790.x>
23. Kristof-Brown, A., Zimmerman, R., & Johnson, E. (2005). Consequences of individuals' fit at work: A meta-analysis of person-job, person-organization, person-group, and person-supervisor fit. *Personnel Psychology*, 58(2), 281–342. <https://doi.org/10.1111/j.1744-6570.2005.00672.x>
24. Kristof-Brown, A., Schneider, B., & Su, R. (2023). Person-organization fit theory and research: Conundrums, conclusions, and calls to action. *Personnel Psychology*, 76(2), 375–412. <https://doi.org/10.1111/peps.12581>
25. Lam, T., & Bengo, P. (2003). A Comparison of three retrospective self-reporting methods of measuring change in instructional practice. *American Journal of Evaluation*, 24(1), 65–80. <https://doi.org/10.1177/109821400302400106>
26. Langer, M., König, C. J., Sanchez, D. R., & Samadi, S. (2020). Highly automated interviews: applicant reactions and the organizational context. *Journal of managerial psychology*, 35(4), 301–314. <https://doi.org/10.1108/JMP-09-2018-0402>
27. Leadership IQ. (2020). Why new hires fail. Leadership IQ. Retrieved from [https://www.leadershipiq.com/blogs/leadershipiq/35354241-why-new-hires-fail-emotional-intelligence-vs-skills?utm_source\[0\]=Partnerstack&utm_source\[1\]=Partnerstack&utm_medium\[0\]=hakimiabduljabar1378&utm_medium\[1\]=hakimiabduljabar1378](https://www.leadershipiq.com/blogs/leadershipiq/35354241-why-new-hires-fail-emotional-intelligence-vs-skills?utm_source[0]=Partnerstack&utm_source[1]=Partnerstack&utm_medium[0]=hakimiabduljabar1378&utm_medium[1]=hakimiabduljabar1378)
28. Lee, B. C., & Kim, B. Y. (2021). Development of an AI-based interview system for remote hiring. *International Journal of Advanced Research in Engineering and Technology*, 12(3), 654–663. DOI:10.34218/IJARET.12.3.2021.060
29. Leung, A. (2003). Different ties for different needs: Recruitment practices of entrepreneurial firms at different development phases. *Human Resource Management*, 42(4), 303–320. <https://doi.org/10.1002/hrm.10092>
30. Liu, G., Hu, P., Zhong, H., ... & Hu, S. (2024). Effects of the acoustic-visual indoor environment on relieving mental stress based on facial electromyography and micro-expression recognition. *Buildings*, 14, 3122. <https://doi.org/10.3390/buildings14103122>

31. Moon, H., & Nam, S. (2022). The asynchronous video interview–artificial intelligence: A review of current studies and directions for future research. *Korean Journal of Industrial and Organizational Psychology*, 35(3), 385–413. <https://doi.org/10.24230/kjiop.v35i3.385-413>
32. Morville, P. (2004). *User experience honeycomb*. Semantic Studios.
33. Nielsen, J. (1994). *Usability engineering*. Academic Press. <https://doi.org/10.1016/B978-0-08-052029-2.50007-3>
34. O'Reilly, C. A., Chatman, J. A., & Caldwell, D. F. (1991). People and organizational culture: A profile comparison approach to assessing person–organization fit. *Academy of Management Journal*, 34(3), 487–516. <https://doi.org/10.2307/256404>
35. Rocha, V., & Grilli, L. (2023). Early-stage start-up hiring: the interplay between start-ups' initial resources and innovation orientation. *Small Bus Econ*, 62, 1641–1668. <https://doi.org/10.1007/s11187-023-00818-7>
36. SAE International. (2021). Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles (*J3016_202104*).
37. Sheridan, T. B., & Verplank, W. L. (1978). *Human and computer control of undersea teleoperators*. MIT Man–Machine Systems Laboratory Technical Report. <https://doi.org/10.21236/ADA057655>
38. Shneiderman, B. (2020). *Human-centered AI*. Oxford University Press.
39. Shneiderman, B. (2020). Human-centered artificial intelligence: Three fresh ideas. *AIS Transactions on Human–Computer Interaction*, 12(3), 109–124. DOI:10.17705/1thci.00131
40. Schneider, B., Ehrhart, M. G., & Macey, W. H. (2013). Organizational climate and culture. *Annual Review of Psychology*, 64, 361–388. <https://doi.org/10.1146/annurev-psych-113011-143809>
41. Son, S., & Oh, J. (2023). A study on the performances of AI recruitment system: A case study on domestic and abroad companies. *Journal of Career and Entrepreneurship Business Association*, 7(2), 137–155. <https://doi.org/10.48206/kceba.2023.7.2.137>
42. Stryker, C. (2025). *What is human in the loop?*. IBM Think. Retrieved from <https://www.ibm.com/think/topics/human-in-the-loop>
43. Tang, Y., Shao, Y.-F., Chen, Y.-J., & Ma, Y. (2021). How to keep sustainable development between enterprises and employees? Evaluating the impact of person–organization fit and person–job fit on innovative behavior. *Frontiers in Psychology*, 12, 653534. <https://doi.org/10.3389/fpsyg.2021.653534>

스타트업과 지원자 간 문화 적합성 평가를 위한 인간-AI 협업 면접 시스템 설계 연구

송승연¹, 임정은¹, 김예진¹, 남유정¹, 구유리^{2*}

¹홍익대학교 영상커뮤니케이션대학원 인터랙션디자인학과, 학생, 서울, 대한민국

²홍익대학교 산업미술대학원 서비스디자인학과, 교수, 서울, 대한민국

초록

연구배경 초기 스타트업의 채용 실패는 종종 지원자와 조직문화 간의 부조화에서 비롯된다. 스타트업은 체계적인 평가 기준과 평가 경험이 충분히 축적되지 않은 상황에서 면접 과정이 주관화되기 쉬우며, 빠르게 변화하는 조직문화와 인력 요구에 비해 판단 기준이 일관되게 유지되기 어렵다. 본 연구는 이러한 한계를 보완하기 위해, 구조화된 평가 기준과 전문 인력이 부족한 상황에서도 문화 적합성을 효과적으로 판단할 수 있도록 인간-AI 협업 기반 면접 시스템을 제안한다. 제안된 시스템은 스타트업의 유동적인 문화 정체성과 후보자의 가치·행동·조직 적합성을 분석하여 채용 결정을 지원한다.

연구방법 본 연구는 스타트업의 채용 실무자를 대상으로 한 심층 인터뷰를 통해, 조직문화 적합성 평가 과정에서의 주요 니즈를 도출하였다. 이후 문화 적합성 면접의 전체 플로우를 분석하여 과업(Task)별 핵심 행동을 구조화하고, 각 단계에서의 AI 협업 가능성을 진단하였다. 이를 바탕으로 설문 조사를 실시하여 AI의 개입 수준과 역할에 대한 사용자 인식 및 적정성을 평가하였으며, 그 결과를 반영하여 프로토타입을 설계·구현하였다. 마지막으로, 프로토타입 기반으로 사용자 경험 검증(연구 1)과 업무 개선도 정량 평가(연구 2)의 두 차례 테스트를 수행하였다. 연구 1은 대면 테스트를 통해 구체적인 사용성 평가를 실시하였으며, 연구 2는 참여자들의 과거 경험과 프로토타입 사용을 동일 시점에서 회상·비교하도록 설계하였다. 이를 통해 사용성 평가와 업무 개선도 기대효과 평가를 통해 제안된 시스템의 실제 효과성과 협업 경험을 검증하였다.

연구결과 프로토타입 기반의 두 차례의 테스트 및 인터뷰 결과, 자동 질문 생성 등의 기능을 통해 사용자의 인지적·시간적 부담이 줄어들었으며, 결과 해석의 투명성이 부족해 신뢰에는 제한이 있었으나, 공동 작업 기능(예: 편집 가능한 질문 초안 등)은 높은 만족도를 보였다. 특히 맥락에 따라 복잡하고 판단이 어려운 업무에서 인간-AI 협업이 효율성과 의사결정 품질을 높이는 데 유의미하게 기여함을 확인했다. 또한 기존의 조직문화 적합성 중심의 면접 경험에 비해, 시간 효율, 업무 효율, 객관성, 그리고 공정성이 모두 개선될 수 있음이 검증되었다.

결론 본 연구는 조직문화 적합성 면접 과정에 AI를 인간 중심으로 통합하는 설계 지침을 제시한다. 평가 기준을 구조화하고 협업 과정을 투명하게 구현함으로써, 채용 의사결정의 불확실성을 줄이고 스타트업이 조직문화에 부합하는 인재를 구성할 수 있도록 지원한다. 향후 연구에서는 장기적 채용 성과에 대한 검증과 다양한 조직 맥락에서의 확장 가능성을 탐색할 필요가 있다.

주제어 인간-AI 협업, 인간중심 인공지능, 개인-조직 적합도, 조직문화 프로파일, 스타트업 채용

*교신저자: 구유리 (yrkoo@hongik.ac.kr)