Check for updates

Trust in AI-Generated Performance Review Summary Interfaces: Focusing on Feedback Sources and UI Frameworks

Jaerin Lee[1], Juhyun Eune[2*]

[1]Department of Design, Master Course, Seoul National University, Seoul, Korea
[2]Department of Design, Professor, Seoul National University, Seoul, Korea

Abstract

Background Recent HR SaaS (Human Resources Software-as-a-Service) platforms increasingly utilize 360-degree feedback data to evaluate employee competencies and to support talent management decision-making. As a result, artificial inteligence (AI)-powered summarization features have been widely adopted to help users efficiently interpret large volumes of qualitative feedback. However, there remains limited understanding of which summarization approaches are most effective in fostering user trust. This study aims to investigate how different user interface (UI) frameworks influence information trust according to feedback sources.

Methods This study first conducted semi-structured interviews with organizational members to explore the contexts in which performance evaluation data are used and how users perceive the trustworthiness of summary UIs. Subsequently, a 2 (Feedback sources) × 4 (UI frameworks) repeated-measures experiment was conducted, with feedback data and UI framework as independent variables and trust in the information as the dependent variable. The study examined how different combinations of feedback data and UI framework influence users' trust in AI-generated performance feedback summaries.

Results The results showed that trust in AI-generated summaries was not determined by a single UI framework but varied depending on the characteristics of the feedback data being interpreted. When reviewing peer feedback based on multiple opinions, participants relatively preferred quotation-based and chart-based summaries, whereas for manager feedback based on authoritative judgments, they preferred keyword-based and interpretive, explanatory summaries. These findings suggest that AI-based performance feedback summaries should provide UI frameworke that are appropriately matched to the characteristics of the feedback data being reviewed.

Conclusions This study empirically demonstrates that trust in AI summary interfaces for HR performance review data is shaped through the interaction of feedback sources and UI frameworks. Design guidelines for trustworthy AI summary interfaces are proposed based on these findings.

Keywords HR Platform, Performance Review, 360-Degree Feedback, AI Summary, Data Trust, XAI

*Corresponding author: Juhyun Eune (jheune@snu.ac.kr)

Citation: Lee, J., & Eune, J. (2026). Trust in AI-Generated Performance Review Summary Interfaces: Focusing on Feedback Sources and UI Frameworks. *Archives of Design Research, 39*(4), 251-261.

http://dx.doi.org/10.15187/adr.2026.11.39.4.251

Received : Jul. 01. 2026 ; **Reviewed :** Jul. 23. 2026 ; **Accepted :** Sep. 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. 연구의 배경 및 목적

최근 기업의 성과관리는 단순한 평가 결과 기록을 넘어 데이터 기반의 인재관리 및 의사결정을 지원하는 방향으로 발전하고 있다. 특히 HR SaaS 플랫폼의 확산과 함께 동료, 관리자, 협업 부서 등 다양한 이해관계자가 참여하는 다면평가(360-Degree Feedback)가 널리 활용되면서, 조직은 대규모의 정성 피드백 데이터를 확보할 수 있게 되었다. 이러한 데이터는 구성원의 역량 진단, 승진 심사, 인재 배치, 육성 계획 수립 등 다양한 인사 의사결정에 활용될 수 있으나, 실제 현장에서는 방대한 텍스트 피드백을 직접 검토하고 해석해야 하는 부담이 존재한다. 이에 따라 HR 플랫폼들은 이를 효과적으로 해석하고 종합하는 것을 중요한 과제로 바라보고 있으며, 이에 대한 해결 방안으로 AI 기반 요약 기능을 적극적으로 도입하고 있는 추세이다. 하지만 대부분은 기능 구현 자체에 초점을 두고 있으며 평가 출처 등 정보의 성격에 따라 어떠한 UI 표현 방식이 사용자의 정보 신뢰 형성에 효과적인지에 대한 실증적 연구는 충분히 이루어지지 않았다. 이에 본 연구는 HR 성과평가라는 특정 상황을 대상으로, 평가 정보와 요약 UI 표현 방식에 따라 사용자가 정보를 어떻게 인식하고 활용하는지를 탐색하고자 한다. 이를 통해 평가 정보의 신뢰성 인식과 업무적 유용성에 영향을 미치는 정보 표현 방식을 탐색하고, HR 성과평가 환경에서 생성형 AI 기반 요약 인터페이스를 설계하기 위한 UI 디자인 가이드라인을 제안하고자 한다.

2. 이론적 배경

2. 1. AI 기반 성과평가 요약과 설명 가능성

생성형 AI와 대규모 언어모델(LLM)의 발전으로 기업은 방대한 텍스트 데이터를 자동으로 분석하고 요약할 수 있게 되었다. AI 요약의 가치는 단순한 정보 축약을 넘어 평가 결과를 이해하기 쉬운 형태로 재구성하는 데 있다. 선행 연구는 설명이 제공된 AI 시스템이 더 높은 신뢰와 수용성을 형성하며(Miller, 2019), 설명의 양과 방식에 따라서도 사용자가 형성하는 시스템의 이해와 인식이 달라질 수 있음을 제시하였다(Kulesza et al., 2013). 또한 동일한 정보라도 내러티브, 키워드, 차트 등 정보를 표현하는 인터페이스 방식에 따라 사용자의 정보 이해와 활용이 달라질 수 있다(Lim et al., 2021). 성과평가를 요약하는 데 있어서도 마찬가지이다. 특히 성과평가와 같이 다양한 평가자의 정성적 의견이 포함된 정보는 요약 과정에서 핵심 내용이 선택되고, 정보가 압축되거나 특정 내용이 강조될 수 있다. 따라서 동일한 평가 정보를 기반으로 하더라도 어떠한 방식으로 요약하고 제시하는지에 따라 사용자가 인식하는 정보의 맥락과 근거성이 달라질 수 있으며 이는 사용자가 정보를 바라보는 데 있어 느끼는 신뢰도의 차이로 이어질 수 있다. 이것은 AI 요약 인터페이스에서 UI 표현 방식이 단순한 시각적 장식이 아니라, 어떤 정보가 강조되고 어떻게 해석되는지를 결정하는 정보 구성의 요소로 기능할 수 있음을 의미한다. 본 연구는 이러한 관점에서, AI 기반 성과평가 요약 과정에서 UI 표현 방식에 따라 정보가 선택·압축·강조되는 방식이 달라질 수 있으며, 이러한 차이가 사용자의 정보 신뢰와 평가 업무에서의 활용에 영향을 미칠 수 있다고 본다. 이에 본 연구는 UI 표현 방식을 주요 설계요인으로 설정하여, 이에 따른 정보 신뢰 및 활용도(활용 의지)의 차이를 검토하고자 한다.

2. 2. 평가 정보에 따른 정보 신뢰

성과평가 요약 인터페이스는 평가 정보(평가 출처)라는 관점에서 고려될 필요가 있다. 실제 HR 평가 환경에서 평가 내용은 다수 기반 의견과 권위 기반 의견이라는 서로 다른 기준으로 요약되어 활용된다. 다수 기반 의견은 동료평가와 같이, 서로 다른 다양한 관점의 정보를 비교하는 과정(London & Smither, 1995)에서 반복적으로 나타나는 평가 패턴을 통해 신뢰를 형성하는 반면, 권위 기반 의견은 관리자 평가와 같이 조직적 판단을 반영한다(DeNisi & Murphy, 2017). 이에 본 연구는 평가 정보의 차이를 독립변수로 두고, 앞에서 설명한 것처럼 UI 표현 방식의 효과를 탐구하고자 한다. 또한 이 과정에서 일반

구성원, 관리자, 인사 담당자 등 정보를 확인하는 조직 구성원 역할에 따른 차이가 있는지를 확인하고자 한다. 한편, 실제로 성과평가를 확인하고 의사결정에 활용하는 행태에는 평가 출처 외에 조직 환경이라는 변수가 영향을 미칠 수 있다. 평가제도의 특성, 조직문화, 평가 결과의 활용방식 및 평가에 대한 심리적 부담 등이 그 예이다. 본 연구에서는 조직 환경이라는 변수를 제외하고 정보 표현 및 인터페이스 차원의 문제에 연구 범위를 한정하고자 한다.

3. 서비스 사례 분석

현재 기업에서 널리 활용되고 있는 글로벌 HR SaaS 플랫폼 중 AI 기반 성과관리 기능을 제공하는 대표 사례를 선정하여 성과평가 내용 요약 현황을 분석해 보았다. 분석 결과, 현재 HR SaaS 플랫폼들은 동일한 상황에서도 플랫폼마다 다른 방식으로 AI 요약 결과를 제공하고 있다. HR 플랫폼 Lattice는 동료평가를 간단한 키워드로 제시하는 UI를 채택한 반면, 또 다른 HR 플랫폼인 Culture Amp와 Deel은 평가 내용을 자연어 형태로 종합하여 제시하는 내러티브 중심의 UI를 채택했다. 이와 달리 HR 플랫폼 Betterworks는 해당 팀원에 대한 구체적인 인용구를 제시해 인물의 역량을 설명하고 있었다.

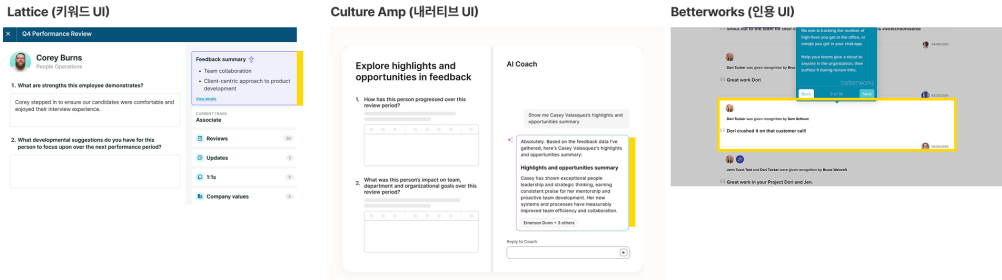

Figure 1 Summary UIs between services

이는 시장에서 공통으로 동의할 수 있는 효과적인 UI가 검증된 바가 없음을 의미하며, 이에 대한 이론적·실증적 연구 결과가 필요하다는 것을 시사한다.

4. 연구방법

4. 1. 인터뷰 연구

본 연구에서는 우선 HR 성과평가 상황에서 사용자가 평가 정보를 활용하는 맥락과 요구를 탐색하기 위해 반구조화 인터뷰를 진행했다. 인터뷰는 성과평가 정보를 실제 업무에서 활용하거나 경험한 조직 구성원을 대상으로 수행하였으며, 일반 구성원 2명, 관리자 3명, 인사 담당자 1명을 모집해 다양한 조직 구성원을 포함하려 했다. 인터뷰는 약 40분간 비대면 반구조화 방식으로 진행되었다. 해당 인터뷰 결과는 사용자별로 실제 상황에서 성과평가 내용을 확인하는 시나리오를 구체화하고 그 과정에서 발생하는 문제를 발견하며, 선호하는 요약 UI 방식을 탐색하는 데 활용했다.

Table 1 Interviewees

	직급	직무	회사규모	인터뷰 내용	해당 경험
P1	팀원	온라인 MD	소기업 (화장품)	피평가자 경험	만 6년
P2	팀원	마케팅	중견기업 (패션)	피평가자 경험	만 4년
P3	팀장	기획	대기업 (여행·IT)	하향평가 경험	만 3년
P4	팀장	디자인	중견기업 (티켓·IT)	하향평가 경험	만 2년
P5	팀장	개발	중견기업 (패션·IT)	하향평가 경험	만 1년
P6	인사담당자	인사	대기업 (바이오)	평가정보 확인 경험	만 8년

4. 2. 실험 설계

본 실험으로는 평가 정보에 따른 UI 표현 방식의 효과를 정량적으로 검증하기 위한 2(평가 출처 기준 : 동료평가, 관리자 평가) × 4(UI 표현 방식 : 뒤에서 설명)의 반복측정 실험을 수행하였다. 독립변수는 평가 정보와 UI 표현 방식이며, 종속변수는 AI가 생성한 성과평가 요약에 대한 신뢰도이다. 여기서 신뢰도는 실제 정보 반영도, 의사결정 확신도 두 항목을 7점 리커트 척도로 측정 및 평균을 합산했는데, 구체적으로 실제 정보 반영도는 AI가 생성한 요약정보가 원래의 평가 내용을 적절하게 반영하고 있다고 생각하는 정도를, 의사결정 확신도는 이를 평가 업무 및 판단에 활용할 수 있다고 인식하는 정도를 응답하도록 했다. 즉 신뢰도는 실제 정보 반영도 및 정보 활용을 포함하는 종합적인 유용성의 개념으로 분석되었다는 점을 밝힌다. 실험에는 인사담당자 27명, 관리자 20명, 일반 구성원 23명 등 총 70명이 참여하였다. 참여자들은 앞선 인터뷰 결과를 바탕으로 구성된 역할별 시나리오에 몰입한 후 8개의 AI 요약 UI를 무작위 순서로 확인하고 각 자극물에 대한 신뢰도 항목에 응답하였다. 실험 데이터는 반복측정 분산분석(Repeated Measures ANOVA)을 통해 분석하였다.

연구 질문은 다음과 같다.

- RQ1. 평가 정보(동료평가 및 관리자 평가)와 UI 표현 방식의 조합은 신뢰도에 어떠한 영향을 미치는가?
- RQ2. 평가 정보에 따라 적합한 성과평가 요약 UI는 무엇인가?
- RQ3. 사용자 역할에 따라 성과평가 AI 요약에 대한 신뢰 형성 방식이 달라지는가?

4. 3. 2x4 실험 자극물

실험 자극물 제작을 위해 생성형 AI를 활용해 가상의 조직 구성원에 대한 동료평가와 관리자 평가 데이터를 만들었다. 실제와 가까운 데이터를 생성하기 위해 구성원의 소속 조직, 진행했던 프로젝트 및 협업했던 동료 등 배경 정보를 AI에 학습시켜 생성했으며, 구성원의 장점과 단점이 해당 배경 정보를 바탕으로 기술되도록 했다. 동료평가는 구성원에 대한 동료 11명의 평가 내용으로 이루어졌으며, 관리자 평가는 구성원에 대한 관리자 1명의 평가 내용으로 이루어졌다. 성과평가 내용을 요약하는 UI 표현 방식은 서비스 사례 분석과 선행연구, 그리고 인터뷰 내용을 바탕으로 네 가지 방식을 선정하였다(Table 2).

Table 2 UI Frameworks

내러티브 UI	문장 중심의 서사 형태로 맥락과 의미를 제공, 가장 해석적인 설명 제시 (Segel & Heer, 2010)
키워드 UI	핵심 내용을 압축해 빠른 탐색과 객관적인 정보 파악을 지원 (Hearst, 2009)
인용 UI	요약 결과와 함께 원문 발화를 제시하여 AI 판단의 근거를 확인할 수 있는 높은 설명 가능성을 제공 (Kulesza et al., 2013)
차트 UI	수치 정보를 시각적 인코딩으로 표현함으로써 정보 간 상대적 크기와 차이를 보다 정확하게 비교할 수 있도록 지원하는 표현 방식 (Cleveland & McGill, 1984)

평가 정보별, UI 표현 방식 별 실험 자극물의 구체적인 요약 UI는 다음과 같다. 우선 내용을 요약하는 기준은 동료평가의 경우 많은 사람들이 언급한 내용이 우선순위로, 관리자 평가의 경우 관리자가 많이 언급한 내용이 우선순위로 작용했다. 내러티브 UI 방식은 구성원에 대해 언급된 장점과 단점을 설명하고, 마지막으로 구성원을 업무 역량(업무 스킬 및 정량적 결과)과 리더십(커뮤니케이션 및 성실도 등) 측면에서 분석하거나 어떤 능력이 돋보이는 구성원인지를 AI가 해석한 결과를 반영해 문장 형태로 설명하도록 했다. 길이가 길지만 맥락적이고 해석적 설명이 들어갔다.



Figure 2 Narrative UI

키워드 UI 방식에서는 장점과 단점을 짧은 키워드 형태로 예시와 함께 제시했다.

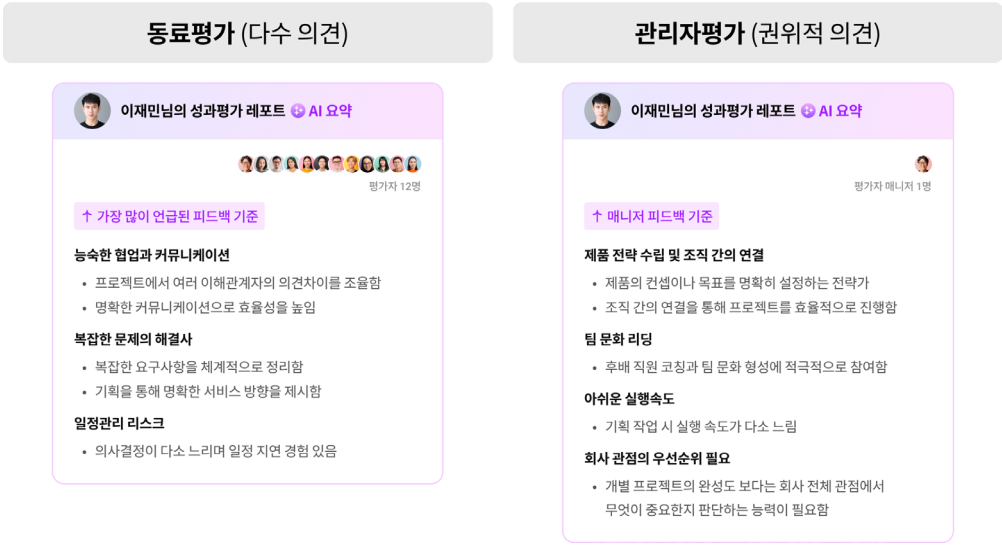

Figure 3 Keyword UI

인용 UI 방식에서는 장점과 단점에 대한 예시를 원문 데이터 가공 없이 인용구 형태로 그대로 보여줬다. 작성한 평가자와 원문 데이터를 보여주는 유일한 방식이었다.

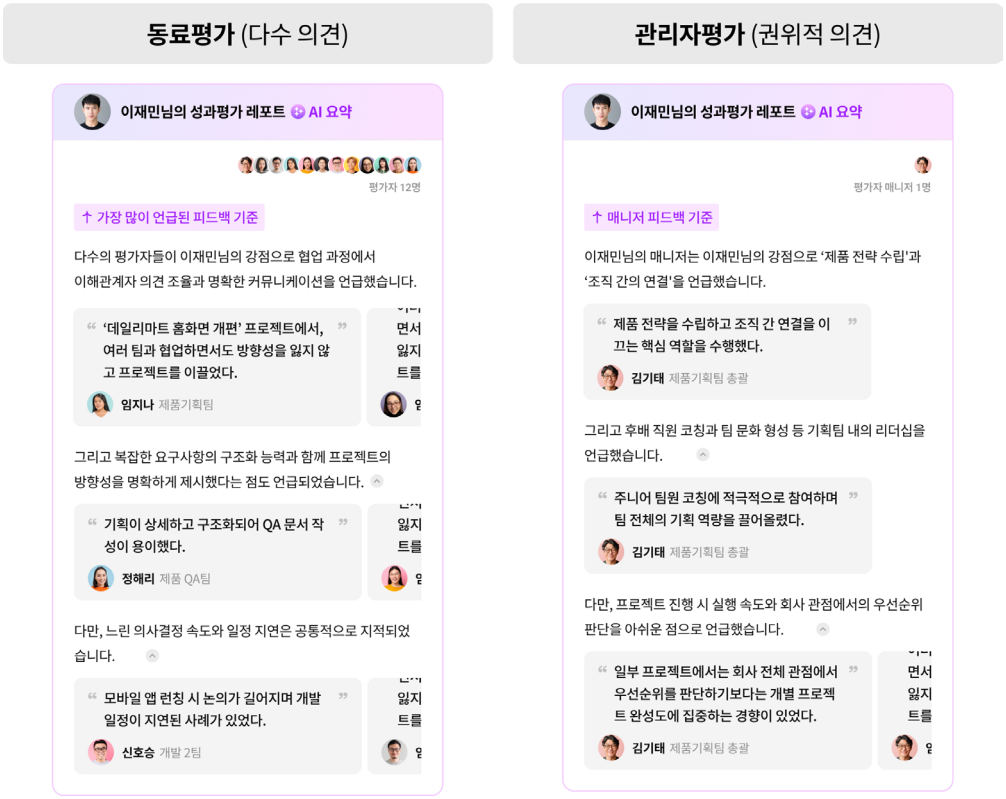

Figure 4 Quotation UI

차트 UI 방식에서는 많이 언급된 키워드 순으로 보여준다는 점은 다른 UI 방식들과 동일했지만 언급된 횟수를 전체 언급 수 대비 비율의 차트 형태로 보여줘 가장 직관적으로 우선순위를 파악 가능하게 했다. 다만 그 외 원문 데이터 등 구체적인 예시가 일절 없었다.

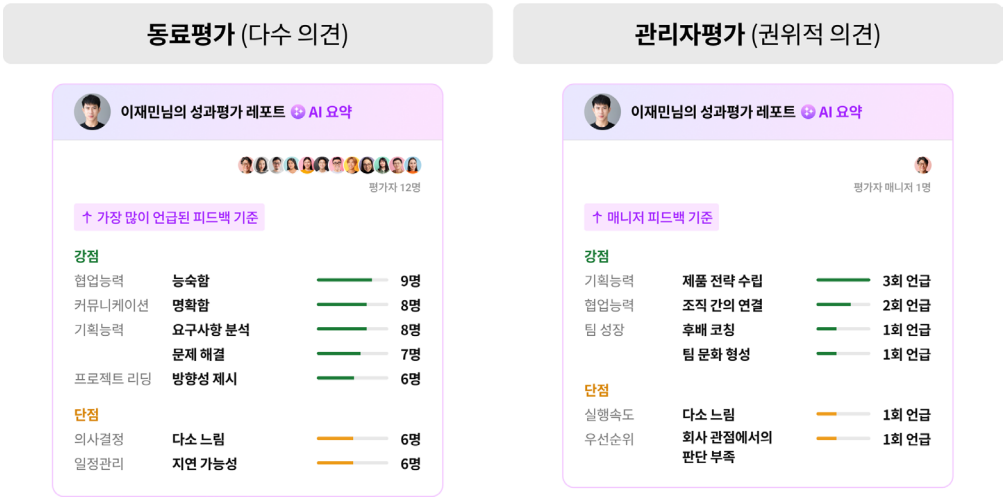

Figure 5 Chart UI

각 평가 정보별로 다루는 원문 데이터를 통제 변수로 설정해 최대한 해당 정보를 요약하는 UI에 따른 실험 참여자 반응 차이만을 확인할 수 있도록 했지만, 각 UI 표현 방식의 형식적 차이로 인해 보이는 정보의 양과 뉘앙스가 달라질 수밖에 없다는 어쩔 수 없는 차이는 존재했다.해당 UI 표현 방식들은 실제 상황에서 반드시 단독으로 사용될 필요는 없으며, 필요시 두개 이상의 표현 방식을 결합해 사용할 수 있다(예 : 키워드 UI를 기본으로 부가 클릭 등을 통해 인용 UI를 확인할 수 있는 요약 인터페이스). 해당 연구에서는 각 UI 표현 방식에 대한 사용자 반응을 통해 우선적으로 제공해야 하는 원칙과 가이드라인을 제공하는 것을 목표로 한다.

5. 연구결과

5. 1. 인터뷰 결과

반구조화 인터뷰를 통해 성과평가 정보가 실제 업무에서 활용되는 상황과, 평가정보를 이해하고 판단하는 과정에서 중요하게 인식되는 정보 표현 요소를 탐색하였다. 인터뷰 분석 결과, 참여자들이 사용자 역할에 따라 평가정보를 활용하는 구체적인 상황에는 차이가 있었으며, 공통적으로 AI가 요약한 평가정보를 그대로 수용하기보다 확인하는 평가 정보와 활용 맥락에 따라 선호하는 요약 정보 표현 방식이 다른 것을 확인할 수 있었다. 일반 구성원은 주로 성과평가 결과 공개 기간에 자신의 평가 내용을 확인하는 상황을 경험하였다. 이들은 본인에 대한 동료평가를 확인할 때에는 부정적 피드백이나 구체적인 평가자의 표현을 직접 확인하는 것에 부담을 느끼며, 여러 평가의 핵심적인 내용과 주요 경향을 빠르게 파악할 수 있는 형태를 선호하였다(“누가 어떤 말을 했는지 구체적으로 알고 싶지 않다. 어떤 내용인지만 대략 알려줬으면 한다”- P1). 반면 본인에 대한 관리자 평가를 확인할 때에는 결과의 근거를 구체적인 설명을 통해 확인하고자 하는 요구를 나타냈다.

관리자는 주로 성과평가 기간에 팀원의 동료평가를 최종 평가를 위한 참고자료로 활용하거나, 조직 재배치 과정에서 신규 팀원의 이전 관리자 평가를 확인하는 상황을 경험하였다. 팀원의 동료평가와 같이 여러 평가자의 의견을 종합한 정보를 활용할 때에는 AI가 제시한 요약의 근거와 평가 맥락을 확인하는 것이 중요하다고 인식하였다(“해당 팀원과 사적으로 얼마나 친밀한지, 그리고 실제 함께 진행했던 프로젝트가 무엇인지 확인해 편향성과 객관성을 따져본다” - P3). 반면 이전 관리자의 평가와 같이 평가자가 명확하게 확인되는 정보에 대해서는 세부적인 발화 내용보다 핵심적인 평가 내용을 우선적으로 파악하고자 하였다.

인사 담당자는 주재원 선정, 승진 결정, 연봉 조정 및 교육 프로그램 대상자 선정 등 조직 차원의 의사결정을 위해 구성원의 동료평가와 관리자 평가를 확인하는 상황을 경험하였다. 이들은 대체로 핵심 내용 위주로 빠르게 확인하고 싶어했으나 동료평가를 확인할 때에는 내용의 이해보다는 내용의 빈도와 경향성 위주로 파악하고 싶어하는 모습을 보였다.

Table 3 Analysis of Interview result

사용자 역할	성과평가 내용 확인 시나리오	요약 UI에 대한 신뢰 형성 요인
일반 구성원	성과평가 결과 확인 기간, 본인에 대한 동료평가 내용 확인 및 관리자의 평가 확인	확인하는 평가 정보에 따라 선호하는 요약 방식이 다름
관리자	하향평가 진행 시 팀원에 대한 동료평가 참고, 조직 재배치 시 신규 팀원에 대한 이전 관리자 평가 참고	
인사 담당자	승진 결정, 연봉 조정 등의 결정 상황에서 조직원의 동료 및 관리자 평가 참고	

5. 2. 실험 결과

앞선 인터뷰 결과를 바탕으로, 평가정보의 출처에 따른 정보 특성에 주목하여 다양한 UI 표현 방식과 결합될 때 정보 신뢰에 어떠한 차이가 나타나는지를 검증하기 위한 2×4 반복측정 실험을 수행하였다. 실험 결과, 평가 정보와 UI 표현 방식 간 상호작용 효과가 유의한 결과를 보였으며($F=6.81$, $p=.001$), 반면 사용자 그룹과

관련된 상호작용 효과는 통계적으로 유의하지 않은 것으로 나타났다(F=1.70, p=.120) 이는 동료평가와 같이 여러 평가자의 의견을 종합한 정보를 확인하는 상황인지, 관리자 평가와 같이 특정 권위자의 판단을 중심으로 구성된 정보를 확인하는 상황인지에 따라 동일한 UI 표현 방식이라도 사용자가 형성하는 신뢰도가 달라질 수 있다는 것을 의미하며, 이 과정에서 정보를 확인하는 사용자 역할별로 달라지는 차이는 크게 없다는 것을 의미한다.평가 정보별 UI 표현 방식에 대한 신뢰도 차이를 보자면, 동료평가를 확인하는 상황에서는 인용 UI(M=5.638)와 차트 UI(M=5.533)의 평균 신뢰도가 상대적으로 높게 나타났으며, 관리자 평가를 확인하는 상황에서는 키워드 UI(M=5.548)와 내러티브 UI(M=5.400)의 평균 신뢰도가 상대적으로 높게 나타났다. 신뢰도의 하위 항목이었던 실제 정보 반영도와 의사결정 확신도로 세분화해서 결과를 분석했을 때도 동일한 결과를 보였다. 이는 다수의 의견을 요약한 동료평가의 경우 다수라는 특성에서 오는 익명성 때문에 구체적인 발화자 및 내용을 확인하고자 하는 요구와 전체적인 패턴을 가장 직관적으로 분석할 수 있는 방식에 대한 요구가 공존하는 반면, 관리자 평가라는 권위를 기반으로 한 확실한 데이터 소스가 있는 경우 요약 근거 확인에 대한 요구보다는 핵심 요약과 질적 해석에 대한 요구가 공존한다는 것으로 해석할 수 있다.

Table 4 Result on frequency based feedback(peer review) (mean) (n=70)

	1. 실제 정보 반영도	2. 의사결정 확신도	신뢰도 (1+2)
내러티브 UI	5.44	5.47	5.45
키워드 UI	5.49	5.21	5.40
인용 UI	**5.63**	**5.66**	**5.64 (1순위)**
차트 UI	**5.52**	**5.56**	**5.53 (2순위)**

Table 5 Result on authority based feedback(manager review) (mean) (n=70)

	1. 실제 정보 반영도	2. 의사결정 확신도	신뢰도 (1+2)
내러티브 UI	**5.36**	**5.47**	**5.40 (2순위)**
키워드 UI	**5.49**	**5.266**	**5.55 (1순위)**
인용 UI	5.39	5.39	5.39
차트 UI	5.08	5.13	5.10

종합해 보자면, 인터뷰 결과에 따르면 사용자 역할에 따라 성과평가 정보를 활용하는 목적은 다르게 나타났으나, 실험 결과를 보았을 때 정보 신뢰를 형성하는 기본 원리는 확인하는 평가 정보의 출처(동료평가 및 관리자 평가)에 따라 달라짐을 확인할 수 있었다. 이는 평가 정보에 따라 적합한 요약 UI 표현 방식을 선택하는 것이 중요함을 시사한다.

6. 평가 출처 중심 성과평가 AI 요약 UI 가이드라인

연구 결과를 바탕으로, 평가 정보를 중심으로 HR 성과평가 AI 요약 인터페이스 설계를 위한 두 가지 디자인 원칙을 제안한다. 여기에서 성과평가는 정성적인 평가 내용을 전제로 한다.

6. 1. 가이드라인 1. 동료평가를 확인하는 상황에서는 평가 근거의 투명성을 강화하는 동시에 경향성을 아주 직관적으로 파악할 수 있도록 한다.

동료평가는 다수의 의견의 종합으로, 평가내용의 근거를 제공하는 것이 중요하게 작동하는 정보이다. 예를 들어 관리자는 성과평가 기간에 자신이 직접적으로 알기 어려운 환경(협업 경험 혹은 이전 팀에서의 업무 경험)에서의 팀원의 역량을 참고해 팀원에 대해 올바른 최종 평가를 내리거나, 평상시에 팀원의 성장을 돕기 위해 이전 평가 내용을 활용한다. 이 경우 AI 요약을 전적으로 믿고 의사결정하기에 부담이 있기 때문에, 인용을 활용하여(인용 UI) 실제 평가 내용과 평가 작성자의 신분 등 정보의 맥락을 함께 제공하는 것이

바람직할 수 있다. 또한 빈도와 경향성을 한눈에 파악할 수 있도록 정보를 함께 제공하면(차트 UI) 더욱 효과가 좋을 것이다. 실험에서는 사용자 역할 별로 유의한 차이는 없었지만, 인터뷰 결과 일반 구성원의 경우 실험 결과와 반대로 본인에 대한 동료 평가를 확인할 때 너무 구체적인 정보를 확인하는 것에 대한 심적 부담감을 언급한 것처럼, 사용자 역할별로 상황에 따른 차이가 있을 수 있기 때문에 사용자 맥락을 고려해 요약 UI를 다르게 보여주는 것도 고려해볼 수 있다.

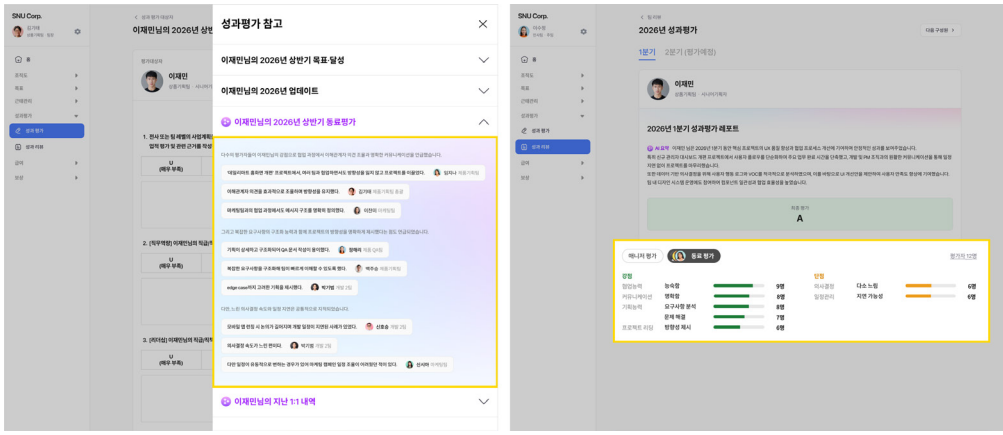

Figure 6 Peer review summary example (Quotation UI(Left), Chart UI(Right))

6. 2. 가이드라인 2. 관리자 평가를 확인하는 상황에서는 핵심 내용을 효과적으로 제공하는 방법을 고민한다.

관리자 평가는 관리자라는 실명 기반의 신뢰가 이미 형성된 정보이다. 따라서 요약 내용의 근거를 제공하기보다는 관리자가 작성한 긴 글의 내용에서 핵심을 일목요연하게 제공하는 것이 중요할 수 있다. 기본적으로는 불렛 형태의 키워드로 요약한 형식(키워드 UI)을 주 방식으로 활용하고 내러티브 형식은 보조적인 설명으로 함께 제공하는 것을 추천한다. 키워드 UI는 예를 들어 특히 인사 담당자처럼 승진 심사, 인재 선발, 조직 진단 등 조직 차원의 의사결정을 위해 여러 구성원에 대한 관리자 평가를 빠르게 확인해야 하는 입장에 최적화된 경험을 제공할 수 있다. 그리고 내러티브 UI는 일반 구성원의 입장에서 본인의 관리자가 지적하거나 긍정적으로 평가한 내용을 업무에 반영해야 하기 때문에 AI가 다시 한번 해석해준 내용으로(내러티브 UI) 본인이 이해한 바를 다시 확인하고 싶어할 수 있다는 점에서 도움이 될 수 있다.

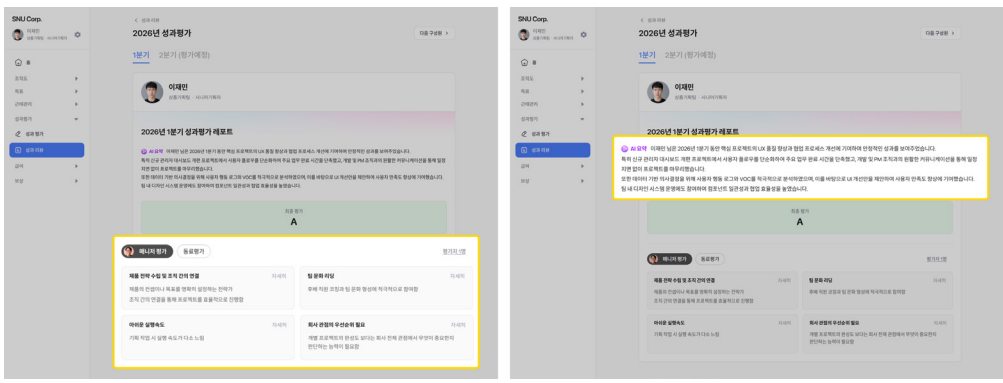

Figure 7 Manager review summary example (Narrative UI(Left), Keyword UI(Right))

7. 결론

본 연구는 HR 성과평가 데이터의 AI 요약 인터페이스를 대상으로 평가 정보와 UI 표현 방식의 조합이 정보 신뢰 형성에 미치는 영향을 탐색하였다. 반구조화 인터뷰, 반복측정 실험 데이터의 분석을 통합한 결과, AI 요약에 대한 신뢰는 특정 UI 표현 방식에 의해 일관되게 결정되지 않았으며, 평가 정보의 출처에 따라 달라지는 것으로 나타났다. 다수 의견을 기반으로 한 동료평가에서는 다수라는 익명성으로 인해 인용 기반의 근거 제시를 중요하게 인식하였으며, 동시에 다수 의견을 직관적인 통계 방식으로 제시하는 것을 선호했다. 반면 권위적 의견을 기반으로 한 관리자 평가에서는 내용에 대한 근거가 동료평가만큼 중요하지 않은 상황인 만큼, 핵심을 정확하게 이해할 수 있는 효율적이면서 질적인 요약 방식을 선호하는 것을 발견했다. 해당 연구는 HR 플랫폼에서 신뢰 형성에 효과적인 AI 요약 UI에 대해 실증적으로 검증했다는 데에 의미가 있다. 특히 최종적으로 제안한 디자인 가이드라인은 앞선 '2.3 서비스 사례 분석'에서 지적한, 모든 HR 플랫폼이 공통으로 활용할 수 있는 검증된 패턴이 부재하다는 문제를 해결하고 실험 결과를 바탕으로 실무적 가이드라인을 제공한다는 점에서 의미가 있다.한편 본 연구는 조직 구조 및 평가 환경 등 다양한 환경적 요인을 고려하지 않고 제한된 환경에서 실험을 진행했다는 점에서 결과를 모든 조직 환경에 일반화하기 어렵다는 한계가 있다. 따라서 향후에는 실제 다양한 조직 구조 및 평가 환경을 분석하고 해당 상황에 최적화된 요약 경험을 제공하기 위한 후속 연구가 필요하다. 이러한 후속 연구는 상황과 사용자 특성에 더욱 적합한 AI 요약 인터페이스 설계 원칙을 발전시키는 데 기여할 것으로 기대된다.

References

1. Cleveland, W. S., & McGill, R. (1984). Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association, 79*(387), 531–554. https://doi.org/10.1080/01621459.1984.10478080
2. DeNisi, A., & Murphy, K. R. (2017). Performance appraisal and performance management: 100 years of progress? *Journal of Applied Psychology, 102*(3), 421–433. https://doi.org/10.1037/apl0000085
3. Hearst, M. A. (2009). *Search user interfaces*. Cambridge University Press. https://doi.org/10.1017/CBO9781139644082
4. Kulesza, T., Stumpf, S., Burnett, M., Yang, S., Kwan, I., & Wong, W.-K. (2013). Too much, too little, or just right? Ways explanations impact end users' mental models. *2013 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*, 3–10. https://doi.org/10.1109/VLHCC.2013.6645235
5. Lim, S., Park, C., Han, H., Ho, J., Hong, J., Lee, S., & Lee, K. (2021). A study on interface design to provide effective meeting summary information: Focused on narrative chart and topic map visualization. *Archives of Design Research, 34*(1), 115–131. https://doi.org/10.15187/adr.2021.02.34.1.115
6. London, M., & Smither, J. W. (1995). Can multi-source feedback change perceptions of goal accomplishment, self-evaluations, and performance-related outcomes? Theory-based applications and directions for research. *Personnel Psychology, 48*(4), 803–839. https://doi.org/10.1111/j.1744-6570.1995.tb01782.x
7. Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence, 267*, 1–38. https://doi.org/10.1016/j.artint.2018.07.007
8. Segel, E., & Heer, J. (2010). Narrative visualization: Telling stories with data. *IEEE Transactions on Visualization and Computer Graphics, 16*(6), 1139–1148. https://doi.org/10.1109/TVCG.2010.179

HR 성과평가 AI 요약 인터페이스에서의 신뢰 형성 연구: 평가 정보 및 UI 표현 방식을 중심으로

이재린[1], 윤주현[2*]

[1]서울대학교 디자인학부 시각디자인과, 석사과정, 서울, 대한민국

[2]서울대학교 디자인학부 시각디자인과, 교수, 서울, 대한민국

초록

연구배경 최근 HR SaaS 플랫폼은 다면평가 데이터를 활용하여 구성원의 역량을 평가하고 인재관리 의사결정을 지원하고 있다. 이에 따라 대량의 정성 피드백을 효율적으로 해석하기 위한 AI 기반 요약 기능이 확산되고 있으나, 어떠한 요약 방식이 효과적인지에 대한 인사이트가 부재한 상황이다. 이에 본 연구는 평가 정보에 따라 어떠한 요약 UI가 정보 신뢰 형성에 효과적인지를 규명하고 디자인 가이드라인을 제공하는 것을 목적으로 한다.

연구방법 본 연구에서는 조직 구성원들을 대상으로 반구조화 인터뷰를 수행하여 성과평가 데이터 활용 시나리오와 요약 UI에 대한 신뢰 인식을 탐색했고, 이후 2(평가 정보) × 4(UI 표현 방식) 반복측정 실험을 진행해 평가 정보 및 요약 UI를 독립변수로, 정보에 대한 신뢰도를 종속변수로 설정해 평가 정보와 UI 표현 방식의 조합이 신뢰도에 미치는 영향을 분석하였다.

연구결과 연구 결과, AI 요약에 대한 신뢰는 단일한 UI 표현 방식에 의해 결정되지 않았으며, 해석하는 정보의 특성에 따라 달라지는 것으로 나타났다. 다수 의견 기반인 동료평가를 확인할 때에는 인용 기반 요약 및 차트 기반 요약을 상대적으로 선호하였으며, 권위적 의견 기반인 관리자 평가를 확인할 때에는 키워드 요약과 해석적이고 설명적인 요약을 선호하는 것을 확인할 수 있었다. 이는 성과평가 AI 요약 설계 시 확인하는 평가 내용을 고려해 적합한 UI 제공해야 함을 시사한다.

결론 본 연구는 HR 성과평가 AI 요약 인터페이스에서 정보 신뢰가 평가 정보 및 UI 표현 방식의 상호작용 속에서 형성됨을 실증적으로 확인하였으며 이를 반영한 디자인 가이드라인을 제안하였다.

주제어 HR 플랫폼, 성과평가, 다면평가, AI 요약, 정보 신뢰, XAI

*교신저자 : 윤주현 (jheune@snu.ac.kr)