Journal of Digital Art Engineering & Multimedia Vol.13, No.3, September (2026), pp.471-478
ISSN 2508-9099(p) 2672-1198(e) *http://dx.doi.org/10.29056/jdaem.2026.09.03*

자율형 AI에이전트의 다단계 작업 수행 과정에서 사용자의 통제감 유지를 위한 UX 가이드라인 연구

Research on UX Guidelines for Maintaining User Control During Autonomous AI Agent Multi-Step Task Execution

이영주[1]*

Young-Ju Lee[1]*

요 약

본 연구는 자율형 AI 에이전트가 사용자의 명시적 개입 없이 복잡한 다단계 작업을 스스로 판단하고 수행하는 과정에서 발생하는 사용자의 불안을 완화하고 만족도를 높이기 위한 UI/UX 가이드라인을 제안하고자 한다. AI 에이전트의 자율성이 극대화될수록 사용자는 고도화된 '블랙박스 현상'으로 인해 시스템에 대한 통제감을 상실하기 쉬우며, 이는 신뢰도 저하와 서비스 이탈로 이어질 수 있다. 이를 해결하기 위해 본 연구에서는 인터럽트 권한의 부여 시점과 시각적 피드백의 수준을 독립변수로 설정하고, 이에 따른 사용자 통제감, 인지적 부하, 시스템 신뢰도의 변화를 측정하기 위해 피험자 60명을 대상으로 이원 반복측정 분산분석을 실시하였다. 실험 결과, 에이전트의 작업 단계별 전환점 및 불확실성이 임계치를 넘어서는 시점에 즉각적인 인터럽트 권한을 제공하고, 단계적 정보 심층화 기반의 구조화된 상태 추적 피드백을 제공할 때 통제감이 가장 높게 유지됨을 규명하였다. 본 연구의 결과는 단순히 알고리즘의 정밀도를 높이는 것을 넘어, 자율형 AI 에이전트 시스템 설계 시 필수적으로 고려해야 할 3대 UI/UX 설계 프레임워크를 정립함으로써 HCI 분야에 정량적이고 실무적인 지침을 제공할 것이다.

핵심어 : 자율형 AI 에이전트, 통제감, 시각적 피드백, 인지 부하, 사용자 경험

Abstract

This study aims to propose UI/UX guidelines to alleviate user anxiety and enhance satisfaction arising from autonomous AI agents independently judging and executing complex multi-step tasks without explicit user intervention. As the autonomy of AI agents is maximized, users are prone to losing a sense of control over the system due to the advanced black box phenomenon, which can lead to decreased trust and service abandonment. To address this, this study established the timing of interrupt authorization granting and the level of visual feedback as independent variables. A two-way repeated measures ANOVA was conducted on 60 subjects to measure changes in user control, cognitive load, and system trust resulting from these factors. The experimental results revealed that the highest level of control was maintained when immediate interrupt authorization was provided at the agent's task transition points and when uncertainty exceeded a threshold, and when structured state tracking feedback based on stepwise information deepening was

1 Department of Multimedia, Chungwoon University, Incheon, Korea [Professor]
e-mail: yjlee@Chungwoon.ac.kr

Received(July 6, 2026), Review Result(1st: August 2, 2026), Accepted(September 9, 2026), Published(September 30, 2026)

provided. Beyond merely improving algorithmic precision, the findings of this study will provide quantitative and practical guidelines for the HCI field by establishing three essential UI/UX design frameworks that must be considered when designing autonomous AI agent systems.

Keyword : Autonomous AI agent, Sense of control, Visual feedback, Cognitive load, UX

1. 서론

1.1 연구의 배경 및 목적

인공지능 기술의 패러다임이 단순한 단발성 프롬프트 반응형 인터페이스에서, 하위 작업을 스스로 생성하고 도구를 선택하여 연쇄 워크플로우를 완수하며 고차원적 목표를 달성해나가는 자율형 AI 에이전트로 급격히 전환되고 있다. 이러한 에이전트 시스템은 데이터 수집, 정제, 모델링, 최종 실행에 이르는 복잡한 전 과정을 독립적으로 처리함으로써 인간의 오퍼레이션 비용을 절감하지만, 인간 사용자는 AI가 내부적으로 어떠한 논리와 절차를 거쳐 과업을 수행하고 있는지 실시간으로 파악하기 어려워지는 고도화된 블랙박스 현상에 직면하게 된다. HCI 영역에서 사용자의 통제감은 시스템에 대한 신뢰성과 지속 이용 의도를 결정하는 최우선적 심리 변수이다 [1][2]. 사용자가 시스템의 작동 메커니즘을 예측하지 못하거나 원치 않는 방향으로 진행될 때 제어할 수 없다고 느낀다면 극심한 인지적 소외감과 불안이 유발된다 [3]. 따라서 자율형 AI의 특성상 긴 실행 시간 동안 사용자가 주도권을 완전히 상실하지 않도록 인터럽트 권한의 시점과 정보 과부하를 유발하지 않는 최적의 시각적 피드백 체계를 구축하는 연구가 절실히 요구되는 시점이다. 이에 본 연구는 자율형 AI 에이전트 환경에서 사용자의 통제감을 유지하고 인지적 부하를 최소화할 수 있는 UI/UX 설계 요소를 실증 분석을 통해 규명하고, 실제 개발 현장에 적용 가능한 정량적 디자인 가이드라인을 정립하는 것을 목적으로 한다.

1.2 연구의 범위와 방법

본 연구는 데이터 소스 수집 및 다단계 정량 분석 보고서를 작성하는 업무 자동화 자율형 AI 에이전트 프로토타입을 연구 범위로 설정한다. 문헌 연구를 바탕으로 인터럽트 시점의 3가지 유형과 시각적 피드백 수준의 3가지 유형을 교차 설계한 3 X 3 요인실험 환경을 구축하였다. 실험은 디지털 환경에 숙련된 60명을 피험자로 모집하여 무작위 배정 및 반복 측정을 수행하였으며, 종속변수인 통제감 점수, NASA-TLX 기반 인지적 부하 점수, 시스템 신뢰도를 정량적으로 측정하였다 [4][5]. 수집된 데이터는 SPSS 26.0 통계 패키지를 활용하여 이원 반복측정 분산분석 및 회귀분석을 실시하였고, 사후 정성 인터뷰 데이터를 코딩 변환하여 양적 결론을 보완하였다.

2. 이론적 배경

2.1 자율형 AI 에이전트와 사용자 통제감의 역학

자율형 AI 에이전트는 LLM 기반의 ReAct 루프나 트리 구조의 사유 방식을 통해 계획을 동적으로 수정하며, 시스템의 주체성이 강화될수록 인간의 주체성은 상대적으로 약화된다 [6]. 노먼(Donald Norman)의 행위 이론에 따르면 사용자는 실행과 평가의 간극을 좁힐 수 있을 때 높은 통제감을 갖는다 [7]. 자율형 에이전트는 인간의 실행 단계를 대행하여 실행의 간극은 줄여주지만, 내부 연산을 은닉함으로써 평가의 간극을 비대하게 넓히므로 연산 흐름을 인지적으로 동기화할 수 있는 인터페이스가 필수적이다 [8].

2.2 인터럽트 권한의 시점 분류

인터럽트 권한이란 시스템의 특정 연산 경로를 정지, 변경, 스킵, 전면 취소할 수 있도록 제공되는 제어 메커니즘이다 [9]. 본 연구에서는 혼합 주도형 인터랙션 이론을 바탕으로 이를 세 가지 유형으로 분류된다 [10]. 즉, 전 과정 동안 제어 버튼이 상시 활성화되어 즉각 개입이 가능한 실시간 항상 개입형(CI), 거시적 하위 작업 완료 시점마다 실행을 멈추고 사용자의 승인을 기다리는 단계 전환점 개입형(MI), 평상시에는 자율 구동되다 내부 판단 신뢰도가 임계치 미만일 때만 개입을 요청하는 불확실성 임계치 개입형(AI)으로 나뉜다.

2.3 시각적 피드백과 정보 가시성의 수준

시각적 피드백은 블랙박스 내부의 다단계 추론 경로를 가시화하여 투명성을 확보하는 수단이다 [6][11]. 이는 정보 밀도와 표현 구조에 따라 세 가지 방식으로 분류된다 [11]. 첫째, 진행률 바 형태로 현재 단계만 간략히 노출하는 최소 요약형(MS), 둘째, 시스템 프롬프트와 API 호출 원시 데이터를 터미널 창 형태로 스트리밍하는 실시간 로그형(RL), 셋째, 목표 수립부터 검증까지의 연쇄 과정을 유향 그래프나 타임라인 노드로 도식화하는 구조화된 상태 추적형(ST)이 있다.

3. 실증 분석 및 실험 설계

3.1 실험 설계 및 절차

본 연구는 앞서 기술한 인터럽트 시점(CI, MI, AI)과 시각적 피드백 수준(MS, RL, ST)의 조합에

따른 사용자 경험적 효과를 검증하기 위해 3 X 3 피험자 내 요인설계를 적용하였다. 피험자는 국내 IT 및 기획 직군 종사자 성인 60명을 대상으로 하였으며 남성 32명, 여성 28명으로 평균 연령은 32.4세이다. 다단계 연쇄 작업이 수행되도록 해외 전기차 시장 데이터 수집 및 특허 분석 보고서 초안 작성을 실험 과업을 제시하고 주관적 통제감(HCI 통제감 척도, α = .89), 인지적 부하(NASA-TLX), 시스템 신뢰도(John Deer 척도 적용)를 측정 도구로 사용하였다 [12][13].

3.2 양적 통계 분석 결과

3.2.1 주관적 통제감과 시스템 신뢰도 분석

수집된 데이터의 상호작용 효과를 검증하기 위해 이원 반복측정 분산분석을 실시하였다. 분석 결과, 인터럽트 시점과 시각적 피드백 형태 간의 강력한 이차 상호작용 효과가 발견되었다($F(4, 236) = 24.81$, $p < .001$). 실험을 통해 도출된 각 독립변수 조합별 수치는 다음 [표 1]과 같다.

[표 1] 인터럽트 시점 및 피드백 형태별 통제감 및 신뢰도 기술통계 (N=60)

[Table 1] Descriptive statistics on control and reliability by interrupt timing and feedback type (N=60)

인터럽트 시점 (A)	피드백 형태 (B)	주관적 통제감 점수 (Mean ± SD)	시스템 신뢰도 점수 (Mean ± SD)
CI (실시간 상시 개입)	MS (최소 요약형)	3.84 ± 0.92	4.12 ± 0.85
	RL (실시간 로그형)	4.21 ± 1.04	4.01 ± 0.99
	ST (구조화 상태 추적)	4.85 ± 0.76	5.11 ± 0.64
MI (단계 전환점 개입)	MS (최소 요약형)	4.91 ± 0.68	5.20 ± 0.71
	RL (실시간 로그형)	5.12 ± 0.88	5.04 ± 0.82
	ST (구조화 상태 추적)	6.42 ± 0.45	6.58 ± 0.39
AI (불확실성 임계치 개입)	MS (최소 요약형)	3.45 ± 1.11	4.78 ± 0.88
	RL (실시간 로그형)	3.88 ± 0.95	4.45 ± 1.02
	ST (구조화 상태 추적)	5.23 ± 0.61	5.84 ± 0.53

통계적으로 가장 우수한 효과를 나타낸 조합은 [단계 전환점 개입형(MI) X 구조화된 상태 추적형(ST)] 구조로, 통제감(M=6.42)과 신뢰도(M=6.58) 모두에서 유의미하게 가장 높은 수치를 기록하였다($p < .001$). 상시 개입형(CI) 조건은 사용자가 개입 시점을 상시 판단해야 하는 심리적 방임 상태를 유발하여 통제감이 낮았으며, 불확실성 임계치 개입형(AI)에 최소 요약형(MS)이 결합될 때 정보의 맥락 결여로 인해 통제감이 가장 낮게(M=3.45) 측정되었다.

3.2.2 인지적 부하 분석

각 인터랙션 방식이 피험자에게 유발하는 인지적 부하의 차이를 분석하기 위해 [표 2]와 같이

분산분석 및 본페로니 사후검증을 수행하였다.

[표 2] 독립변수에 따른 NASA-TLX 인지적 부하 분석 및 사후 검증 결과
[Table 2] Analysis of NASA-TLX Cognitive Load and Post-hoc Verification Results Based on Independent Variables

변수 구분	세부 실험 조건	인지적 부하 평균 (SD)	F 값	p 값	사후검증 (Bonferroni)
인터럽트 시점 (A)	A_1 : 실시간 상시 개입 (CI)	5.62 ± 0.81	38.42	< .001	$A_1 > A_2$, A_3
	A_2 : 단계 전환점 개입 (MI)	3.12 ± 0.54			
	A_3 : 불확실성 임계치 (AI)	4.24 ± 0.73			$A_3 > A_2$
피드백 형태 (B)	B_1 : 최소 요약형 (MS)	4.68 ± 0.65	52.14	< .001	$B_2 > B_1 > B_3$
	B_2 : 실시간 로그형 (RL)	5.94 ± 0.88			
	B_3 : 구조화 상태 추적 (ST)	2.86 ± 0.41			

실시간 상시 개입형(A1, M=5.62)과 실시간 로그형 피드백(B2, M=5.94)은 높은 인지적 과부하를 전가하는 결함이 있음이 입증되었다. 무차별적인 텍스트 스트리밍 노출은 주도적 제어가 아닌 강제된 모니터링 노동으로 변질됨을 의미한다. 반면 구조화된 상태 추적형 피드백(B3)은 인지적 부하가 2.86으로 가장 낮아, 시각적으로 정제된 그래프가 제한된 작업 기억 공간을 효율적으로 보조하고 있음을 실증한다.

3.3 회귀분석을 통한 매개 효과 검증

주관적 통제감과 인지적 부하가 최종 결과 변수인 시스템 신뢰도 및 지속 사용 의도에 미치는 영향력을 규명하기 위해 다중회귀분석을 실시하였다. 결과는 [표 3]과 같다.

[표 3] 시스템 신뢰도에 대한 다중회귀분석 결과
[Table 3] Results of multiple regression analysis on system reliability

독립변수	비표준화 계수(B)	표준화 계수(β)	t 값	p 값	다중공선성 (VIF)
(상수)	1.241	-	3.14	.002	-
주관적 통제감	0.584	.521	7.84	<.001	1.24
인지적 부하	-0.312	-.342	-5.12	<.001	1.18
정보 투명성 지각	0.211	.185	2.94	.004	1.31

종속변수: 시스템 신뢰도 / 모델 적합도: R^2 = .648, 수정된 R^2 = .634, F = 44.12 (p<.001)

회귀모형의 설명력은 64.8%(R^2 = .648)로 매우 높게 나타났다. 시스템 신뢰도를 예측하는 가장 강력한 정(+)의 변수는 주관적 통제감(β = .521, p<.001)이었으며, 인지적 부하(β = -.342, p<.001)는 강력한 부(-)의 변수로 작동하였다. 즉, 자율형 AI 에이전트의 신뢰도를 높이기 위해서는 단순히 알

고리즘의 정밀도를 올리는 것뿐만 아니라, 사용자의 주관적 통제감을 높이고 인지 부하를 낮추는 인터랙션 설계가 필수적이다.

4. 자율형 AI 에이전트 UI/UX 가이드라인

4.1 단계적 정보 심층화기반 가시성 설계

인터페이스의 기본 레이어는 매크로 워크플로우를 직관적으로 파악할 수 있는 구조화된 상태 추적 노드 그래프로 단일화하여 노출해야 한다. 구체적으로 대시보드 상단 레이어에서는 현재 실행 중인 에이전트의 하위 목표, 소요 시간, 공정률만을 시각화하여 거시적 가시성을 확보하고, 인터랙티브 하위 레이어에서는 사용자가 특정 노드를 선택하는 주도적 탐색 행위를 했을 때에만 드롭다운 형태로 백엔드 로그 및 상세 추론 체인을 전개한다. 이러한 이중 레이어 구조는 정보의 투명성을 완벽히 보장하면서도, 가공되지 않은 실시간 로그형 피드백이 유발하던 심각한 인지 부하 문제를 원천적으로 제어하는 효과를 거둔다.

4.2 예측 가능한 인터럽트 앵커링 구축

사용자가 에이전트의 구동 과정을 상시 감시하지 않도록, 연산이 시작되기 전 전체 타임라인 상에 유저가 개입하여 의사결정을 수정할 수 있는 정량적 검토 포인트를 가시적으로 선언해야 한다. 에이전트 다단계 태스크의 단계별 가이드라인 매트릭스는 [표 4]와 같다.

[표 4] 에이전트 다단계 태스크 단계별 가이드라인 매트릭스

[Table 4] Agent Multi-stage Task Step-by-Step Guideline Matrix

작업 단계	주요 에이전트 행동	UI/UX 피드백 표현	인터럽트 권한 설계
계획 수립 (Planning)	목표 분할 및 하위 태스크 트리 생성	구조화된 노드 다이어그램으로 분할 경로 최초 시각화	사용자의 하위 태스크 삭제 및 순서 변경 권한 부여
도구 활용 (Executing)	외부 API, 웹 크롤러, DB 쿼리 실행	실행 중인 노드의 색상 변화(애니메이션) 및 활성 도구 아이콘 표시	상시 정지 버튼 제공 지양, 해당 노드 완료 시점으로 인터럽트 예약 대기 큐 적용
검증 및 정제 (Refining)	수집 데이터 크로스 체크 및 무결성 검사	신뢰도 점수 게이지 바 노출, 미달 항목 붉은색 강조	[단계 전환점 앵커] 최종 출력을 확정하기 전 인간의 최종 승인 UI 강제 활성화

사용자는 다음 인터럽트 시점이 명확히 설계되어 있음을 인지할 때 비로소 브라우저 화면을 계속 주시해야 하는 불안에서 벗어나며, 시스템에 대한 고차원적 심리적 주도권을 확보하게 된다.

4.3 동적 샌드박스 제어 제공

에이전트 인터럽트 발생 시 시스템 제어는 전면 리셋이 아닌 샌드박스형 분기 수정 방식을 취해야 한다. 이를 위해 오류 시 직전의 정상 노드 상태로 복원하는 상태 롤백과 일시 정지 상태에서 참조 변수나 URL을 직접 변경하는 매개변수 인라인 수정 기능을 제공해야 한다. 최종적으로는 수정 반영 시점부터 에이전트가 변경된 분기를 인지하여 이후 과정을 자율 연쇄 실행할 수 있도록 하는 분기 재개기반의 부분 재실행 메커니즘을 구축해야 한다.

5. 결론 및 향후 과제

본 연구는 인공지능이 스스로 워크플로우를 전개하는 자율형 AI 에이전트 환경에서 인간 사용자의 통제감 유지를 위한 UI/UX 상호작용 설계 방안을 실증적으로 규명하였다. 3 X 3 요인실험 및 분산분석을 통해, 사용자는 언제든 개입할 수 있는 상시적 권한보다 명확히 구조화되고 예측 가능한 단계 전환점 중심의 개입 권한에서 인지 부하를 대폭 줄이면서 최고의 통제감(M=6.42)을 얻는다는 점을 통계적으로 입증하였다. 또한 텍스트 중심의 실시간 로그 스트리밍은 인지적 피로만을 가중시킬 뿐이며, 단계적 정보 심층화가 결합된 구조화 상태 추적 인터페이스만이 시스템 신뢰도를 안전하게 인양할 수 있음을 확인하였다. 본 연구가 도출한 가이드라인은 향후 수많은 엔터프라이즈 자동화 솔루션 및 AI 에이전트 서비스가 범할 수 있는 기술 편의주의적 블랙박스 설계를 방지하고, 인간과 자율 에이전트 간의 조화로운 협업 파트너십을 구축하는 강력한 설계 표준이 될 것이다. 본 연구는 정형화된 오피스 업무 자동화 환경을 중심으로 실험을 진행하였다는 한계가 있으므로, 향후 연구에서는 오작동 시 위험 비용이 극도로 높은 의료 진단 AI 에이전트, 자율 금융 트레이딩 에이전트 등 고위험 도메인별 특수성을 반영하여 신뢰도 임계치 제어 인터페이스를 세분화 및 고도화해 나갈 필요가 있다.

References

[1] S. Amershi, D. Weld, M. Vorvoreanu, A. Fourney, B. Nushi, P. Collisson, J. Suh, S. Iqbal, P. N. Bennett, and K. Inkpen, "Guidelines for Human-AI Interaction," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, Glasgow, UK, May 4-9, 2019, Art. no. 3, pp. 1-13, doi: 10.1145/3290605.3300233.

[2] B. Shneiderman, "Human-Centered Artificial Intelligence: Reliable, Safe & Trustworthy," *International Journal of Human-Computer Interaction*, vol. 36, no. 6, pp. 495-504, Mar. 2020, doi: 10.1080/10447318.2020.1741118.

[3] S. Hwang and D. Kwak, "A Study on Functional Sportswear Design for Enhancing Sensory-Based UX: Focusing on the Analysis of Flow Experience Elements," *Journal of Next-generation Convergence Information Services Technology*, vol. 14, no. 6, pp. 785-796, Dec. 2025, doi: 10.29056/jncist.2025.12.05.

[4] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research," *Advances in Psychology*, vol. 52, pp. 139-183, Nov. 1988, doi: 10.1016/S0166-4115(08)62386-9.

[5] D. Johnston, D. Holtz, A. M. Richmond, C. Ong, P. Tambe, and A. Chatterji, "The Shift to Agentic AI: Evidence from Codex," Jun. 2026, arXiv:2606.26959, doi: 10.48550/arXiv.2606.26959.

[6] S. Yao et al., "React: Synergizing reasoning and acting in language models," Oct. 2022, arXiv:2210.03629, doi: 10.48550/arXiv.2210.03629.

[7] D. A. Norman, *The Design of Everyday Things*, Revised and expanded ed., Basic Books, 2013.

[8] Y. J. Lee, "A Study on Information Architecture & User Experience of the Smartphone," *Journal of Digital Convergence*, vol. 13, no. 11, pp. 383-390, Nov. 2015, doi: 10.14400/JDC.2015.13.11.383.

[9] M. A. Goodrich and A. C. Schultz, "Human-Robot Interaction: A Survey," *Foundations and Trends in Human-Computer Interaction*, vol. 1, no. 3, pp. 203-275, Jan. 2007, doi: 10.1561/1100000005.

[10] F. Flemisch, M. Heesen, T. Hesse, J. Kelsch, A. Schieben, and J. Beller, "Towards a Dynamic Balance between Humans and Automation: Authority, Ability, Responsibility and Shared Control," *Cognition, Technology & Work*, vol. 14, no. 1, pp. 3-18, Mar. 2012, doi: 10.1007/s10111-011-0191-6.

[11] M. Hoque, J. Shin, and N. Elmqvist, "Visualization for Human-Centered AI Tools," Jun. 2026, arXiv:2404.02147, doi: 10.48550/arXiv.2404.02147.

[12] J. Starkova, I. Zhou, and J. O. Stark, "CoAutoML: User Interface Framework for Machine Learning Novices using LLM-based AutoML and Test-Driven Machine Teaching," *Journal of Human-Computer Interaction Studies*, vol. 18, no. 2, pp. 142-156, Apr. 2026, doi: 10.1145/3613904.3642738.

[13] A. Abdul, J. Vermeulen, D. Wang, B. Y. Lim, and M. Kankanhalli, "Trends and trajectories for explainable AI in HCI," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, Montreal, Canada, Apr. 21-26, 2018, pp. 1-18, doi: 10.1145/3173574.3174134.