

특집논문 (Special Paper)

방송공학회논문지 제31권 제5호, 2026년 9월 (JBE Vol.31, No.5, September 2026)

<https://doi.org/10.5909/JBE.2026.31.5.948>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

인간과 Agentic AI 협업을 위한 여유시간(Margin) 기반 적응형 UX 프레임워크 연구

안 성 희^{a)†}

A Study on a Margin-Based Adaptive UX Framework for Human - Agentic AI Collaboration

Sunghee Ahn^{a)†}

요 약

오늘날 스마트 팩토리, 자율주행 모빌리티 등 고도 자동화 환경의 AI 기반 인터페이스는 여전히 정보 중심(Information-Oriented) 패러다임에 머물러 있다. 그러나 인간의 인지 처리 능력을 초과하는 속도와 규모로 의사결정을 수행하는 Agentic AI의 등장으로, 정보 중심 인터페이스는 소통의 한계를 드러내고 있으며 - 최적화된 개입에는 행동에 앞선 준비 시간이 필수적이나, 단순한 투명성 증가만으로는 업무 효율이 향상되지 않는다. 본 연구는 이를 투명성 역설(Transparency Paradox)로 정의하고, Agentic AI 시대 AI-UX의 목표가 투명성의 ‘극대화’가 아니라 ‘최적화(Optimization)’에 있음을 제안한다. 이를 위해 상황 인지(Situation Recognition)에서 판단 형성(Decision Formation), 행동 준비(Action Preparation), 제어권 전환(Control Transfer), 네 단계의 인지·행동 과정을 기반으로 AI의 예측 시점과 인간의 개입 사이의 여유시간, 즉 Margin(Time)을 변수로 정의하고 인터페이스가 동적으로 재구성되는 Margin 기반 Decision-Oriented UX Framework를 제안한다. 본 프레임워크는 기존 Explainable AI의 설명 중심 접근을 의사결정 중심(Decision-Oriented) AI-UX로 확장함으로써, Agentic AI 환경의 인간 중심 인터페이스 설계를 위한 새로운 설계 원칙과 이론적 토대를 제시한다.

Abstract

AI-based interfaces in highly automated environments, such as smart factories and autonomous mobility systems, remain in an information-oriented paradigm. However, with the emergence of Agentic AI, which performs decision-making beyond human cognitive capacity, information-oriented interfaces face significant limitations. Although effective intervention requires sufficient preparation time before action, increasing transparency alone does not necessarily improve operational efficiency. This study defines this phenomenon as the Transparency Paradox and argues that the goal of AI-UX in the Agentic AI era should shift from maximizing transparency to Transparency Optimization. To address this challenge, this study defines Margin(Time) as the temporal difference between the AI prediction horizon and the time required for human intervention, and proposes PHAIM (Predictive Human-Adaptive Interface Model), which dynamically reconfigures the interface according to this Margin across four stages: Situation Recognition, Decision Formation, Action Preparation, and Control Transfer. The proposed framework extends Explainable AI from an explanation-oriented paradigm to decision-oriented AI-UX, providing new design principles and a theoretical foundation for human-centered interface design in Agentic AI environments.

Keyword : Agentic AI, HAI Interaction, AI-UX, Transparency Paradox, Adaptive Interface

Copyright © 2026 Korean Institute of Broadcast and Media Engineers. All rights reserved.

“This is an Open-Access article distributed under the terms of the Creative Commons BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited and not altered.”

1. 서론

1. 연구 배경

기존 AI 시스템은 정보 제공이나 추천 수준에 머물렀으나, Agentic AI는 환경을 인식하고 미래를 예측하며 자율적으로 의사결정을 내려 실제 행동까지 수행하는 수준으로 발전하고 있다^[2]. 이에 따라 인간의 역할도 협업 주체로 변화^[3,10]하고 있으나 대부분의 인터페이스는 여전히 현재 시스템 상태를 전달하는 정보 중심 설계에 머물러, 인간의 실제 의사결정 과정을 충분히 반영하지 못한다^[4]. 이는 인간이 상황을 인지하고 행동으로 연결하기까지 일정한 인지적·물리적 시간을 필요로 하기 때문이다^[11]. 실제로 자율주행 분야의 제어권 전환 연구는 인수(take-over) 소요 시간이 긴급성과 요청 양식에 따라 수 초 단위로 달라짐을 보여준다^[12]. 반면 Agentic AI는 인간의 인지 처리 능력을 초과하는 속도와 규모로 의사결정을 수행하여 정보 과부하와 인지 부담을 증가시킬 수 있다^[21]. 신뢰도 정보와 설명을 함께 제공해도 협업 정확도가 항상 향상되지는 않았으며^[13], 오히려 투명한 AI 모델이 사용자의 오류 탐지 능력을 저하시키는 경우^[15]도 선행 연구에서 발표되었다. 이는 AI의 판단 속도와 인간의 의사결정 속도 사이에 시간적 비대칭(Time Asymmetry)이 커지고 있음을 의미한다. 따라서 인간이 제한된 시간 안에 적절한 판단을 내릴 수 있도록 정보를 선별해 적절한 시점에 제공하는 인터페이스가 필요하며, Agentic AI 시대의 AI-UX는 정보 전달 중심 설계를 넘어 인간의 의사결정을 지원하는 새로운 설계 원리를 요구한다.

2. 연구 목적

본 연구의 목적은 Agentic AI 시대 인간-AI 협업을 위한

새로운 인터페이스 패러다임을 제안하는 것이다. 기존 인터페이스가 현재 시스템 상태를 전달하는 데 초점을 맞추었다면, 본 연구는 인간의 인지 및 행동에 필요한 시간을 추정하고, AI의 예측 시점과 인간의 개입 가능 시간 사이의 여유(Margin)에 따라 동적으로 재구성되는 Human-Adaptive Interface 개념을 제안한다.

이를 위해 인터페이스는 단순히 AI의 내부 연산을 설명하는 데 그치지 않고, 인간이 위험을 인지하고 판단을 형성하며 행동을 준비하고 필요한 경우 제어권을 전환할 수 있도록 시간적 조건에 따라 정보를 선택적으로 재구성해야 한다. 본 연구는 이를 구체화하기 위해 PHAIM(Predictive Human-Adaptive Interface Model)을 제안하며, Margin(Time)을 핵심 제어 변수로 사용하여 인간의 의사결정 준비도(Decision Readiness)를 지원하는 인터페이스 설계 원리를 제시한다.

3. 연구 기여

본 연구의 주요 학술적 기여는 다음과 같다.

첫째, Agentic AI 환경에서 인간-AI 협업의 핵심 문제를 정보 부족이 아니라 ‘시간적 비대칭(Time Asymmetry)’의 문제로 새롭게 정의한다.

둘째, 인간의 상황 인지, 판단 형성, 행동 준비 시간을 통합적으로 고려하는 Human Temporal Characteristics 개념을 만들고, 이를 AI-UX 설계의 새로운 변수로 도입한다.

셋째, AI의 예측 시점과 인간의 개입 소요 시간 사이의 여유시간(Margin)에 따라 인터페이스가 동적으로 정보를 재구성하는 PHAIM(Predictive Human-Adaptive Interface Model)을 제안한다.

넷째, 기존의 상태 모니터링 중심 인터페이스를 Explanatory, Focused, Urgent, Escalation 네 상태로 확장하고, 이로부터 여섯 가지 설계 원칙(P1 - P6)으로 구성된 Human-Adaptive Interface Design Framework를 도출함으로써 Agentic AI 시대의 새로운 인터페이스 설계 방향을 제시한다.

본 연구는 AI-UX의 미래가 AI가 얼마나 많은 정보를 설명하는가에 의해 결정되는 것이 아니라, 인간이 Agentic AI와 협업하는 과정에서 효과적으로 의사결정하고 적절한 시

a) 홍익대학교 디자인컴퓨터전공학부(HongIk University)

* Corresponding Author : 안성희(Sunghee Ahn)

E-mail: ahnsunghee9@gmail.com

Tel: +82-44-860-2137

ORCID: <https://orcid.org/0000-0002-7717-2412>

· Manuscript received July 22, 2026; Revised September 9, 2026;

Accepted September 9, 2026.

점에 개입할 수 있도록 지원하는가에 의해 결정된다는 관점을 제시하며, 이를 위한 이론적 형식화를 시도한 초기 연구(early-stage formalization)로 자리매김한다.

II. Agentic AI 시대의 Human-AI 인터랙션의 변화

1. Evolution of Human - AI Interaction in the Agentic AI Era

HCI는 인간이 컴퓨터 시스템을 컨트롤하는 인터페이스 설계 중심으로 발전해 왔으나, 최근 Agentic AI의 발전은 이 패러다임에 근본적 변화를 요구한다. Human - AI Interaction 연구는 AI를 도구가 아닌 협업 대상으로 재정의하는 방향으로 전환되고 있다^[1].

Agentic AI는 스스로 목표를 설정하고 환경을 지속적으로 인식하며 자율적으로 의사결정과 행동을 수행한다는 점에서 기존 생성형 AI와 질적으로 다르다^[2]. 자동화 단계 모델에 따르면 인간의 역할은 이러한 자율성 위임 정도에 비례해 조작자(Operator)에서 감독자(Supervisor)로 이동하며^[3], 디지털 트윈 기반 인간-로봇 협업처럼 실시간 의사결정이 요구되는 환경에서는 이 역할 분담이 더욱 복잡해진다^[4]. 그 결과 Human - AI Interaction은 Information Interaction에서 Decision Interaction으로 진화하고 있으나, 기존 연구는 인터페이스가 인간의 의사결정 특성을 어떻게 반영해야 하는지에 대한 이론적 틀을 충분히 제시하지 못하고 있다. 이는 XAI 문헌에서도 반복적으로 확인된다. 데이터 과학자를 관찰한 연구는 시각화된 설명이 모델 이해보다 기존 가설의 확인이나 결과 정당화 용도로 쓰이는 경향을 보고했다^[5]. Social Transparency 연구는 기존 XAI가 알고리즘 중심 설명에 치우쳐 조직적·사회적 맥락을 반영하지 못한다고 지적하며^[6], AI 설명에 대한 요구는 전문가·비전문가 구분이 아니라 사용자의 지식 유형·목적·책임·과업에 따라 달라진다는 연구도 제시되었다^[7]. 최근 체계적 문헌 고찰들 역시 XAI 평가가 기술적 정확성보다 인간의 이해·유용성·신뢰 조절·협업 성과를 중심으로 이루어져야 한다고 강조한다^[8,9].

그러나 이러한 인간 중심 XAI 연구는 Agentic AI가 연속적으로 수많은 결정을 생성하는 환경에서 어떤 정보를 생략하고 무엇을 우선 제공해야 하는지, 인간의 정보처리 한계에 따라 설명을 선택적으로 제공하는 원리는 아직 충분히 논의하지 못하고 있다.

2. Human Temporal Characteristics and Decision Readiness in Agentic AI Collaboration

Agentic AI로의 전환은 인간-AI 협업의 시간적 구조를 근본적으로 변화시킨다. 협업의 병목은 연산 능력이 아니라, 시스템의 의사결정 주기와 인간의 인지 처리 간격 사이의 시간적 불균형에서 비롯된다.

AI의 자율성이 높아질수록 인간의 의미 있는 감독(meaningful oversight)은 어려워진다^[10]. 자동화 단계 모델에 따르면 정보 습득·분석·의사결정·행동 실행이 시스템에 위임될수록 인간의 역할은 조작자에서 감독자로 전환되며^[3], 이에 따라 인간은 자율화된 프로세스를 감시하다가 위험 발생 시 선택적으로 개입하는 역할을 수행하게 된다^[3,10]. 이러한 역할 전환이 어려운 이유는, AI가 밀리초 단위로 판단을 생성하는 반면 인간의 개입은 지각(Perception)·이해(Comprehension)·예측(Projection)의 순차적 인지 과정에 의존하기 때문이다^[11]. 시간적 간극은 제어권 전환(Control Transfer)에 관한 129개 연구를 종합한 메타분석에 따르면, 인수(take-over) 소요 시간은 상황의 긴급성, 부차 과제 유형, 요청 양식 등에 따라 체계적으로 달라진다^[12]. 이는 자율주행뿐 아니라 스마트 팩토리나 같은 Agentic AI 환경에서도 인간의 인지·개입 시간이 인터페이스 설계의 핵심 변수여야 한다는 본 연구 가설을 뒷받침한다. 또한 신뢰도 점수와 설명을 함께 제공해도 협업 정확도가 항상 향상되지는 않으며, 그 효과는 상황과 과업에 따라 달라진다^[13]. 투명한 AI 모델이 오히려 사용자의 오류 교정 능력을 저하시키는 경우도 연구되었으며^[14,15], 설명의 정확성과 무관하게 AI 권고 수용률만 높이는 경향도 지적된 바 있다^[16]. 설명과 정보량이 형식적 투명성은 높이지만 인지 부하가 임계치를 넘으면 실질적 이해와 개입 판단을 오히려 저해한다는 투명성의 모순(Transparency Paradox)이 여기서 성립한다. 즉 설명의 효과는 정보량이 아니라 사용자의 과업·시점·개입

표 1. Agentic AI 시스템에서 인간 제어권 전환을 위한 Margin(Time) 4단계

Table 1. Four Margin-Based Stages for Human Control Transition in Agentic AI Systems

Code	Margin (Time)	Stage Title	Human Role	AI Agent Role	Interface Design Guide
S1_W	High Margin	Situation Recognition Time	Supervisor / Watch	Autonomous Planner	Adjustable Autonomy
S2_R	Moderate Margin	Decision Formation Time	Shared Decision Maker / Ready	Collaborative Partner	Shared Control
S3_J	Low Margin	Action Preparation Time	Decision Validator / Join	Primary Decision Executor	Human-on-the-Loop
S4_C	Critical Margin	Control Transfer Time	Emergency Override / Control	Autonomous Safety Controller	Human Control

요구와의 적합성에 의해 결정된다는 것으로 평가 기준 역시 단순한 정보량에서 시간적·과업적 적합성으로 전환될 필요가 있다. 사용자의 조건과 상황을 통합적으로 고려하는 UX 설계는 Agentic AI 협업 환경에서 핵심 설계 요소가 된다.

본 연구는 이러한 인간의 시간적 특성을 Human Temporal Characteristics로 정의하며, 다음 네 단계로 구성한다. 표 1은 AI 자동화 환경에서 인간의 제어권 전환에 필요한 시간적 과정을 네 단계로 구분하고, 이를 Margin(Time)에 따른 인간 중심 정보 전달 단계로 제시한다.

AI 시스템에서 사람에게 제어권이 최종적으로 넘어가기까지의 네 단계는 위험을 인지하기 전에는 의사결정이 성립할 수 없고, 의사결정이 완료되어도 물리적 행동이 수행되지 않으면 제어권은 전환되지 않는다. 따라서 Agentic AI 시대의 인터페이스는 현재 상태를 전달하는 역할(Information Interface)을 넘어, 인간의 시간적 요구를 지원하는 시스템(Human-Adaptive Interface)으로 발전 설계되어야 한다.

3. Human-Adaptive Interfaces for Agentic AI Collaboration

최근에는 지능형 사용자 인터페이스(Intelligent User Interface) 연구가 맥락 인식, 개인화된 시각화, AI 기반 의사결정 지원으로 그 범위를 확장하고 있다. Explainable AI 연구 역시 사용자의 전문성·목적·인지적 요구에 따라 설명을 동적으로 제시하는 방법을 제안해 왔다^[6,7]. 그러나 이러한 접근들은 공통적으로 ‘어떤 정보를 보여줄 것인가(What

to display)’에 초점을 두고 있으며, ‘언제(When)’ 정보를 재구성해야 인간의 개입이 효과적인지는 충분히 다루지 않는다. 기술적으로 정확한 정보라도 인간의 개입 가능 시점을 지난 뒤에 제시된다면 운영상의 가치는 제한적이다. 디지털 트윈 기반 대시보드 연구 역시 센서 데이터와 예측 결과를 동시에 제공해야 하는 상황에서 인터페이스의 복잡성과 사용자 인지 부하가 함께 증가하는 문제를 지적한다^[4]. Shneiderman의 Human-Centered AI(HCAI) 프레임워크는 인간에 대한 의미 있는 통제(meaningful control)를 자동화 수준이 높아지는 상황에서도 유지해야 할 핵심 요건^[17]으로서 AI가 발전할수록 인간의 역할이 감독과 증강(augmentation)으로 이동하며^[18], 인간이 AI에 과업을 위임하는 과정에서 메타인지적 한계로 인해 역할 분배가 최적화되지 않을 수 있음을 실험적으로 증명^[19]한 선행 연구들도 있다. Amershi 등이 제시한 Human-AI Interaction 설계 가이드라인 역시 인터페이스가 AI의 불확실성과 실행 시점을 사용자에게 명확히 전달해야 한다는 원칙을 포함한다^[20]. 이들 연구는 공통적으로 ‘무엇을 설명할 것인가’가 아닌 인간이 실제로 개입할 수 있는 시간을 확보하는 것이 통제권 유지의 전제 조건임을 시사한다.

III. Conceptual Framework of the Decision-Oriented UX Model

1. Design Rationale

앞서 II.3절에서 제안한 Human-Adaptive Interface를 구

체화하기 위해, 본 연구는 AI의 예측 시점과 인간의 개입 소요 시간을 비교하고 그 여유시간(Margin)에 따라 인터페이스를 동적으로 재구성하는 PHAIM(Predictive Human-Adaptive Interface Model)을 제안한다. AI Prediction Module은 Agentic AI 시스템이 내부적으로 보유한 계획 지평(planning horizon)으로부터 위험 또는 개입 필요 시점까지 남은 시간(T_{AI})을 산출한다. Human Temporal Estimation Module은 II장에서 정의한 인간의 결정까지 가는 단계들의 합으로 T_{human} 을 산출한 결과이다.

이 값들은 고정값이 아니라 상황에 따라 변할 수 있다. Cao 등은 의사결정의 단계(정보 탐색, 판단, 확정)에 따라 시간 압박이 미치는 영향이 다르다는 것을 실증했고^[23], Swaroop 등은 AI-보조 의사결정에서 정확도와 시간 사이에 구조적인 트레이드오프를 실험을 통해 정량화했다^[24].

이 두 연구는 ' T_{human} '이 변하는 추정 가능한 변수라는 PHAIM의 전제를 뒷받침하는 최신 연구의 근거다. 두 모듈의 출력은 Margin Calculator로 수렴하여 ' $Margin = T_{AI} - T_{human}$ '을 계산한다. 이 값이 PHAIM 전체를 관통하는 단일 제어 변수이다. Margin은 Interface Adaptation Engine으로 전달되어 표출될 인터페이스의 정보 우선순위와 시각적 강조를 재구성하고, 그 결과가 Human-Adaptive Interface를 통해 인간 운용자에게 제시된다. 그림 1과 같이 운용자의 개입은 제어권 전환(Control Transfer)을 통해 다시 Agentic AI 시스템으로 피드백되며, 이 순환 구조가 본 논문의 개념 모델인 PHAIM의 전체 동작 루프를 이룬다.

PHAIM 아키텍처 다이어그램의 전체 흐름은 위에서 아래로 순환하는 루프 구조이다. 맨 위 Agentic AI system은 출발점으로 이 시스템이 지속적으로 미래를 예측하고 행동

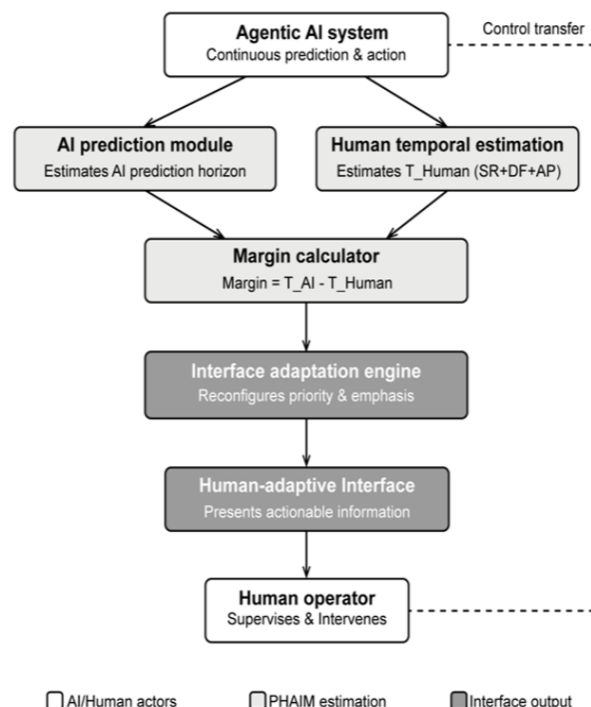


그림 1. PHAIM(Predictive Human-Adaptive Interface Model)의 개념적 아키텍처

Fig. 1. Conceptual Architecture of the Predictive Human-Adaptive Interface Model (PHAIM)

(하단 범례: 흰색 = AI/인간이라는 실제 행위 주체

회색 = PHAIM 내부의 추정·계산 모듈

진회색 = 그 계산 결과를 실제 인터페이스로 바꾸는 출력 모듈

본 UX 모델 프레임워크의 논리 구조(행위자 → 계산 → 출력 → 행위자)임)

을 실행하는 구조를 보여준다. 두 번째 줄 왼쪽은 AI가 위협이나 개입이 필요한 시점까지 남은 시간, 즉 T_{AI} (예측 지평)를 계산하고 오른쪽은 인간이 상황을 인지·판단·행동하는 데 필요한 시간, 즉 T_{human} (Situation Recognition + Decision Formation + Action Preparation의 합)을 계산해서 작동한다. 이 두 모듈은 독립적으로 값을 산출한다. 세 번째 줄은 남은 Margin을 추정하고 위 두 값이 수렴하는 지점이다. 수식은 ' $Margin = T_{AI} - T_{human}$ '을 계산하며 이 수치가 PHAIM 전체를 움직이는 핵심 변수다. 네 번째 줄은 Margin 값을 받아서 실제로 무엇을 정보로 보여줄지를 결정하는 단계다. 여기서 '실행/출력' 단계로 넘어간다. 다섯 번째 줄은 실제로 화면에 나타나는 결과물, 즉 동적으로 반응하며 재구성된 인터페이스다. 맨 하단은 Human operator로 인간 운용자가 이 인터페이스를 보고 개입으로 전환 후 결정에 영향을 주는 정보의 원리를 설명한다. 모델에서 오른쪽 점선 화살표(Control transfer)는 Human operator에서 시작해 맨 위 Agentic AI system 방향의 순환 루프이다. 인간 개입 결과(제어권) 후 다시 AI 시스템으로 회귀하고, 다음 사이클의 T_{AI} 계산에 반영한다. 점선은 인간의 개입이 일어났을 때만 발생하는 일시적 생성 흐름이다.

2. Formalization of Human Temporal Characteristics and Margin

본 논문은 인간의 개입 소요 시간(T_{human})과 이를 기반으로 한 여유시간(Margin)을 개념화하였다. 이는 이론적으로 정의된 원자적 처리 단위의 시간을 결합하여 복합 과업의 전체 소요 시간을 추정하는 수식화 접근이다. 다만 본 연구의 개념적 수식은 아직 검증 단계를 거치지 않은 가설 생성 단계(hypothesis-generating stage)의 모델이다.

2.1 Definition of T_{human}

Human Temporal Characteristics 중 개입 이전에 소요되는 세 단계(Situation Recognition, Decision Formation, Action Preparation)의 합을 인간이 개입을 시작하기까지 필요한 총 시간(T_{human})으로 정의한다.

$$T_{human} = T_{SR} + T_{DF} + T_{AP} \quad (1)$$

2.2 Theoretical Lower Bound

Card(1981)의 Model Human Processor는 인간의 정보처리를 지각 사이클($\tau_P \approx 100ms$), 인지 사이클($\tau_C \approx 70ms$), 운동 사이클($\tau_M \approx 70ms$)로 분해한다^[22]. SR, DF, AP는 각각 Card(1981)의 Model Human Processor가 정의하는 지각, 인지, 운동의 처리 사이클에 대응하도록 개념적으로 설정하였다. 다음 수식은 그 물리적 하한값으로서 가장 단순한 자극-반응 과업에서도 인간이 이보다 빠르게 개입하기 어렵다는 이론적 한계점으로 사용한다. 실제 인터페이스 조작에는 지각·인지·운동 처리뿐 아니라 입력 행동 시간이 포함될 수 있다^[28]. (근거 수준: τ_P, τ_C, τ_M 은 Card(1981)의 실측값을 인용한 실증적(empirical) 근거 수집.)

$$T_{SR} \geq \tau_P (\approx 100ms), T_{DF} \geq \tau_C (\approx 70ms), \quad (2)$$

$$T_{AP} \geq \tau_M (\approx 70ms)$$

$$T_{human} \geq \tau_P + \tau_C + \tau_M \approx 240ms \quad (3)$$

2.3 Compositional Parameterization

실제 소요 시간은 이론적 하한에 과업·도메인·개인의 상태에 따른 변동 배율이 적용되는 것으로 가정하였다. α 는 상황별 배율이다. 각 단계 고유의 과업 복잡도 배율($\alpha_{SR}, \alpha_{DF}, \alpha_{AP}$)에, 세 단계에 공통으로 작용하는 도메인 배율(α_{domain})과 개인 상태 배율($\alpha_{individual}(t)$)이 곱해진다. 단순 결정에서부터 인지적 처리 상황으로 갈수록 필요 시간은 배가 된다.

$$T_{SR} = \tau_P \times \alpha_{SR}, T_{DF} = \tau_C \times \alpha_{DF}, \quad (4)$$

$$T_{AP} = \tau_M \times \alpha_{AP}$$

(근거 수준: α_{SR} (지각 단계 배율)은 단순 경고 신호 대비 복수의 센서 정보를 종합해야 하는 상황일수록 증가하며 Endsley의 Situation Awareness 3단계 이론(지각→이해→예측)의 처리 깊이(processing depth)에 비례한다^[11]). 본 연구에서 α 는 실증적으로 확정된 상수가 아니라 기본 처리시간에 대한 상대적 증가를 표현하기 위한 개념적 조정계수(adjustment coefficient)로 정의한다. 예를 들어 복수 정보

원의 통합이나 높은 판단 복잡성이 요구되는 상황에서는 해당 단계의 α 가 증가하는 것으로 해석한다. α 의 구체적인 수치를 실증적으로 산정하지 않으며, 실제 계수의 추정과 UX 검증이 추가로 필요하다.

2.4 Definition of Margin

AI Prediction Module이 산출하는 T_{AI} (위험 또는 개입 필요 시점까지 남은 시간)와 T_{human} 의 차이를 Margin으로 정의한다.

$$Margin(t) = T_{AI}(t) - T_{human}(t) \quad (5)$$

Margin은 세 가지 상태로 해석한다. $Margin > 0$ 은 인간이 개입을 완료할 수 있는 시간적 여유가 남아 있음을 의미하고, $Margin \approx 0$ 은 개입 가능 시간이 임계치에 도달했음을 의미한다. $Margin < 0$ 은 인간이 필요한 개입을 완료하기 전에 AI의 실행이 완료될 가능성이 높은 상태를 의미한다. 본 연구는 이러한 Margin의 변화에 따라 인터페이스를 재구성한다.

2.5 T_{human} Parameterization

식 (1)에서 (5)까지를 종합하면 T_{human} 은 하나의 파라미터화된 함수로 표현된다.

$$T_{human}(t) = \alpha_{domain} \alpha_{individual}(t) (\tau_P \alpha_{SR} + \tau_C \alpha_{DF} + \tau_M \alpha_{AP}) \quad (6)$$

3. Interface Adaptation Logic (Decision-Oriented Reconfiguration)

Margin 값의 구간에 따라 인터페이스는 서로 다른 상태로 전환된다. Margin이 클수록 인간은 충분한 정보를 바탕으로 검토할 시간이 있으므로 설명 중심의 뷰가 유지되지만, Margin이 줄어들수록 인터페이스는 정보를 축소하고 개입 가능한 조작 요소를 전면 배치하는 방향으로 재구성된다.

그림 2의 네 단계의 상태는 III.2의 Margin 정의를 세분화한 것이며, 두 개의 임계값 θ_1 , θ_2 를 도입하여 다음과 같이

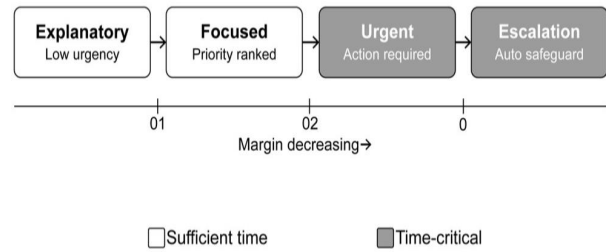


그림 2. Margin 기반 4단계 Interface Adaptation Logic
Fig. 2. Four Margin-based Interface Adaptation Logic

구분한다. 이때 θ_1 과 θ_2 는 고정된 보편 상수가 아니라 T_{human} 및 도메인별 안전 요구에 따라 조정되는 개념적 임계값으로 정의한다.

$$State(t) = \begin{cases} \text{Explanatory} & Margin(t) > \theta_1(t) \\ \text{Focused} & \theta_2(t) < Margin(t) \leq \theta_1(t) \\ \text{Urgent} & 0 < Margin(t) \leq \theta_2(t) \\ \text{Escalation} & Margin(t) \leq 0 \end{cases} \quad (7)$$

첫 번째 단계인 Explanatory ($Margin > \theta_1$)는 시간적 여유가 충분한 상태이다. AI의 판단 근거와 대안적 설명을 함께 제공하여 인간이 상황을 깊이 있게 이해하도록 지원한다. 이 구간에서는 정보의 양보다 인간의 지각의 수준과 이해의 질을 우선한다.

두 번째 단계인 Focused ($\theta_2 < Margin \leq \theta_1$)는 Margin(여유시간)이 줄어들기 시작하는 상태이다. 인터페이스는 정보의 우선순위를 재정렬하여 개입과 직접 관련된 항목을 상단에 배치하고 이때부터는 부차적인 설명은 축소하거나 접어서(collapse) 인지 부하를 낮춘다^[20].

세 번째 단계인 Urgent ($0 < Margin \leq \theta_2$)는 개입 가능 시간이 임계치에 도달한 상태이다. 인터페이스 가이드는 인터페이스의 시각 요소와 데이터 값 그리고 현황 설명을 최소화하고 관리자(인간)가 조작가능한 개입 요소(정지, 승인, 재조정 등)를 시각적으로 UI를 강조한다. 이 구간에서 추가 정보 제공은 오히려 판단을 지연시킬 수 있으므로^[13-16], 정보량이 아니라 ‘개입 동선 단축’이 설계 목표가 된다.

네 번째 Escalation ($Margin \leq 0$)은 인간의 개입이 AI의 실행 완료 전에 이루어질 가능성이 낮아진 상태이다. 이 구간에서는 인터페이스가 인간의 판단을 기다리지 않고 사전

에 rule-based로 정의된 안전 정책(예: 자동 정지, 보수적 기본 행동)을 실행하며, 사후 검토를 위한 로그와 알림만을 제공한다. 이때 Θ_1 , Θ_2 는 고정 상수가 아니라 III.2절에서 정의한 T_{human} 의 함수로 설정된다. 즉 임계값 자체가 인간의 시간적 특성에 비례하여 설정되므로, 인간이 느려지면(예: 피로, 인지 부하 상승) Θ_1 과 Θ_2 도 함께 커져 더 이른 시점에 Focused·Urgent 상태로 전환이 필요하다.

4. Scenario / Walkthrough

디지털 트윈 기반 스마트 제조 라인에서, 로봇 팔이 예정에 없던 경로로 이동하며 근처 작업자와의 충돌 위험이 예측된 상황을 가정한다. 이 시나리오는 2장에서 언급한 디지털 트윈 기반 인간-로봇(Agentic AI System) 협업 환경에 해당한다^[4].

초기 조건은 Agentic AI 시스템이 $t=0$ 에서 위험을 최초 감지하고, 이 시점의 AI 예측 지평 $T_{AI}(0)=16$ 초인 것으로 설정한다. 해당 인간 작업자의 T_{human} 은 $T_{SR} + T_{DF} + T_{AP} = 1.2 + 1.8 + 1.0 = 4.0$ 초로 시뮬레이션을 위해 가정한다. 이 수치는 제조 현장에서 실증된 값이 아니라 모델의 작동 논리를 설명하기 위한 예시값이다.

그림 4는 AI가 위험을 감지한 시점부터 예상 충돌 시점까지 남은 Margin(Time)의 변화에 따라 인간의 인지 및 개입 과정이 시간 순서대로 표현되어 있다. PHAIM은 남은 Margin에 따라 인터페이스를 단계적으로 전환하여 적절한 시점의 인간 개입과 안전한 제어권 전환을 지원한다.

- $t=0s$: $Margin = 16 - 4 = 12$ 초. $12 > \Theta_1(10)$ 이므로 Explanatory 상태로 시작한다. AI가 위험 판단 근거

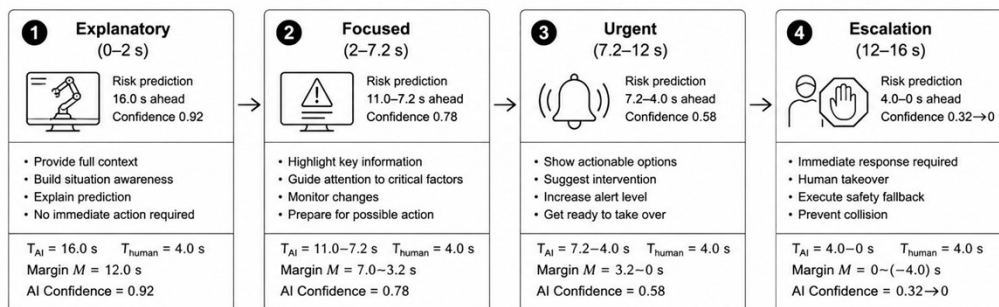


그림 3. PHAIM 설계가 적용된 스마트 제조 시나리오

Fig. 3. Smart Manufacturing Scenario Applying the PHAIM Design Framework

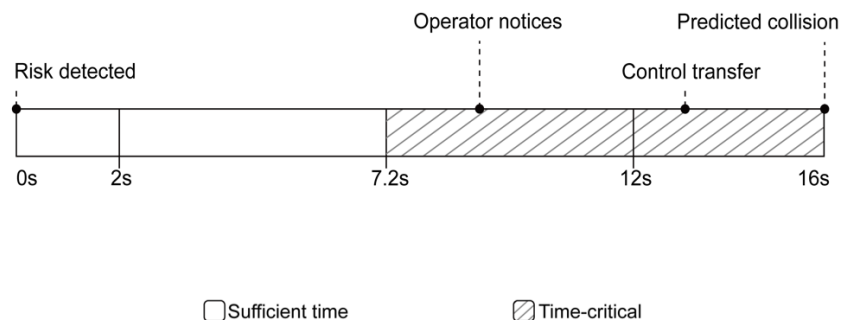


그림 4. PHAIM 기반 스마트 제조 시나리오에서의 인간 제어권 전환 타임라인

Fig. 4. Timeline of Human Control Transition in a PHAIM-Based Smart Manufacturing Scenario

(쾌적 이탈 원인, 대안 경로)를 함께 제시한다.

- $t=2s$: 남은 T_{AI} 가 감소하며 $Margin$ 이 Θ_1 아래로 내려가 **Focused** 상태로 전환된다. 인터페이스 설명이 축소되고 기계 혹은 기기의 예상 경로와 작업자 위치를 강조한다.
- $t=7.2s$: $Margin$ 이 $\Theta_2(4.8)$ 아래로 내려가 **Urgent** 상태로 전환된다. 인터페이스는 정지·경로 재조정 버튼을 화면 중앙에 배치하고 부차 정보를 모두 제거한다.
- $t=9s$: 작업자가 **Urgent** 상태의 시각적 강조 인지 후 개입을 시작한다(Situation Recognition 개시). 이는 이상적인 $t=0$ 대응이 아니라 실제로는 지연된 반응이지만, 남은 T_{AI} 가 아직 7초 있어 개입이 물리적으로 가능한 구간이다.
- $t=10.2s$: $T_{SR}(1.2s)$ 완료. 이어서 $T_{DF}(1.8s)$ 가 진행되어 $t=12.0s$ 에 Decision Formation이 끝난다. 그러나 이 시점은 $Margin$ 이 정확히 0이 되는 순간이기도 하다 – 즉 인터페이스는 작업자가 아직 행동을 준비하는 도중에 자동으로 Escalation 상태로 전환되며, 사전 정의된 안전 정책(예: 로봇 팔 자동 감속)이 함께 발동된다.

- $t=13.0s$: $T_{AP}(1.0s)$ 까지 완료되어 작업자의 Control Transfer가 실제로 이루어진다. 이 시점에 남은 T_{AI} 는 여전히 3초($16-13$)이므로, 자동 안전 정책과 작업자의 수동 개입이 중첩된 상태로 충돌을 회피한다.
- $t=16s$: 개입이 없었을 경우의 예측 충돌 시점이다(또는 오류 시점). 실제로는 $t=13s$ 에 제어권이 이미 이전 되었으므로 이 시점은 인간 작업자에게는 도달하지 않는다.

IV. Human-Adaptive Interface Design Framework

1. Principles of the Human-Adaptive Interface Design Framework

앞서 III장에서 제시한 PHAIM의 수식은 개념적 형식화이며, 실무 설계자가 모든 상황에서 $Margin$ 을 직접 계산하기는 어렵다. 다음의 표 2는 본 논문의 개념적 논리로부터 도출되는 6개 설계 원칙을 제시한다. 본 연구가 제안하는 여섯 개 원칙은 Amershi 등의 Human-AI Interaction 가이

표 2. 적응형 인터페이스 설계 프레임워크의 6대 원칙

Table 2. 6 Principles of Human-Adaptive Interface Design Framework

No	Principle	Design Implications	Design Adoption Framework (PHAIM Component)
P1	Temporal Primacy	Time as the primary design variable (Human-Centered Time Simulation)	-Define the remaining time before human intervention prior to determining information content or presentation. -Design information only after the human time simulation has been established. -Introduce available human supervision time as a design parameter for determining automation levels ^[3] .
P2	Information Reduction	Maintaining meaningful human control	-Quantify the level of information compression based on the available Margin (Time) ^[20] . -As the available Margin decreases, reduce the overall information volume, rather than merely reorganizing its presentation.
P3	Graduated Escalation	Progressive reallocation of human supervisory roles	-Avoid abrupt transitions directly to warning or emergency states. -Calculate the Margin(remaining time) and progressively transition by 4 interface states: Explanatory → Focused → Urgent → Escalation.
P4	Dynamic Individual Calibration	Real-time personalization based on human processing speed	-Adapt the interface according to individual differences in cognitive processing speed. -Continuously update user models during operation using physiological and behavioral signals (e.g., fatigue and attention level).
P5	Safe Degradation	Ensuring standard safe default behavior	-Prepare predefined AI automation strategies for situations where human control transfer fails (e.g., no response or emergency conditions). -When the available Margin reaches zero, execute predefined fail-safe actions (e.g., automatic stop).
P6	Mode Transition	Making interface state transitions perceptible	-Interface state transitions should be explicitly communicated to the human operator. -Silent or unannounced automation transitions may lead to failures in mode awareness ^[27] . -Provide dedicated notification mechanisms for Human-Adaptive Interface transitions.

드라인^[20]이나 Shneiderman의 HCAI 원칙^[17]에 반하는 것이 아니라 오히려 그 원칙들이 다루지 않았던 ‘언제’라는 시간의 축을 덧붙여 제안하는 보완적 방향성을 가짐과 동시에 실무 설계지침이 된다. 예를 들어 Amershi 등의 ‘AI의 불확실성을 명확히 전달하라’는 원칙은 P1 - P2와 결합될 때 ‘불확실성을 언제, 얼마나 압축해서 전달할 것인가’라는 설계 프레임워크를 제안하는 것이다.

V. Discussion and Limitation

1. Discussion

본 논문의 PHAIM은 최종적으로 완성된 시스템이나 검증된 알고리즘이 아니라, Margin을 중심으로 Human-Adaptive Interface의 상태 전환을 형식화한 개념 모델(conceptual model)이다. 논문의 학술적 기여는 인터페이스 설계에서 인간의 개입 소요 시간을 명시적 변수로 다루어야 한다는 원칙을 제안하는 데 있다. 실무 설계자는 최소한 III.3절의 상태 전이 로직(Explanatory → Focused → Urgent → Escalation)을 정성적 체크리스트로 적용하여, 위험 근접도에 따라 화면 요소를 단계적으로 축소하는 설계를 고려할 수 있다. 여기서 중요한 개념 도출은 이 결정에 시간 차원을 추가한다는 것이다. 동일한 기능이라도 도메인의 T_AI가 인간의 T_human보다 항상 작다면(구조적으로 Margin이 음수), 그 기능은 애초에 인간 감독이 아니라 완전 자동화 대상으로 재분류되어야 한다는 설계 판단 기준을 제공할 수 있는 중요한 개념 모델이 된다.

Shneiderman의 Human-Centered AI 프레임워크는 인간에 대한 의미 있는 통제(meaningful control)를 자동화 수준이 높아지는 상황에서도 반드시 유지해야 할 핵심 요건으로 제시한다^[17]. 그렇지만 이 프레임워크는 통제가 유지되는 조건을 다루지 않는다. 본 논문의 모델들은 이 공백에 대해 ‘의미 있는 통제는 정보 접근성만으로 성립하지 않으며, 개입을 완료할 수 있는 물리적 시간이 확보되어야 성립한다’는 조작적 정의를 추가한다. 즉 $\text{Margin} > 0$ 은 Shneiderman이 말하는 meaningful control의 시간적 필요조건으로 재해석될 수 있다.

Amershi 등의 Human-AI Interaction 가이드라인은 AI의 불확실성과 실행 시점을 사용자에게 명확히 전달해야 한다는 원칙을 제시한다^[20]. 본 논문의 Interface Adaptation Engine(3장)은 이 원칙을 ‘언제, 어떤 형태로’ 전달할지에 대한 구체적 시스템의 상태 값(state machine)을 개념적으로 추출해 본 연구 사례로 볼 수 있다. 다만 본 가이드라인 자체가 시간 임계값을 명시하지 않으므로, PHAIM의 Θ_1 , Θ_2 는 이 가이드라인을 정량화하려는 시도이다. 이에 덧붙여 Raisch와 Krakowski가 제시한 자동화-증강 역설(automation-augmentation paradox)은 AI가 발전할수록 인간의 역할이 감독과 증강으로 이동한다고 설명한다^[18]. 본 논문은 여기에 ‘시간적 조건’을 추가하여, 인간이 감독자 역할을 실제로 수행할 수 있는지 역할을 부여할 뿐 아니라 Margin의 부호와 크기에 의해서도 제약될 수 있음을 제안한다.

기존의 설명 중심 인터페이스가 주로 AI의 판단 근거와 불확실성을 사용자에게 어떻게 전달할 것인가에 초점을 둔다면^[13,20], PHAIM은 정보의 내용뿐 아니라 인간이 실제로 개입할 수 있는 남은 시간을 인터페이스 재구성 of 핵심 조건으로 사용한다. 따라서 PHAIM의 차별점은 설명 정보를 단순히 추가하는 데 있지 않고, Margin의 변화에 따라 Explanatory-Focused-Urgent-Escalation으로 정보량과 상호작용 방식을 단계적으로 변화시키는 시간 기반 적응 논리에 있다. 다만 이러한 상대적 효과는 본 연구에서 실증적으로 비교하지 않았으므로, 현 단계에서는 이론적 차별성으로 한정하여 해석할 필요가 있다.

2. Limitations

본 연구는 개념 모델(PHAIM)을 형식화하고 그 설계 논리를 시나리오를 통해 예시적으로 제시하였으나, 실증적 검증은 수행하지 않았다. 따라서 T_human의 파라미터와 Θ_1 , Θ_2 의 임계값은 이론적 근거와 선행 연구를 바탕으로 한 개념적 값이며, 제조·디지털 트윈 도메인에서 실증된 값으로 해석해서는 안 된다. 향후 연구에서는 실제 과업별 반응시간, 인지 부하 및 인터페이스 조작 시간을 측정하여 파라미터와 임계값을 검증할 필요가 있다. 또한 시나리오에서 단일 운용자-단일 위험 시나리오만 다루었으며, 복수의 위험

이 동시에 발생하거나 복수 운용자가 개입을 분담하는 상황은 모델에 반영되지 않았다.

Ⅵ. 결 론

본 연구는 Agentic AI의 등장으로 인간-AI 협업 구조가 ‘AI가 일상적으로 판단하고 인간이 감독·개입하는 관계’로 전환되는 변화에 대응하여 Human-Adaptive Interface의 설계 원리를 제안하였다. 기존의 반응형 인터페이스가 시스템 상태나 사용자 특성에 따라 반응했다면, Human-Adaptive Interface는 운용자의 시간적 준비 상태(temporal readiness)를 지속적으로 추정하고, 개입까지 남은 시간에 따라 정보 우선순위·시각적 강조·상호작용 복잡도·운영 권고를 동적으로 재구성했다. 또한 인간의 개입 준비 시간(T_{human})과 AI의 예측 시점(T_{AI}) 사이의 여유시간(Margin)을 인터페이스 적응의 핵심 변수로 정의하고, 이를 기반으로 PHAIM(Predictive Human-Adaptive Interface Model)과 6개의 설계 원칙(P1 - P6)으로 구성된 Human-Adaptive Interface Design Framework를 연구하여 제안하였다.

이는 개념적 프레임워크를 제안하는 연구로서 실증적 검증 단계까지는 포함하지 않았으나 제안된 수식과 임계값(Θ_1 , Θ_2 , k_1 , k_2)은 Card(1981)의 인간 정보처리 모델을 이론적 기반으로 하여 근거를 기반으로 인터페이스 적응 메커니즘을 개념화한 것이다. 다만 Θ_1 과 Θ_2 는 아직 실증적으로 검증된 고정 상수가 아니라 후속 연구에서 도메인별로 검증되어야 할 개념적 임계값이다.

본 연구의 기여는 완성된 시스템을 제시하는 데 있기보다, 제어실의 인지 부하 모니터링, 자율주행의 안전 정지 등 개별적으로 연구되어 온 시간 기반 개입 메커니즘을 ‘Margin’이라는 공통 개념으로 통합하고, 이를 인터페이스 상태 전환의 일반화된 설계 프레임워크로 구조화했다는 점에 있다. 향후 연구에서 검증 가능한 설계 변수와 가설을 명확한 모델의 형태로 제안했다는 점에서 학술적 의의를 갖는다.

Agentic AI 시대의 인터페이스 설계는 ‘무엇을 보여줄 것인가’보다 ‘언제, 그리고 얼마나 남은 시간 안에 보여줄

것인가’를 중심으로 재구성될 필요가 있다. 본 연구는 이러한 관점의 전환을 제안함으로써, Agentic AI 환경에서 인간 중심 인터페이스 설계를 위한 새로운 이론적 출발점을 제시한다.

참 고 문 헌 (References)

- [1] M. Chignell, L. Wang, A. Zare, and J. Li, “The Evolution of HCI and Human Factors: Integrating Human and Artificial Intelligence,” *ACM Transactions on Computer-Human Interaction*, Vol.30, No.2, pp.1-30, 2023. (accessed Jul. 2, 2026)
doi: <https://doi.org/10.1145/3557891>
- [2] D. B. Acharya, K. Kuppan, and B. Divya, “Agentic AI: Autonomous Intelligence for Complex Goals—A Comprehensive Survey,” *IEEE Access*, Vol.13, pp.18912-18936, Jan. 2025. (accessed Jul. 2, 2026)
doi: <https://doi.org/10.1109/ACCESS.2025.3532853>
- [3] R. Parasuraman, T. B. Sheridan, and C. D. Wickens, “A Model for Types and Levels of Human Interaction with Automation,” *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans*, Vol.30, No.3, pp.286-297, 2000. (accessed Jul. 21, 2026)
doi: <https://doi.org/10.1109/3468.844354>
- [4] A. Baratta, A. Cimino, F. Longo, and L. Nicoletti, “Digital Twin for Human-Robot Collaboration Enhancement in Manufacturing Systems: Literature Review and Direction for Future Developments,” *Computers & Industrial Engineering*, Vol.187, p.109764, 2024. (accessed Jul. 10, 2026)
doi: <https://doi.org/10.1016/j.cie.2023.109764>
- [5] H. Kaur, H. Nori, S. Jenkins, R. Caruana, H. Wallach, and J. Wortman Vaughan, “Interpreting Interpretability: Understanding Data Scientists’ Use of Interpretability Tools for Machine Learning,” *Proceeding of ACM CHI Conference on Human Factors in Computing Systems*, pp.1-14, 2020. (accessed Jul. 10, 2026)
doi: <https://doi.org/10.1145/3313831.3376219>
- [6] U. Ehsan, Q. V. Liao, M. Muller, M. O. Riedl, and J. D. Weisz, “Expanding Explainability: Towards Social Transparency in AI Systems,” *Proceeding of ACM CHI Conference on Human Factors in Computing Systems*, Article 82, 2021. (accessed Jul. 10, 2026)
doi: <https://doi.org/10.1145/3411764.3445188>
- [7] H. Suresh, S. R. Gomez, K. K. Nam, and A. Satyanarayan, “Beyond Expertise and Roles: A Framework to Characterize the Stakeholders of Interpretable Machine Learning and Their Needs,” *Proceeding of ACM CHI Conference on Human Factors in Computing Systems*, 2021. (accessed Jul. 2, 2026)
doi: <https://doi.org/10.1145/3411764.3445088>
- [8] J. Kim, H. Maathuis, and D. Sent, “Human-Centered Evaluation of Explainable AI Applications: A Systematic Review,” *Frontiers in Artificial Intelligence*, Vol.7, p.1456486, 2024. (accessed June. 21, 2026)
doi: <https://doi.org/10.3389/frai.2024.1456486>
- [9] Y. Rong, T. Leemann, T. T. Nguyen, L. Fiedler, P. Qian, V. Unhelkar,

- T. Seidel, G. Kasneci, and E. Kasneci, "Towards Human-Centered Explainable AI: A Survey of User Studies for Model Explanations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.46, No.4, pp.2104-2122, 2024. (accessed June. 21, 2026)
doi: <https://doi.org/10.1109/TPAMI.2023.3331846>
- [10] T. O'Neill, N. McNeese, A. Barron, and B. Schelble, "Human - Autonomy Teaming: A Review and Analysis of the Empirical Literature," *Human Factors*, Vol.64, No.5, pp.904-938, 2022. (accessed June. 21, 2026)
doi: <https://doi.org/10.1177/0018720820960865>
- [11] M. R. Endsley, "Toward a Theory of Situation Awareness in Dynamic Systems," *Human Factors*, Vol.37, No.1, pp.32-64, 1995. (accessed Jul. 18, 2026)
doi: <https://doi.org/10.1518/001872095779049543>
- [12] B. Zhang, J. de Winter, S. Varotto, R. Happee, and M. Martens, "Determinants of Take-over Time from Automated Driving: A Meta-Analysis of 129 Studies," *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol.64, pp.285-307, 2019. (accessed Jul. 2, 2026)
doi: <https://doi.org/10.1016/j.trf.2019.04.020>
- [13] Y. Zhang, Q. V. Liao, and R. K. E. Bellamy, "Effect of Confidence and Explanation on Accuracy and Trust Calibration in AI-Assisted Decision Making," *Proceeding of ACM Conference on Fairness, Accountability, and Transparency (FAT*)*, pp.295-305, 2020. (accessed Jul. 18, 2026)
doi: <https://doi.org/10.1145/3351095.3372852>
- [14] Z. Buçinca, M. B. Malaya, and K. Z. Gajos, "To Trust or to Think: Cognitive Forcing Functions Can Reduce Overreliance on AI in AI-Assisted Decision-Making," *Proceeding of ACM on Human-Computer Interaction (CSCW1)*, Article 188, 2021. (accessed Jul. 18, 2026)
doi: <https://doi.org/10.1145/3449287>
- [15] F. Poursabzi-Sangdeh, D. G. Goldstein, J. M. Hofman, J. Wortman Vaughan, and H. Wallach, "Manipulating and Measuring Model Interpretability," *Proceeding of ACM CHI Conference on Human Factors in Computing Systems*, Article 237, 2021. (accessed Jul. 2, 2026)
doi: <https://doi.org/10.1145/3411764.3445315>
- [16] G. Bansal, T. Wu, J. Zhou, R. Fok, B. Nushi, E. Kamar, M. T. Ribeiro, and D. S. Weld, "Does the Whole Exceed Its Parts? The Effect of AI Explanations on Complementary Team Performance," *Proceeding of ACM CHI Conference on Human Factors in Computing Systems*, 2021. (accessed Jul. 2, 2026)
doi: <https://doi.org/10.1145/3411764.3445717>
- [17] B. Shneiderman, "Human-Centered Artificial Intelligence: Reliable, Safe & Trustworthy," *International Journal of Human - Computer Interaction*, Vol.36, No.6, pp.495-504, 2020. (accessed Jul. 2, 2026)
doi: <https://doi.org/10.1080/10447318.2020.1741118>
- [18] S. Raisch and S. Krakowski, "Artificial Intelligence and Management: The Automation - Augmentation Paradox," *Academy of Management Review*, Vol.46, No.1, pp.192-210, 2021. (accessed Jul. 18, 2026)
doi: <https://doi.org/10.5465/amr.2018.0072>
- [19] A. Fügener, J. Grahl, A. Gupta, and W. Ketter, "Cognitive Challenges in Human - Artificial Intelligence Collaboration: Investigating the Path Toward Productive Delegation," *Information Systems Research*, Vol.33, No.2, pp.678-696, 2022. (accessed Jul. 18, 2026)
doi: <https://doi.org/10.1287/isre.2021.1079>
- [20] S. Amershi, D. Weld, M. Vorvoreanu, A. Fourney, B. Nushi, P. Collisson, J. Suh, S. Iqbal, P. N. Bennett, K. Inkpen, J. Teevan, R. Kikin-Gil, and E. Horvitz, "Guidelines for Human-AI Interaction," *Proceeding of ACM CHI Conference on Human Factors in Computing Systems*, 2019. (accessed Jul. 2, 2026)
doi: <https://doi.org/10.1145/3290605.3300233>
- [21] J. Sweller, "Cognitive Load During Problem Solving: Effects on Learning," *Cognitive Science*, Vol.12, No.2, pp.257-285, 1988. (accessed Jul. 2, 2026)
doi: https://doi.org/10.1207/s15516709cog1202_4
- [22] S. K. Card, "The Model Human Processor: A Model for Making Engineering Calculations of Human Performance," *Proceeding of Human Factors Society Annual Meeting*, Vol.25, No.1, pp.301-305, 1981. (accessed Jul. 2, 2026)
doi: <https://doi.org/10.1177/107118138102500180>
- [23] S. Cao, C. Gomez, and C.-M. Huang, "How Time Pressure in Different Phases of Decision-Making Influences Human-AI Collaboration," *Proceeding of ACM on Human-Computer Interaction (CSCW2)*, pp.1-26, 2023. (accessed June. 21, 2026)
doi: <https://doi.org/10.1145/3610068>
- [24] S. Swaroop, Z. Buçinca, K. Z. Gajos, and F. Doshi-Velez, "Accuracy-Time Tradeoffs in AI-Assisted Decision Making under Time Pressure," *Proceeding of International Conference on Intelligent User Interfaces (IUI)*, pp.138-154, 2024. (accessed June. 21, 2026)
doi: <https://doi.org/10.1145/3640543.3645206>
- [25] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research," in *Human Mental Workload*, P. A. Hancock and N. Meshkati, Eds., Amsterdam, Netherlands: North-Holland, 1988, pp.139-183. (accessed Jul. 19, 2026)
doi: [https://doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9)
- [26] P. Bhavsar, B. Srinivasan, and R. Srinivasan, "Pupillometry Based Real-Time Monitoring of Operator's Cognitive Workload to Prevent Human Error during Abnormal Situations," *Industrial & Engineering Chemistry Research*, Vol.55, No.12, pp.3372-3382, 2016. (accessed Jul. 19, 2026)
doi: <https://doi.org/10.1021/acs.iecr.5b03685>
- [27] N. B. Sarter and D. D. Woods, "How in the World Did We Ever Get into That Mode? Mode Error and Awareness in Supervisory Control," *Human Factors*, Vol.37, No.1, pp.5-19, 1995. (accessed Jul. 19, 2026)
doi: <https://doi.org/10.1518/001872095779049516>
- [28] S. K. Card, T. P. Moran, and A. Newell, "The Keystroke-Level Model for User Performance Time with Interactive Systems," *Communications of the ACM*, Vol.23, No.7, pp.396-410, 1980. (accessed Jul. 19, 2026).
doi: <https://doi.org/10.1145/358886.358895>

저 자 소 개



안 성 희

- 2005년 : University of London, Goldsmiths College, Research Fellow
- 2013년 : Brunel Univ. (UK), School of Design & Engineering, Design Innovation & Strategy 전공 (Ph.D)
- 2013년 ~ 2016년 : Hunan Univ. China Engineering & Design College 디자인학부 조교수
- 2016년 9월 ~ 현재 : 홍익대학교 IDAS 대학원 및 디자인컨버전스학부(세종) 부교수
- 한국방송 미디어공학회 / 한국HCI학회 이사
- ORCID : <https://orcid.org/0000-0002-7717-2412>
- 주관심분야 : HCI, XR-UX, Human-centered AI