

# From Gap to Synergy: Enhancing Contextual Understanding through Human-Machine Collaboration in Personalized System

UIST '23

Paper Review by Eunseo Yang

2024.01.17



# Background

- **시스템의 개인화의 필요성이 두각**

시나리오, 선호도, 처해있는 컨텍스트에서 상당한 차이가 나타남

- **AI, IoT 기술의 발전**

강력한 컨텍스트 인식 능력 생기기 시작

- **스마트폰 센서를 통한 다양한 정보 수집 가능**

물리적 맥락 (위치, 활동), 소셜 맥락 (연락처, 메시지) 감지/예측하여 특정 자동화된 서비스를 제공함

- **데이터 기반 기계 학습 알고리즘의 한계**

맞춤형 컨텍스트에서 사용자의 선호도를 학습하기가 아직 어려움

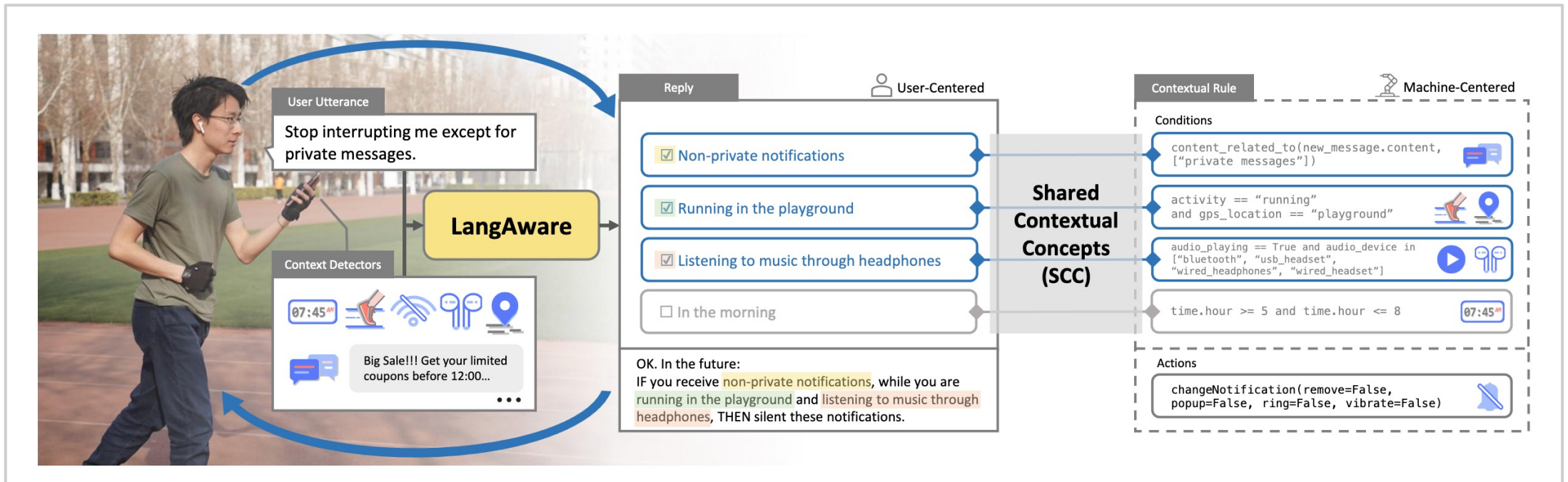
(개인화 지식을 활용하는 데 사용자를 효과적으로 사용하지 못함) → 인간과 기계 간의 컨텍스트 구축 상당한 간극이 존재

# Previous Attempts

- 미리 정의된 간단하고 제한된 맞춤 설정 → 사용자의 다양 요구 수용 x
- 단축키 및 포커스 모드와 같은 솔루션은 제한된 솔루션
- 사용자 친화적이지 않는 복잡한 작업

# Idea

- 자연어를 사용해서 컨텍스트 규칙을 설정하자. **“Shared Contextual Concepts (SCC)”**
- LLMs 활용 + 센서 정보해서 user's context를 좀 더 잘 이해하고, 그들의 언어를 이해해보자!



# Related Work

## Context-aware computing

: 사용자의 컨텍스트를 기반으로 맞춤형 서비스를 제공하는 기술

- 시간, 위치, 활동, 앱 사용 기록, 근처 Wi-Fi/Bluetooth 기기, 연락처와 메시지 내 내장된 소셜 정보 (물리적 감지 정보)
- 풍부한 컨텍스트 정보는 권장 시스템, 메시지 관리, 앱 추천, 배터리 관리, 볼륨/밝기 관리 등 다양한 개인화된 자동 서비스 제공
- 인식 기술로서 데이터 기반 ML 접근법을 주로 사용해 옴

## 문제 및 개선 점

- 몇 달에 걸친 기간 동안 수집된 많은 양의 사용자 데이터 필요
- 맞춤형 컨텍스트 및 행동 데이터가 희소할 때에는 개인화된 지식을 얻는 데 비효율적
- 사용자에게 불투명한 시로 작동하여 시스템의 행동 이해/제어 어려움

→ **사용자를** 효과적으로 컨텍스트 **분석에 참여시키는** 것이 중요!

**이는 사용자가 개인화된 지식을 제공하고 시스템에 대한 이해와 제어를 높임**

# Related Work

## End User Context Construction

- 사용자들을 컨텍스트 분석에 참여시키는 연구도 있긴 했음
- Trigger-action programming “이벤트 발생하면 조건 충족 동안 조치를 취하는” 패러다임을 주로 사용
- 주로 사용자들은 자신의 필요사항을 표현하기 위해 IF-THEN 규칙을 사용하고, 이를 이용해서 스마트 홈에서 요구를 표현했다고 함

## Main Findings

- 대부분의 참여자들이 프로그래머가 아니라서 프로그래밍 threshold 낮춰야 했음.  
기계의 세부 사항을 알아야 하고, 컨텍스트 조건을 작성하기가 애초에 어렵다는 점이 문제.
- 이를 위해 고수준 추상화를 제안 했어야 했음 → GUI 개선 → 높은 비용
- 자연어 시도도 있었는데, 모든 시도들이 정해진 컨텍스트 조건에 접근하기 위함을 다뤘고,  
새롭게 현장에서 사용자마다 다른 새로운 컨텍스트를 만드는 문제는 다루지 않음

# Study Design

## 인간과 컴퓨터의 컨텍스트 인식 방법 간에는 이해의 간극이 존재

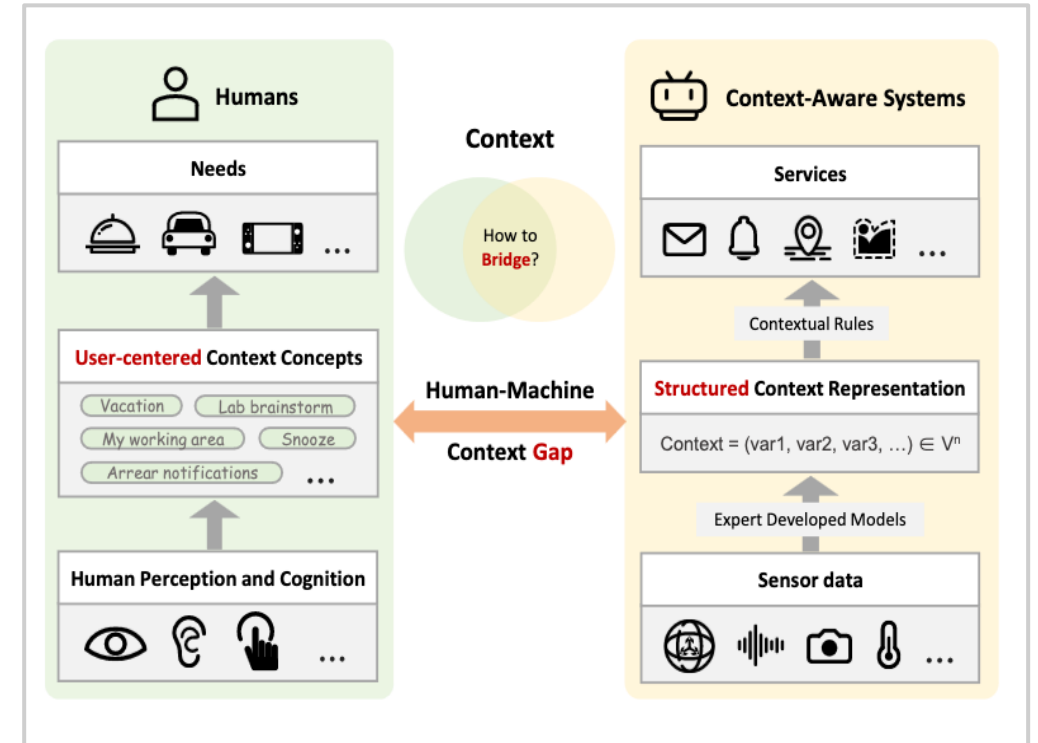
(1) 인간과 컴퓨터는 서로 다른 지각 채널을 가지고 있어 수집하는 입력 정보의 차이가 존재

(2) 인간은 상황을 추상적인 개념 (일하는 중), 구체적인 개념(사무실에서)을 통해 이해  
이에 반해 기계는 구조화된 컨텍스트 변수와 서비스 인터페이스를 사용해서 맥락에 규칙을 부여

이러한 불일치가 최종 사용자에게는 기계 능력에 대한 부족한 이해,  
기계 규칙에서 결과를 예측하는 데에 어려움, 사용 용이성 x 문제로 이어짐

### 문제

- 기계는 고수준 의미를 요약하지 못해 사용자가 맞추어서 서비스를 받아야 함
- 간극을 줄이기 위해서는 "효과적인 상호작용 방법" 고민
- 스마트폰 자동화 작업에서 형성적 연구해서 실제로 어떻게 컨텍스트 선호도 표현하고 규칙 생성에 협력하는지 이해해보기로 함



# Study Design

- 다양한 일상 생활에서 자신을 상상하고 상호작용하는 Wizard of Oz 디자인 채택
- 6명의 참가자 (각 40분)
- 컨텍스트 프로파일을 기반으로 한 자동화 서비스에 대한 경험 질문 ex. 방해 금지 모드 등등 (1명 제외 모두 경험)
- 고급 컨텍스트 인식 및 인간과 유사한 수준의 이해력 가진 채팅 서비스 제공 후 선호도 조사 + 인터뷰
- 이야기와 배경 사진 정해준 뒤에 선호도에 맞게 작업 조정하라고 말함  
(환경: 실험실 작업, 도서관 공부, 기숙사 휴식, 놀이터 스포츠 배경 → 조작: 메세지 관리, 일정관리, 볼륨 제어, 네트워크 전환 서비스)
- 사용자의 자연어와 컨텍스트 모두 고려해서 구현 가능한 IF-THEN 규칙을 생성하고, 자연어로 전달하는 과정

★ Formative study: 새로운 시스템, 서비스 설계/개발 전에 수행되는 초기 연구로서 새로운 아이디어나 디자인의 초기 버전에 대한 피드백을 수집하고 사용자의 요구사항과 반응을 이해하기 위해 수행됨 (개발 초기에 반복됨)



# Study Design

## Findings

- 사용자가 일반적으로 필요한 서비스를 정확히 설명할 수 있었다. → 자연어 Input의 잠재력 확인
- 그러나 컨텍스트의 조건을 전달하는 데는 약간의 어려움이 있었다. → 자연어 처리 기술의 문제점 확인

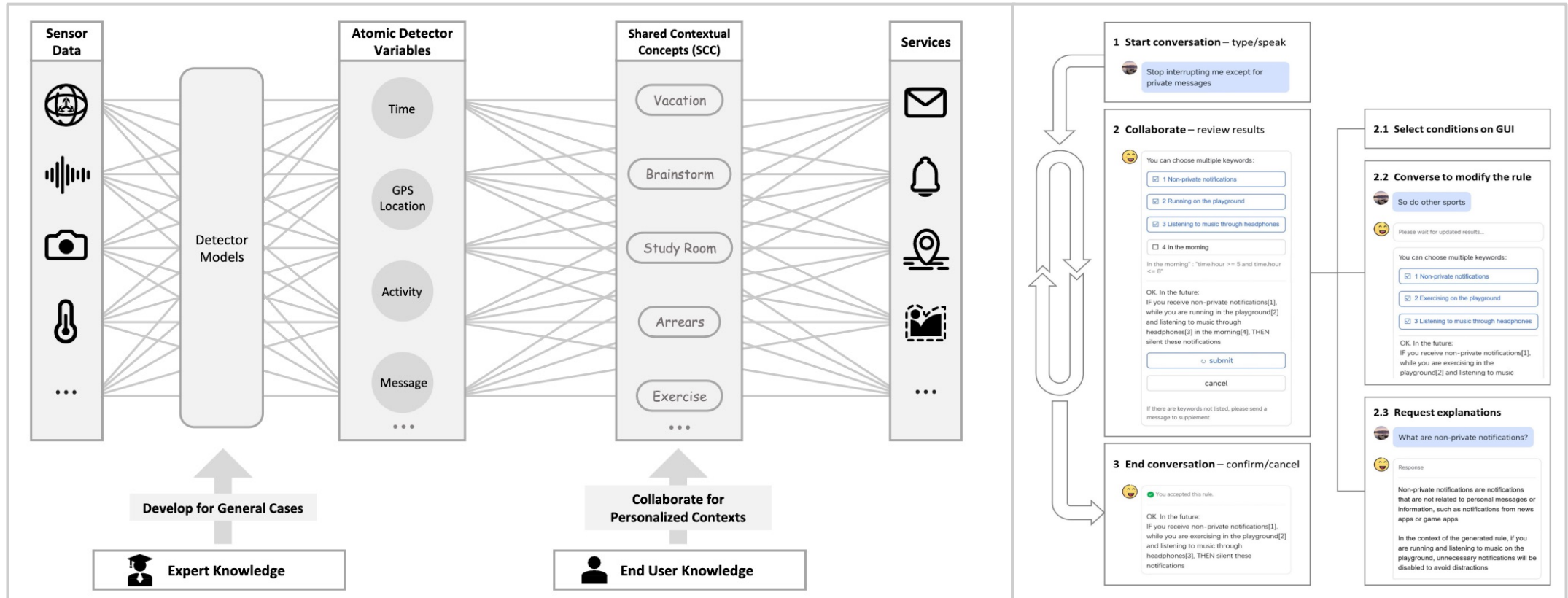
## 발견한 도전 과제

- 컨텍스트 표현에서의 사용자 말의 모호성 → 강력한 언어 이해 능력 필요
- 사용자 입력 언어에서의 컨텍스트 정보 생략 → 모든 걸 말해야 한다면 언어 표현 비용이 늘어날 것이라는 인터뷰
- 이해 가능한 인터페이스를 제공 해야 한다는 점
  - 내가 말한 것이 기계가 제대로 인식했는지 이해하고 싶다는 의견 → 기계 규칙을 해석하고 설명할 수 있는 필요성
  - 기계가 정한 규칙대로 사용자가 말하지 않음 (컨텍스트 특정 단어에서 뽑아온다면 → 유연하게 알아채야 함)

# LangAware

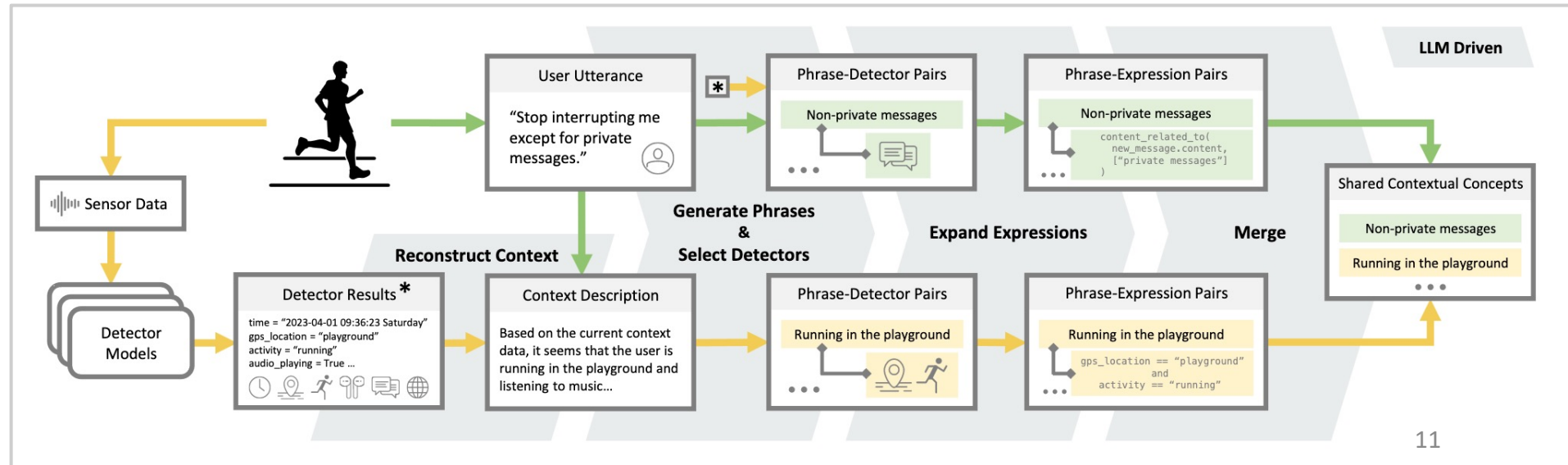
SCC는 자연어 구문과 기계 부울 표현의 쌍으로 정의하여 규칙 기반으로 진행

시스템에서 인식할 수 없는 사용자 발화의 경우, 해당 표현이 없다면 SCC 생성 불가능 + 피드백이 제공됨



# LangAware

- 활용한 특징들: 현재 시간, 위치, 전화 활성 앱, 사용자의 활동 등 (하드웨어 센서 뿐 아니라 GUI, 메시지, 다른 SW 센서도 포함)
  - \* 이미 전문가가 구축한 계산 모델이 존재 (시계열 데이터에서 사용자 동작 상태 계산하기 위한 신경망) or 화이트 박스 (GPS 좌표 기반 사용자 위치 정보 결정)
- 이벤트: 새로운 메시지 수신, 위치 변경, 소음 변경 등
- 5분 간격으로 Variable\_name 설정 (컨텍스트 상태 값)
- 센서 정보는 있는데 구문이 감지 되지 않으면 후속 피드백 들어감



# Implementation

- Feishu Bot을 사용해서 프론트엔드 대화 인터페이스 구축 (사용자는 타이핑, 음성 통해 텍스트 인풋 가능)
- 원격 서버에서 제어(컨텍스트 규칙 생성, 서비스 매핑)되며 스마트폰 백그라운드에서 실행
- 서비스 액션 구성은 API 통해서 볼륨 조절, 새 메시지에 대한 통지 모드 수정, 통화 모드 변경, 네트워크 전환 등에서 선택 (or LLM 자체에서 Python 함수 호출 구문 따라 생성)
- ChatGPT API 3.5 turbo 모델 사용

# Experiment

- 16명의 참가자 대상 안드로이드 스마트폰을 사용하여 진행
- 현장 컨텍스트 센서 정보 통합하지 않고, 언어로만 사용하는 기존 시스템과의 성능 평가
- 한 사람당 4가지 task 부여 (기능 무/유)

## 1번의 trial 끝나면, 다음 질문에 대한 답변을 얻음

- How **mentally demanding** was it to use this assistant?
- How **physically demanding** was it to use this assistant?
- How **hurried or rushed** was the pace of the interaction with this assistant?
- How well do you think this assistant **understands your intentions**?
- How **satisfactory** do you think the final result generated by this assistant is?

# Experiment

8회가 모두 끝나면,

- To what extent did this assistant **increase your understanding of its context-aware capabilities?**
- **How easy was it to learn** how to use this assistant?
- **How easy was it to generate** context rules using this assistant?
- **How likely** are you to use this assistant again in the future?

## 전문가 대상 추가 성능 테스트 진행

- 컨텍스트 인식 시스템 설계/개발 경험의 전문가 3명 초청해서 생성된 컨텍스트 규칙을 평가.  
SCC의 충분성(머신 표현이 구문을 의미적으로 얼마나 잘 나타내는지) 및 필요성(구문이 의미를 얼마나 잘 나타내는지)

# Experiment

## Tasks: 12개의 캠퍼스 시나리오 (4가지 큰 범주)

- **조용한 도서관/스터디 룸에서 주변 학생들과 함께할 때:**

동영상 시청 중에 갑작스런 오디오 출력 방지; 헤드폰을 착용하고 포커스 앱을 사용할 때 자동으로 경음악 재생; 블루투스 헤드셋이 갑자기 연결 해제될 때 오디오 출력하지 않음; 광고 알림 차단하여 공부에 방해받지 않도록 함

- **시끄러운 운동장에서:**

헤드폰을 착용하고 달리며 음악을 듣는 중에 비공개 채팅 메시지에 방해받지 않도록 함; 운동장에 들어가면서 달리기를 시작하면 자동으로 운동 체크인 기록을 시작하고 운동을 멈추면 자동으로 기록을 종료함.

- **룸메이트가 있는/없는 기숙사에서:**

선생님의 메시지에 대한 벨소리 알림 설정; 음식 배달 관련 메시지에 대해 크게 알림; 아침 수업이 있는 경우 알람이 켜져 있는지 확인; 기숙사에 다른 사람이 있을 때 비디오에서 오디오를 출력하지 않음.

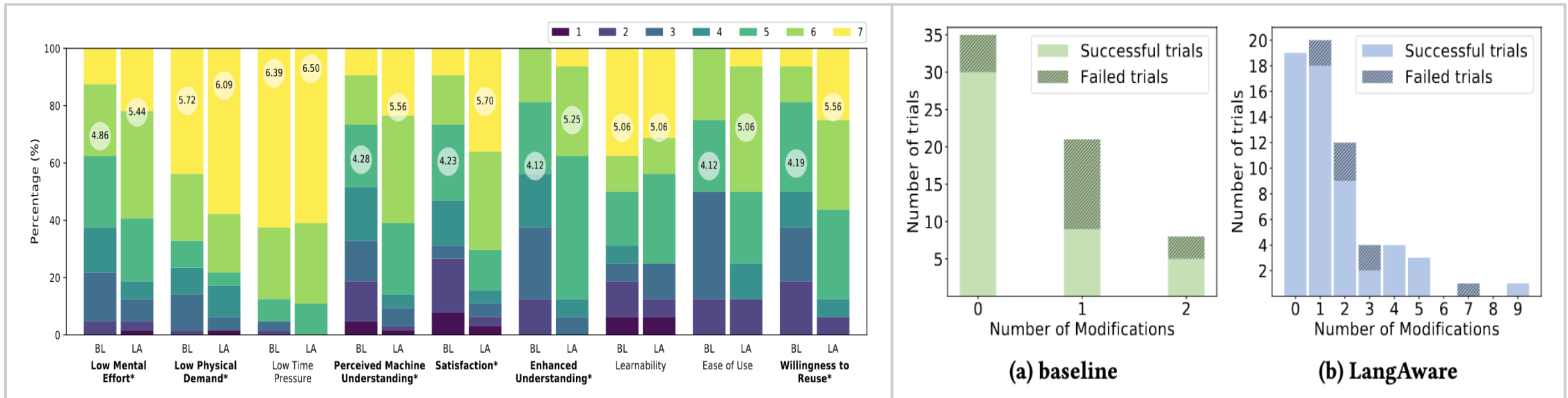
- **시끄러운 카페에서:**

메시지 수신 시 벨소리와 진동; 음악을 듣거나 전화 통화 시 볼륨을 높임.

# Experiment

## Results

- 사용자가 느끼기에 고품질에 SCC (평균 전문가 만족도 점수 4.50/5 및 필요성 점수 4.50/5)
- 성공률 87.5% 개선
- 사용자 만족도, 기계 이해, 정신적 노력, 신체적 수요와 같은 주관적인 평가 메트릭에서도 기존 시스템을 크게 능가





# Discussion

- 각 작업에서 참여자들은 각각 다양한 개인화된 요구와 서비스를 바랍
- 각 작업에서 참여자들의 컨텍스트 추론 방식이 다 다름.

누구는 표현에 숫자 값을 조절하는 것을 선호, 누구는 추상화된 언어로만 소통 선호

Ex. "조용한 공공 장소" 환경을 주변 센서로부터 추론 해야 하는 사람이 있었고,

"주변 사람이 있는" 환경을 근처 전화번호 감지기를 사용하여 SCC를 만들어내는 사람이 있었음

즉, 추론 방식과 추론 결과가 다 달라진다는 것 (사람들이 말하는 것에 따라서) 유연성 필요

- 각기 다른 감지기 변수의 선택을 통해 추론된다는 의미!
- 자연어를 통한 텍스트 의미론 "content related to" 감지기의 도입은 키워드 자체가 본질적으로 개인화 되어 있어서 충분히 개인화를 제공 ex. 발신자를 상사, 동료, 멘토 제외 메시지 금지! → 각 개인의 중요성에 대해 이해할 수 있다는 것

# Limitation

- 사용자가 명시하지 않은, 인식하지 못한 **근본적 목표를** 다루지 않음  
사용자가 SCC 추적 후에 SCC 의미론을 분석하고 클러스터링할 필요  
**장기 작동을 위한 데이터 기반 사용자 행동 패턴 마이닝 방법을 조사할 필요**  
이로서 근본적 목표를 파악하고 사용자의 실제 문제에 대한 맞춤형 솔루션을 제공하는 능력이 필요
- LLM 적용이 아직 한계. 비교적 느린 계산 속도는 사용자 경험을 저해  
LLM의 토큰 제한으로 인한 프롬프트의 예제 수를 제한하는 것이 처리 능력을 제한할 수 있음  
미세 조정된 모델을 사용하여 완화 가능

# Conclusion

- 사용자가 자연어를 사용하여 현장에서 맞춤형 컨텍스트를 구성할 수 있도록 하는 혁신적인 협력적 접근을 하자!
- 이 작업은 인간-기계 상호작용의 직관성과 적응성을 향상시키며,  
IoT 시나리오에서 더 많은 맞춤형 컨텍스트 인식 애플리케이션의 개발을 용이하게 하는 접근법
- 이를 통해 저수준 센서 감지기를 고수준 컨텍스트에 의미적 연결,  
효과적인 사용자 참여를 위해 이해하기 쉬운 자연어 피드백을 제공
- 시스템 개인화를 위한 연구 흐름에 기여

# Q&A, Takeaways

- 사람들은 생각보다 더 원하는 것이 다르고, 원하는 것을 요구하는 방법도 다르다.  
“개인화”하기 위해서는 보다 세심한 정보 수집 및 처리가 필요하다.
- 사용자의 근본적인 목표를 기계가 이해하고, 그에 맞는 서비스를 반응해주기 위해 필요한 피쳐들을 고민해야 한다.  
사람들은 자연어로 모든 것을 말해주지 않는다. 정확한 상황 설명 및 구체적인 정보 제시를 하지 않기 때문에 추가적인 정보 수집이 필수적이다. (사용자의 행동 패턴, 검색 기록 등의 정보 퓨전을 통한 세심한 개인화 방법이 필요하다.)
- 사용자의 상황 파악에 있어서 단순히 센서 정보와 사용자의 말을 사용하는 것이 아니라 그 둘을 함께 사용하고, 사용자가 확인하는 과정까지 거친다면 더욱 정확한 상황 이해가 가능하다. 시스템 적의 성능 향상도 좋지만, 인간의 협력도 중요하기 때문에 함께 최종안을 결정할 수 있는 매커니즘이 필요하다.