



# PaTAT: Human-AI Collaborative Qualitative Coding with Explainable Interactive Rule Synthesis

2023.02.19

오하나

# PaTAT: Human-AI Collaborative Qualitative Coding with Explainable Interactive Rule Synthesis

Simret Araya Gebreegziabher\*

University of Notre Dame  
Notre Dame, IN, USA  
sgebreeg@nd.edu

Yihao Meng

Xi'an Jiaotong University  
Xi'an, China  
ymeng2@nd.edu

Zheng Zhang\*

University of Notre Dame  
Notre Dame, IN, USA  
zzhang37@nd.edu

Elena Glassman

Harvard University  
Cambridge, MA, USA  
glassman@seas.harvard.edu

Xiaohang Tang

University of Liverpool  
Liverpool, UK  
sgxtang4@liverpool.ac.uk

Toby Jia-Jun Li

University of Notre Dame  
Notre Dame, IN, USA  
toby.j.li@nd.edu

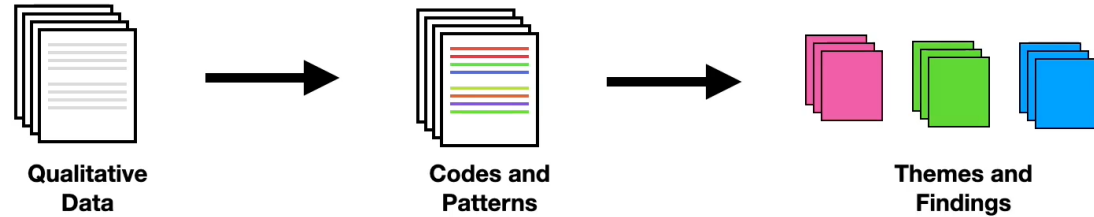
ACM CHI'23 April 인용 5회, 다운로드 858회  
Google scholar에서 인용 18회

# PaTAT: Human-AI collaborative coding



## Abstract

- 배경
  - AI를 활용해 데이터에 주석 다는 작업이 많다.
- 문제
  - Thematic analysis(텍스트 자료, 인터뷰 내용, 설문조사 응답에서 주제나 패턴을 식별하고 분류하는 과정)에서 Human-AI 협업이 어렵다.
- 방법
  - 인터뷰를 통해 중요한 디자인 요소 선택하여 human-ai 협업 툴을 만들.
  - 사용자가 데이터에 주석을 다는 것을 토대로 배우고, 실시간으로 텍스트에 적용하는 상호작용 프로그램 PaTAT을 설계
- 결과
  - 규칙 기반 레이블링 보조 시스템(Cody)보다 표현력이 넓다.
  - BERT과 비교시 예시가 50~100개 이하일때 PaTAT이 우위를 차지함. Thematic analysis는 많은 데이터가 많이 있지 않기에 PaTAT이 적합하다고 판단.
  - User study: Bert, manual과 비교했을때 PaTAT는 사용의 용이성에서 1위를, 유용성에서 2위로 평가.
- 의의
  - 사용자 연구를 통해 PaTAT의 유용성과 효과성을 입증
  - 모호함과 불확실성을 다루는 thematic analysis의 human-AI 협력 설계에 대한 시사점 제시



# PaTAT: Human-AI collaborative coding



## Overview

paTAT

Rank By: Default | Group By: None | RETRAIN IN 10

3% You | 20% Model

Theme: price | Multi-label:  | Binary:

Patterns for price RENAME MERGE SPLIT DELETE

Prediction Set				
Pattern	Weight	Fscore	Recall	Precision
[price]\$MONEY	1.00	0.86	0.75	1.00
\$MONEY+NUM++NOUN	0.70	0.40	0.25	1.00

All Patterns

Pattern	Fscore	Recall	Precision
NUM [price]	0.86	0.75	1.00
[price]\$MONEY	0.86	0.75	1.00
NUM (price)	0.75	0.75	0.75

The bbq is wonderful and they have maybe the best collard greens i have had in the us .  
product

The staff was very friendly and they have a wide selection of drinks .  
service product

The place was beautiful , immaculately clean , comfortable bedding ( lovely sheets ) and breakfast was delicious .  
atmosphere

The food was not as good as i expected it to be , it was overpriced  
price

Had a really awkward time ordering a beer because nothing on the menu was stocked .  
product

It is a little pricy though , i drop about \$ 200 per visit .  
Predicted price: ACCEPT REJECT  
+ Add Label SIMILAR

[price]\$MONEY: 1.00  
\$MONEY+NUM++NOUN: 0.70

I tried them out today for the first time . food was great and all the staff was very friendly .

# PaTAT: Human-AI collaborative coding



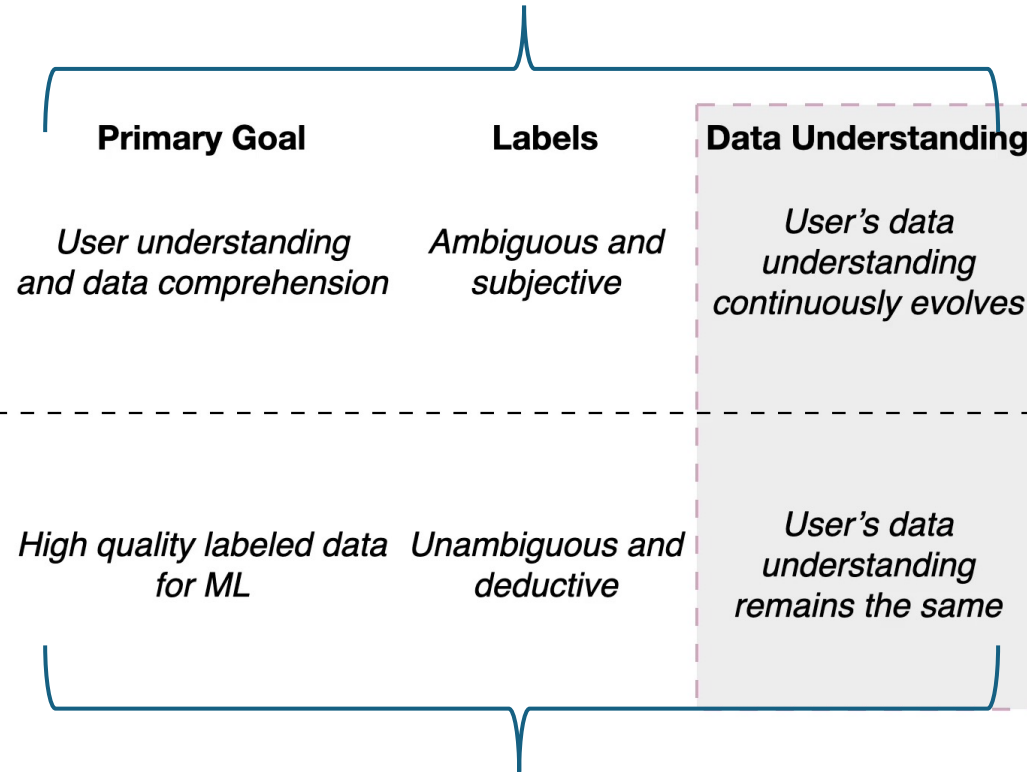
## Introduction: Problem

데이터를 검토하는 과정에서 새로운 레이블을 유동적으로 추가함.  
그래서 레이블이 자주 수정되거나, 병합되거나, 분할된다.

Data annotation &  
qualitative coding  
차이점

이 논문의 focus

Data Annotation  
기존 ML 방식



ML은 qualitative coding의 복잡한 특성을 수용하고 지원하는데 적합하지 않다.

”정답”이 있는 레이블을 사용  
예) 이미지 내 객체 분류/ 텍스트의 개체 추출 → 레이블 집합이 고정적

# PaTAT: Human-AI collaborative coding



## Introduction: Solution

- **기존 QDA 도구** (예: NVivo, MAXQDA, ATLAS.ti)
  - 키워드 검색, 데이터 포인트 일괄 주석 기능 제공.
  - labeling 과정 가속화에 도움이 되지만, 사용자 학습에는 한계가 있음.
- **ML 지원 QDA 도구** (예: INCEption, BRAT)
  - 모델의 효율성과 정확성은 좋지만, 새로운 label을 찾는 유연성 부족.
- **규칙 기반 텍스트 분석**
  - Tempura, Cody와 같은 도구에서 사용됨.
  - 사용자의 직관적 이해와 데이터 조작 가능성 제공.
  - 구조화된 데이터에 초점을 맞춤: 유연성 한계.
- **PaTAT**
  - Program synthesis: 자동으로 데이터 positive/negative 예시에서 패턴을 인식해주는 기술 사용
  - Generalize할 수 있는 부분을 데이터에서 표시할 수 있도록 함.
  - 사용자가 모델 학습 상태를 탐색하고, label 추천 이유를 이해할 수 있는 인터페이스 제공.

# PaTAT: Human-AI collaborative coding



## Formative Study

### 해결

- PaTAT은 실시간으로 패턴을 학습해서 데이터에 주석을 추천해준다.
- 2년 이상 qualitative coding 경험이 있는 5명 인터뷰.

### 3.5 Design Goals

We summarize the following five key design goals from the findings of our formative study, which helped guide the design of PaTAT described in Section 4:

- **DG1:** Improve user's understanding of the dataset, i.e., help them (1) discover low-level nuance and commonalities across data items and (2) develop high-level mental models of the dataset (Section 3.4.1).
- **DG2:** Generate explainable coding recommendations that allow users to understand the AI's rationale, verify the soundness of the AI's reasoning, and provide affordances for the user to correct the AI's current model (Section 3.4.3).
- **DG3:** Accommodate the iterative and ambiguous nature of qualitative coding at different granularity levels (Section 3.4.1)
- **DG4:** Ensure user agency in the coding process so that users have the authority to make the ultimate annotation decisions when necessary (Section 3.4.2)
- **DG5:** Facilitate effective human-AI collaboration to resolve uncertainty and ambiguity in annotations (Section 3.4.4)

### PaTAT 설계 목표

- DG1: 데이터셋에 대한 사용자의 이해 개선
- DG2: 해석 가능한 label 추천 생성
- DG3: 질적 label의 반복적 및 모호한 특성 수용
- DG4: label 과정에서의 사용자 주도권 확보
- DG5: 효과적인 인간-AI 협업 지원

# PaTAT: Human-AI collaborative coding

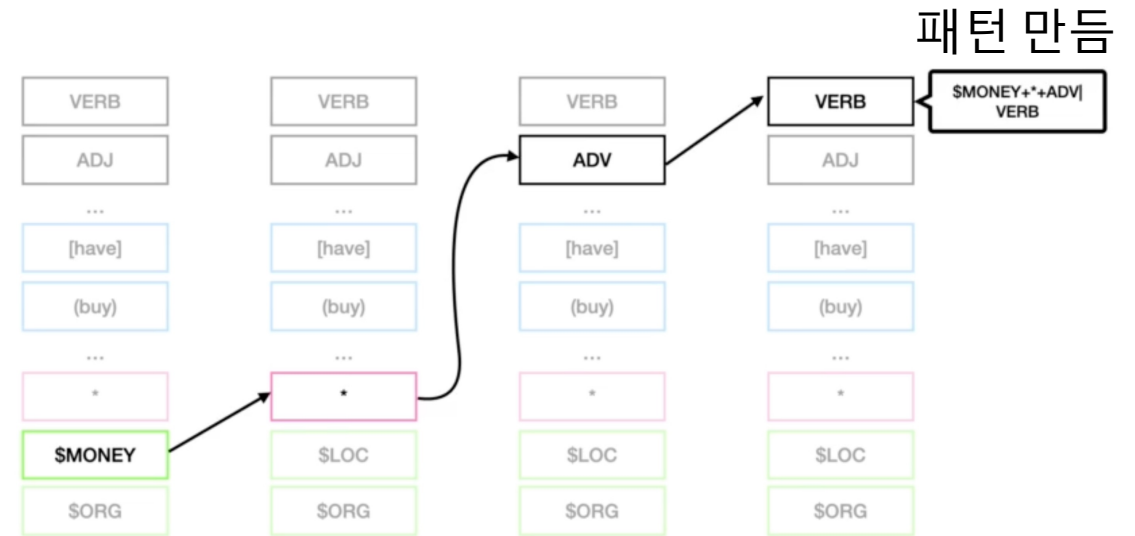


## PaTAT system

### 패턴 인식

- 데이터의 어휘, 구문, 의미 특성에 대한 패턴 규칙을 학습하는 대화형 프로그램 방식 제안

- **Part-of-speech (POS) tags:** VERB, PROPN, NOUN, ADJ, ADV, AUX, PRON, NUM
- **Word stemming:** [WORD] (e.g., [have] will match all variants of have, such as *had*, *has*, and *having*)
- **Soft match:** (word) (e.g., (pricey) will match synonyms such as *expensive* and *costly*, etc.)
- **Entity type:** \$ENT-TYPE (e.g., \$LOCATION will match phrases of location type, such as *Houston, TX* and *California*; \$DATE will match dates; \$ORG will match names of organizations)
- **Wildcard:** \* (will match any sequence of words)



# PaTAT: Human-AI collaborative coding



## Function

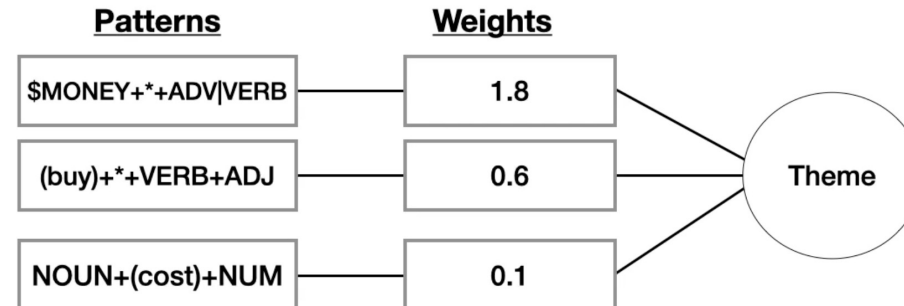
**\$MONEY+\*+ADV|VERB**

### Positive Examples

- I bought a living social deal \$30 for \$60 a while back.
- The prices were reasonable, between \$5 - \$10 per roll, but the rolls were just ok.
- One complaint is they're expensive (\$6.25 for 8) and the sizing is not consistent

### Negative Examples

- Great customer service, great food that is fresh and cooked to order.
- Everything fresh and full of flavor
- Nice open, clean and crisp atmosphere

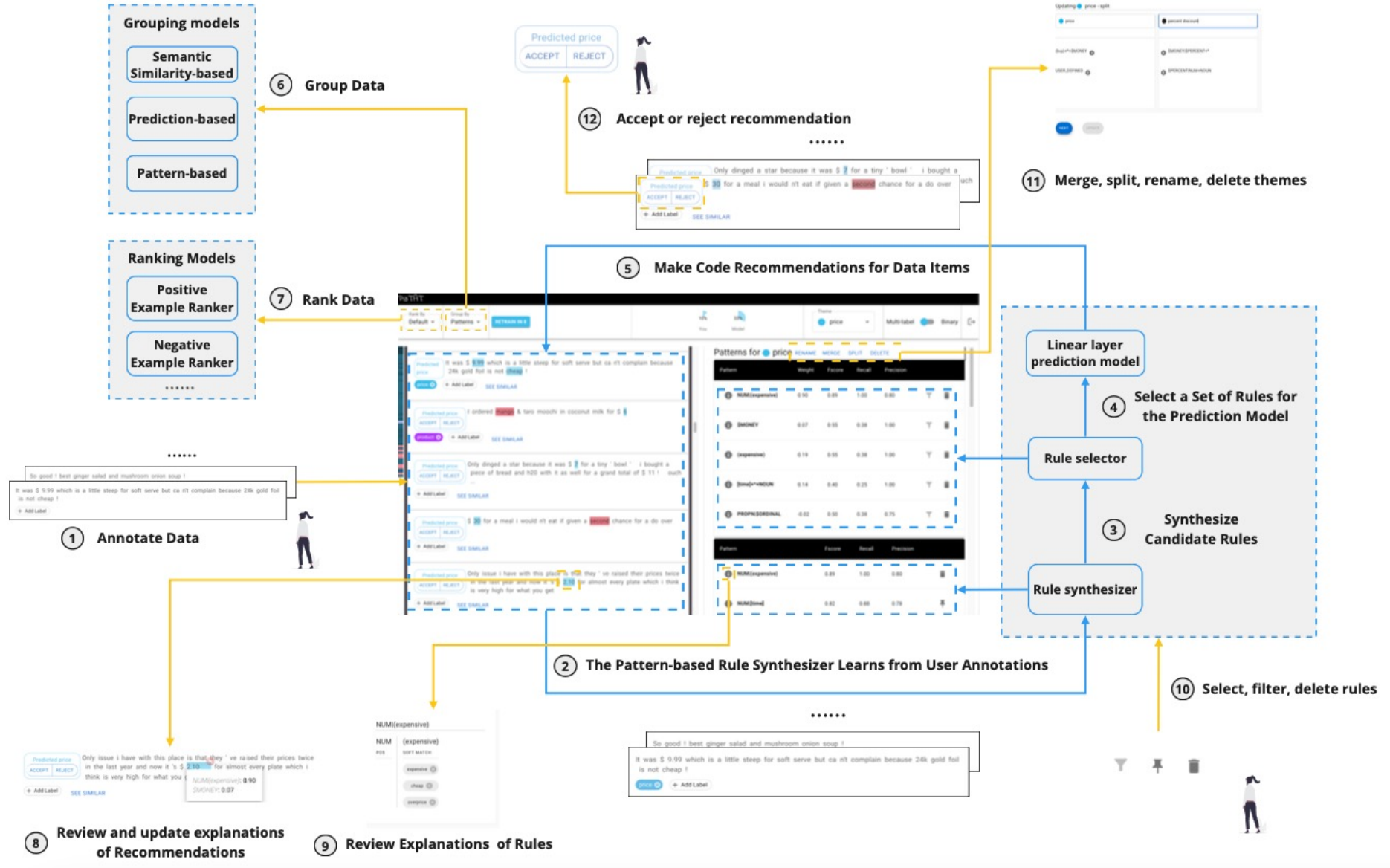


상위 5개의 가중치가 높은 패턴을 사용해서 새로운 데이터의 theme(label)을 예측

# PaTAT: Human-AI collaborative coding



## System Architecture



# PaTAT: Human-AI collaborative coding



## System Architecture

The figure illustrates the PaTAT system architecture through several key components:

- (1) Dataset Grouping:** The main interface shows a list of text snippets grouped into thematic categories like 'Group-2' and 'Group-3'. Each snippet is annotated with predicted labels such as 'price' or 'product'.
- (2) Pattern Filter:** A table titled 'Patterns for price' displays a 'Prediction Set' of regular expressions. The table includes columns for 'Pattern', 'Weight', 'Fscore', 'Recall', and 'Precision'.

Pattern	Weight	Fscore	Recall	Precision
(pay)(sale)	0.87	0.94	0.89	1.00
(sale)(food)*+AUX	0.65	0.67	0.66	0.88
(place)*+VERB	0.64	0.36	0.22	1.00
(customer)*+service	0.04	0.36	0.22	1.00
- (3) Item Interaction:** A tooltip for a highlighted item shows the predicted price and the associated pattern: `(expensive)(buy): 0.37`.
- (4) Theme Splitting:** A 'Updating price - split' panel allows users to refine a theme by selecting sub-themes like 'money', 'percent', or 'money' from a list.
- (5) Pattern Manipulation:** A panel for the pattern `(expensive)(buy)` shows 'SOFT MATCH' options for synonyms: 'expensive', 'cheap', 'overprice', 'buy', 'deal', and 'purchase'.

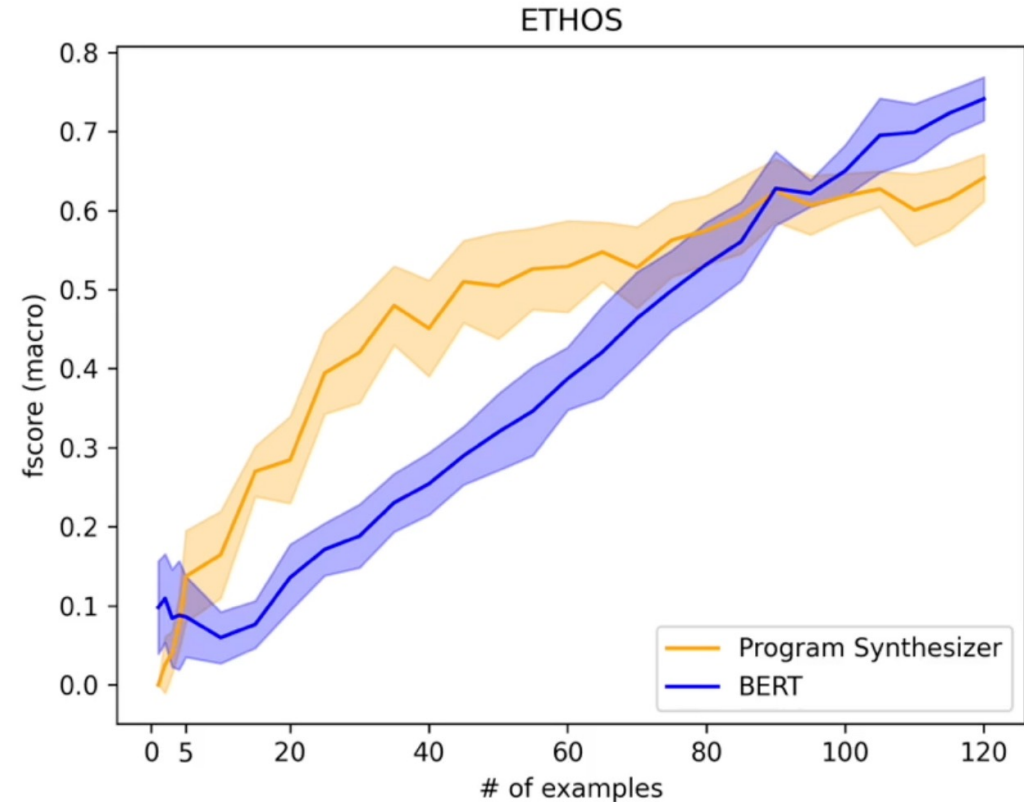
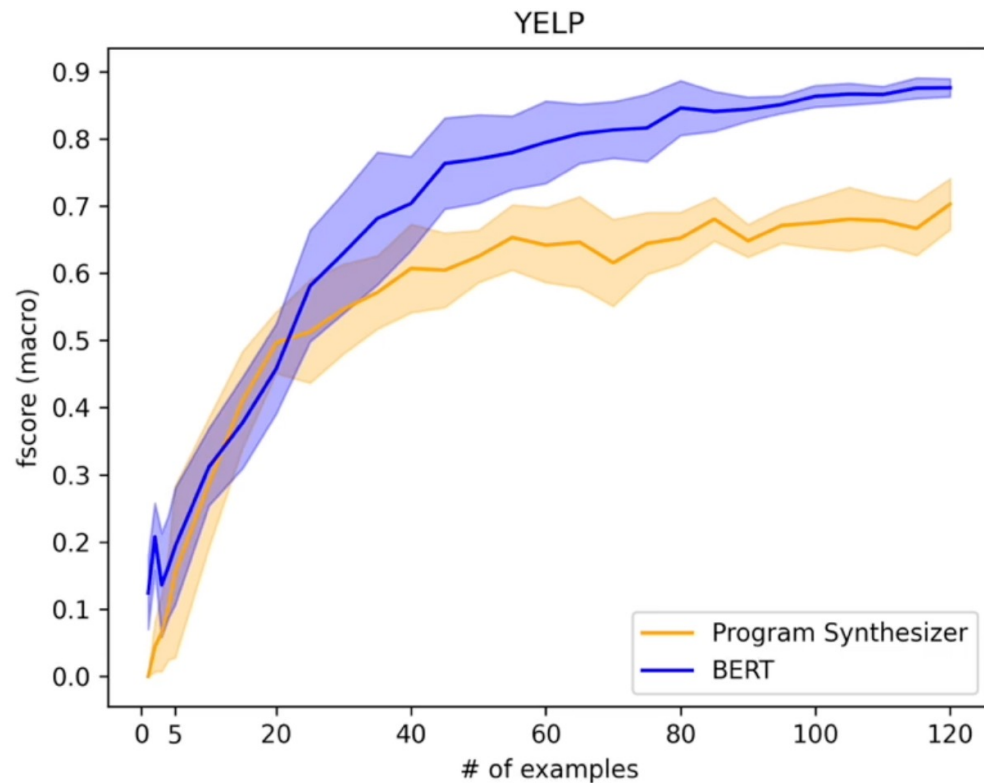
**Figure 3: PaTAT provides different methods for organizing and interacting with data: (1) The dataset can be grouped by thematic similarity. (2) The pattern filter shows all the items that matched that pattern. (3) When a user hovers over a highlighted item, all matching patterns are displayed. (4) As the user discovers new sub-themes within a theme, they have the option to split the theme using the patterns associated with the theme. (5) Users can directly manipulate patterns by dictating what synonyms (soft-matches) are allowed to match.**

# PaTAT: Human-AI collaborative coding



## Performance

Compare with pretrained BERT through few shot learning



Thematic analysis(인터뷰, 설문조사 응답 ..등)은 많은 데이터가 많이 없기에 PaTAT이 적합함

# PaTAT: Human-AI collaborative coding



## User Study

- **참가자:** qualitative analysis 경험이 있는 8명의 참가자를 대상으로 사용성 실험
- **연구 디자인:** 각 연구 세션은 약 120분 지속
- **조건:** 참가자는 MANUAL(수동 주석), BERT, PaTAT조건 하에서 시스템을 사용했습니다.
- **결과:** PaTAT은 코드 추천의 예측 정확도가 높으며, 사용자에게 데이터에 대한 새로운 통찰력을 제공하고, 질적 코딩 과정에서의 모호성과 동적 특성을 잘 처리한다.

Condition	Factor	1st Place	2nd Place	3rd Place
Manual	Ease of use	<u>6</u>	1	1
	Usefulness	0	2	6
	Perceived user control	<u>4</u>	3	1
	Facilitating human learning of data	0	4	4
BERT	Ease of use	1	3	4
	Usefulness	3	4	1
	Perceived user control	1	4	3
	Facilitating human learning of data	1	3	4
PaTAT	Ease of use	1	4	3
	Usefulness	<u>5</u>	2	1
	Perceived user control	3	1	4
	Facilitating human learning of data	<u>7</u>	1	0

# PaTAT: Human-AI collaborative coding



## Discussion & future works

### Discussion

- 데이터에 대한 친숙도와 레이블링 형성 단계에 따라 시스템으로부터 다양한 형태의 보조 필요.
- 사용자들은 AI 보조를 통해 수동으로 주석을 다는 것과 유사한 자율성을 원함.
- 초기 레이블링 형성 후, 참가자들은 PaTAT의 제안을 사용하여 레이블링의 일관성 확보
- 사용자는 모델이 무엇을 배웠는지, 왜 특정 label을 추천 하는지 이해하고, 모델 결과를 효과적으로 검증하고 수정 필요
  - 전체 패턴의 검색 범위를 지정
  - 패턴을 직접 선택하고 조정

### Future works

- 현재 PaTAT은 규칙 기반 코딩 보조 시스템(Cody)보다 표현력이 있고 유연하지만, sentiment, co-references, and commonsense knowledge와 같은 자연어 처리는 부족.
- PaTAT는 현재 한 번에 한 명의 사용자만 지원하지만, 앞으로 여러 사용자가 협력할 수 있도록.
- PaTAT의 사용이 qualitative coding 과정의 효율성에 미치는 영향에 대한 정량적 결과 부족.

# PaTAT: Human-AI collaborative coding



## Cited by

Gao J, Choo K, Cao J, Lee R and Perrault S. (2023). **CoAlCoder: Examining the Effectiveness of AI-assisted Human-to-Human Collaboration in Qualitative Analysis**. *ACM Transactions on Computer-Human Interaction*. **31**:1. (1-38). Online publication date: 29-Feb-2024.

<https://doi.org/10.1145/3617362>

Zhang Z, Ning Z, Xu C, Tian Y and Li T. **PEANUT: A Human-AI Collaborative Tool for Annotating Audio-Visual Data**. Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology. (1-18).

<https://doi.org/10.1145/3586183.3606776>

Gao J, Guo Y, Li T and Perrault S. **CollabCoder: A GPT-Powered WorkFlow for Collaborative Qualitative Analysis**. Companion Publication of the 2023 Conference on Computer Supported Cooperative Work and Social Computing. (354-357).

<https://doi.org/10.1145/3584931.3607500>

Shen H, Li T, Li T, Park J and Yang D. **Shaping the Emerging Norms of Using Large Language Models in Social Computing Research**. Companion Publication of the 2023 Conference on Computer Supported Cooperative Work and Social Computing. (569-571).

<https://doi.org/10.1145/3584931.3606955>

Ning Z, Zhang Z, Sun T, Tian Y, Zhang T and Li T. **An Empirical Study of Model Errors and User Error Discovery and Repair Strategies in Natural Language Database Queries**. Proceedings of the 28th International Conference on Intelligent User Interfaces. (633-649).

<https://doi.org/10.1145/3581641.3584067>

## Human-AI collaboration in

- CoAlCoder: Qualitative analysis
- collabCoder: qualitative analysis
- PEANUT: Annotating audio-visual data
- LLM in social computing research
- Discovery and repair strategies in NLP queries.

## Human-AI collaboration 이 어려운 분야가 있을까?

- 인터뷰에서 언어적 뉘앙스/강조점을 이해하고 반영할 수 있는 AI
- Annotating qualitative /audio-visual data 말고 다른게 있을까?

# PaTAT: Human-AI collaborative coding



Other study

## On the Design of AI-powered Code Assistants for Notebooks

Authors: Andrew M Mcnutt, Chenglong Wang, Robert A Deline, Steven M. Drucker [Authors Info & Claims](#)

CHI '23: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems • April 2023 • Article No.: 434 • Pages 1-16 • <https://doi.org/10.1145/3544548.3580940>

ACM 인용 8회, 다운로드 1226회  
Google scholar 인용 28회

문제 의식:

IDE에 사용할 수 있는 코딩 보조 서비스는 많은데 Jupyter와 같은 Notebook 형태는 Code, Markdown, Table 같이 여러 형태가 가능하기 때문에 코딩 보조 서비스를 만들기 어려움.

- 다른 연구의 기반이 될 수 있는 새로운 문제 의식 발굴
- 인터뷰를 통해 디자인 방법 선택



# Q&A