NEURAL INFORMATION PROCESSING SYSTEMS

EHRCon: Dataset for Checking Consistency between Unstructured Notes and Structured Tables in Electronic Health Records

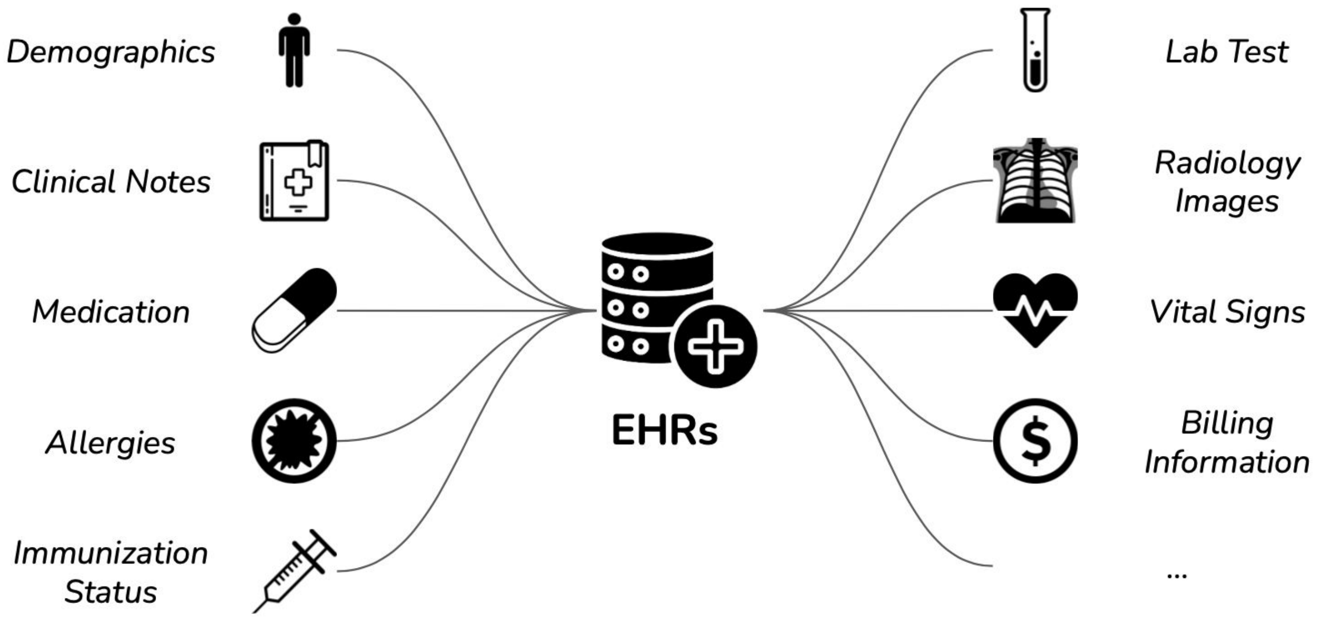
Yeonsu Kwon[1*] Jiho Kim[1*] Gyubok Lee[1] Seongsu Bae[1] Daeun Kyung[1]
Wonchul Cha[2] Tom Pollard[3] Alistair Johnson[4] Edward Choi[1]

[1]KAIST [2]Samsung Medical Center [3]MIT [4]University of Toronto

KAIST AI Kim Jaechul Graduate School SAMSUNG SAMSUNG MEDICAL CENTER Massachusetts Institute of Technology UNIVERSITY OF TORONTO

Background: EHR

- Electronic Health Records (EHRs)



Motivation: Consistency check between Clinical Notes and EHR Tables

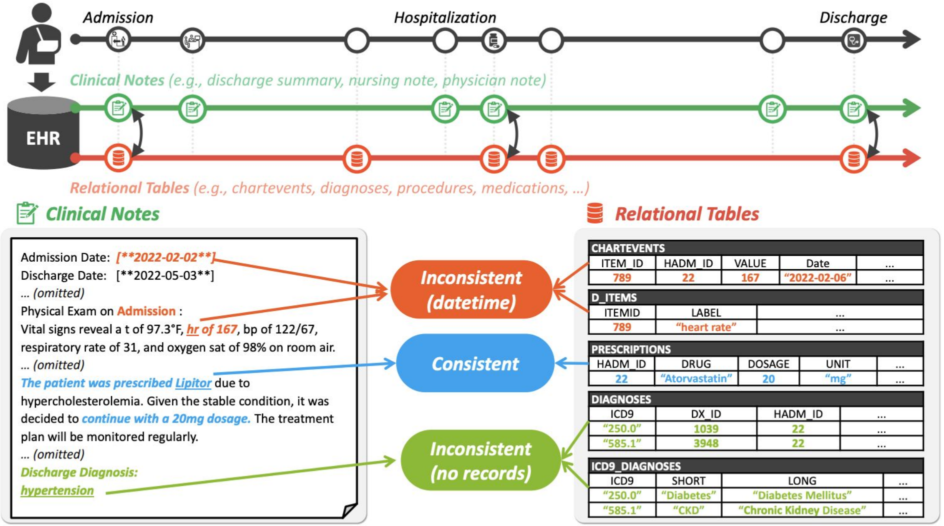
- Inconsistencies can arise between the two sets of data for several reasons
  a. EHR interfaces are often designed with a focus on administrative and financial tasks, which makes it difficult to accurately document clinical information
  b. Overburdened practitioners might unintentionally introduce errors by importing incorrect medication lists, copying and pasting outdated records, or entering inaccurate test results
- Theses errors can lead significant discrepancies between the structured data and clinical notes in the EHR
  a. Potentially jeopardizing patient safety and leading to legal complications

Motivation: Consistency check between Clinical Notes and EHR Tables

- Manual scrutiny of these records is both time-intensive and costly, underscoring the necessity for automated interventions
  a. Despite the need for automated systems, previous studies on consistency check between tables and text have primarily focused on single claims and small-scale single tables
  b. These approaches are not designed for the complex and large-scale nature of EHRs, which require more comprehensive and scalable solutions

New Task and Dataset: EHRCon

- EHRCon
  - We propose a new task and dataset
    - which is designed to verify the consistency between clinical notes and large-scale relational databases in EHRs

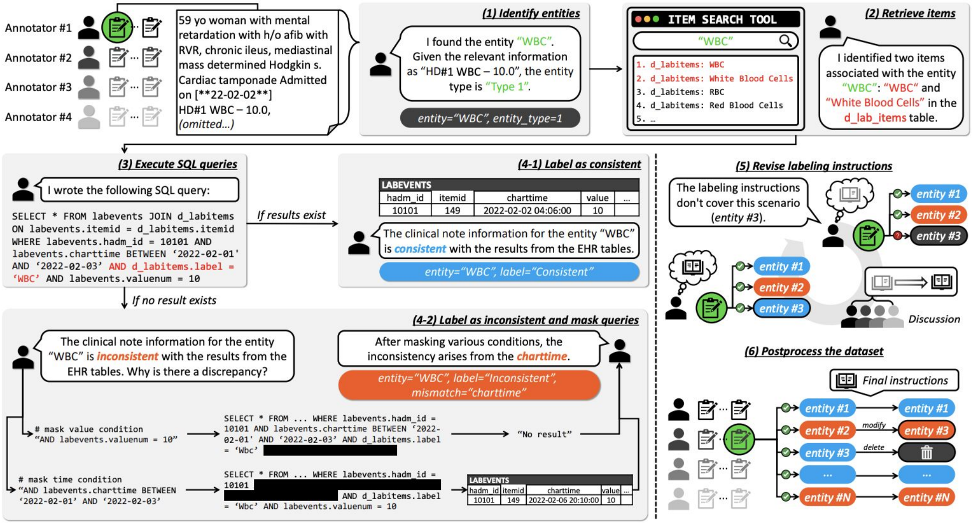


New Task and Dataset: EHRCon

- EHRCon
  - Scope
    1. Consistency check across 13 tables for three types of clinical notes
    2. Label - Consistent / Inconsistent
    3. Definition of Entity Types
       1. Entities with numerical values,
       2. Entities without values but whose existence can be verified in the database
    4. Time Expression
       1. Event time written in a standard time format
       2. Event time described in a narrative style
       3. Time information of the event not written

Annotation Process: EHRCon

1. Identify entities and relevant information within the note
2. Extract information related to the entity from the notes (e.g., dates, values, units)
3. Use extracted information to generate SQL queries and execute them.
4. Review the results to label the entity as either CONSISTENT or INCONSISTENT
5. Upon completing all annotations
    - The annotators engaged in a post-processing phase to ensure high-quality data

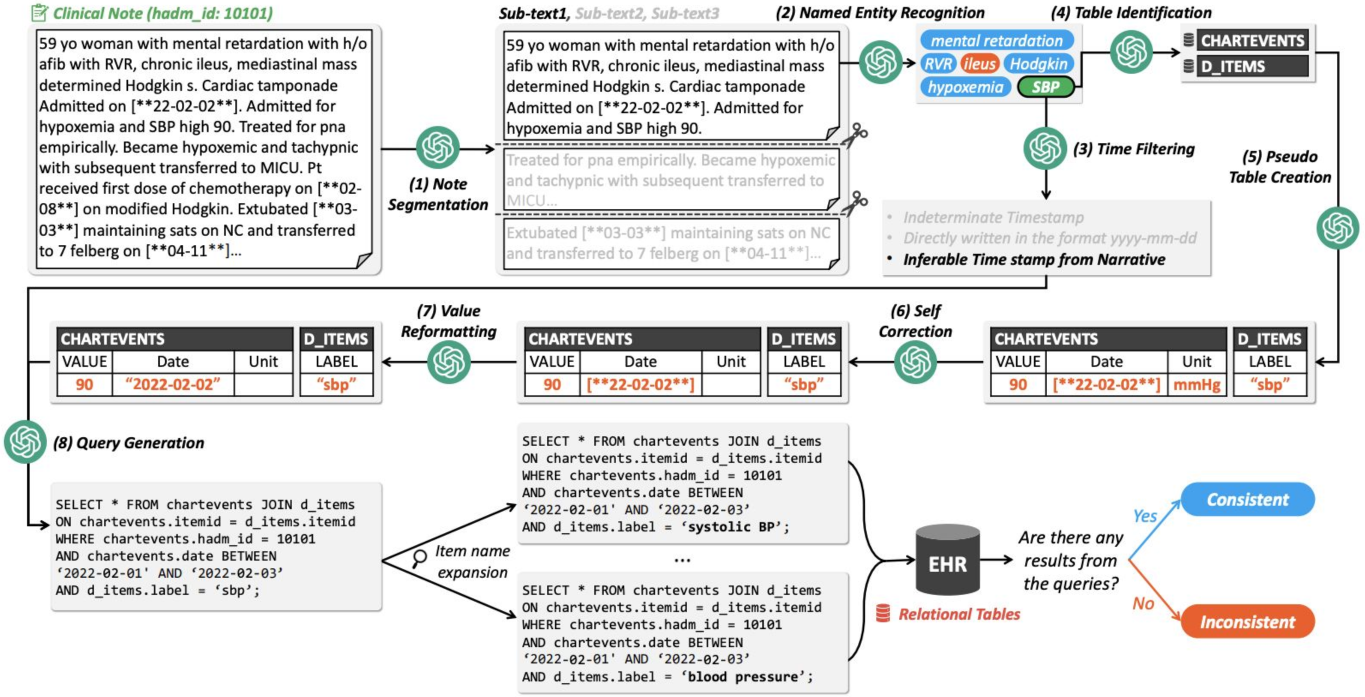


Statistics

Note Type	Entity			Labels	Note	
	Mean Num	Total Num	Type 1 / 2	Con. / Incon.	Total Num	Mean Length
Discharge Summary	50.21	1,908	1,400 / 508	1,181 / 727	38	2,789
Physician Note	46.36	1,530	1,111 / 419	1,230 / 300	33	1,859
Nursing Note	19.50	663	500 / 163	522 / 141	34	1,111
Total	39.06	4,101	3,011 / 1,090	2,933 / 1,168	105	1,953

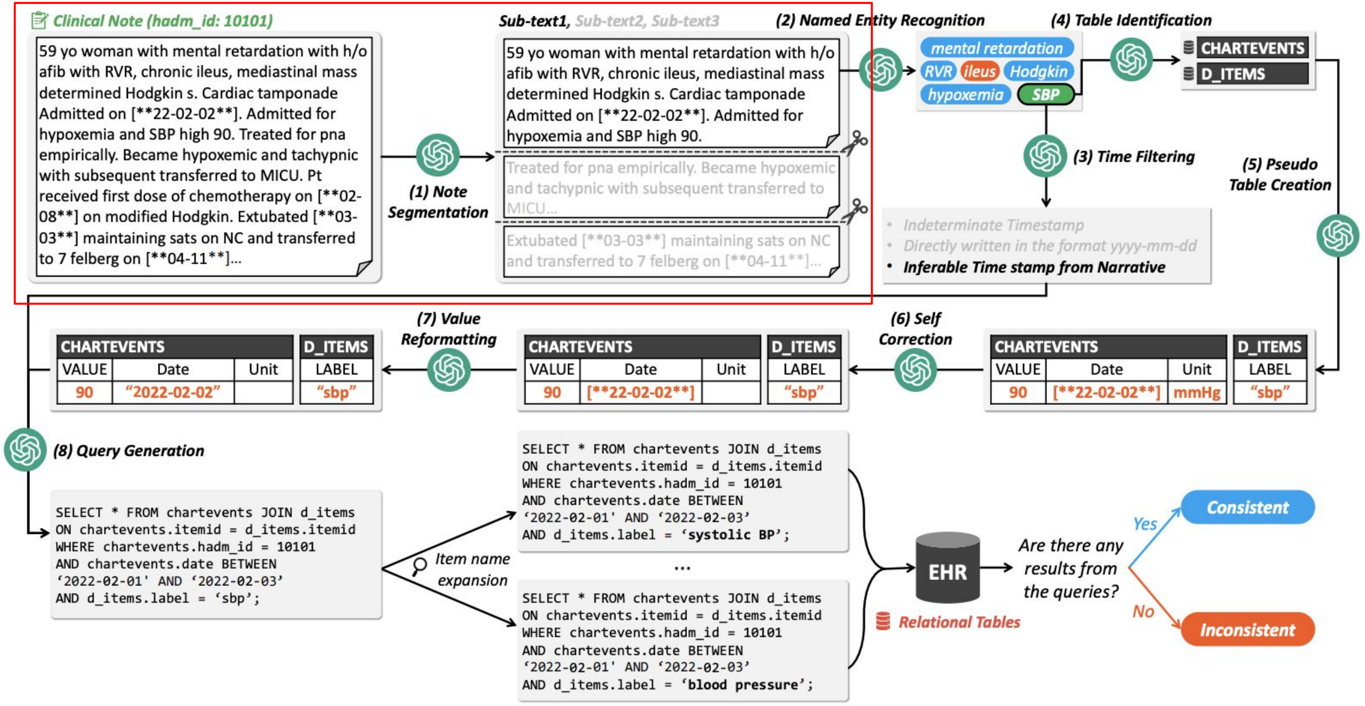
Novel Framework: CheckEHR

- Overall



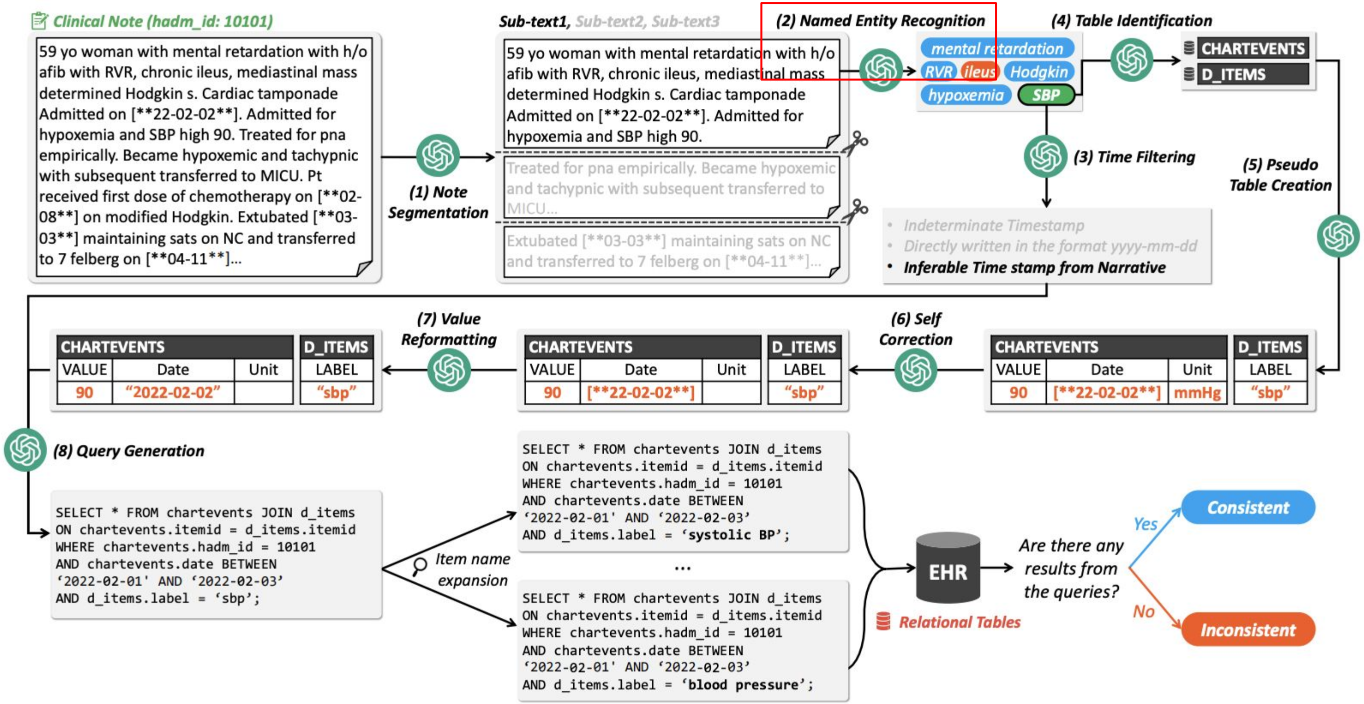
Novel Framework: CheckEHR

- Note Segmentation
  - Divides the entire clinical note into smaller sub-texts that each focus on a specific topic



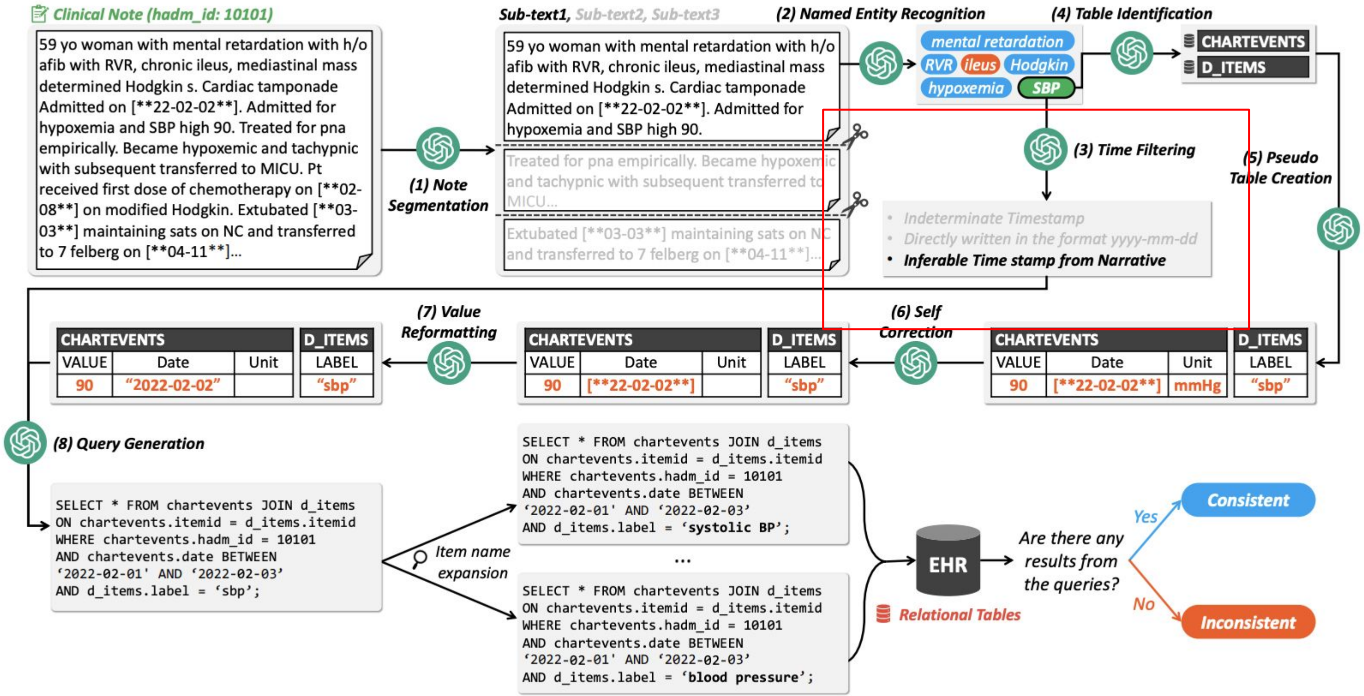
Novel Framework: CheckEHR

- Named Entity Recognition
  - Extract entities



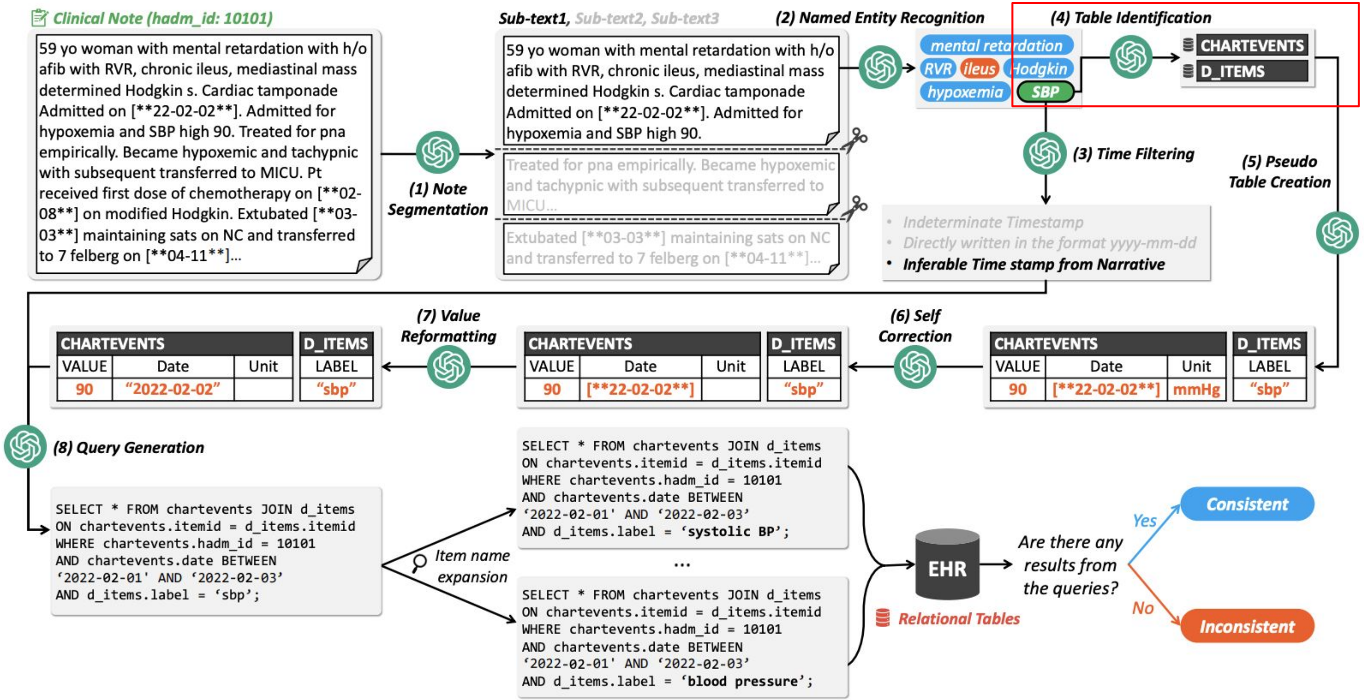
Novel Framework: CheckEHR

- Time Filtering
  - LLM determines whether the time expression of a clinical event
    - specific time format, written in a narrative style, or if the time is not specified.



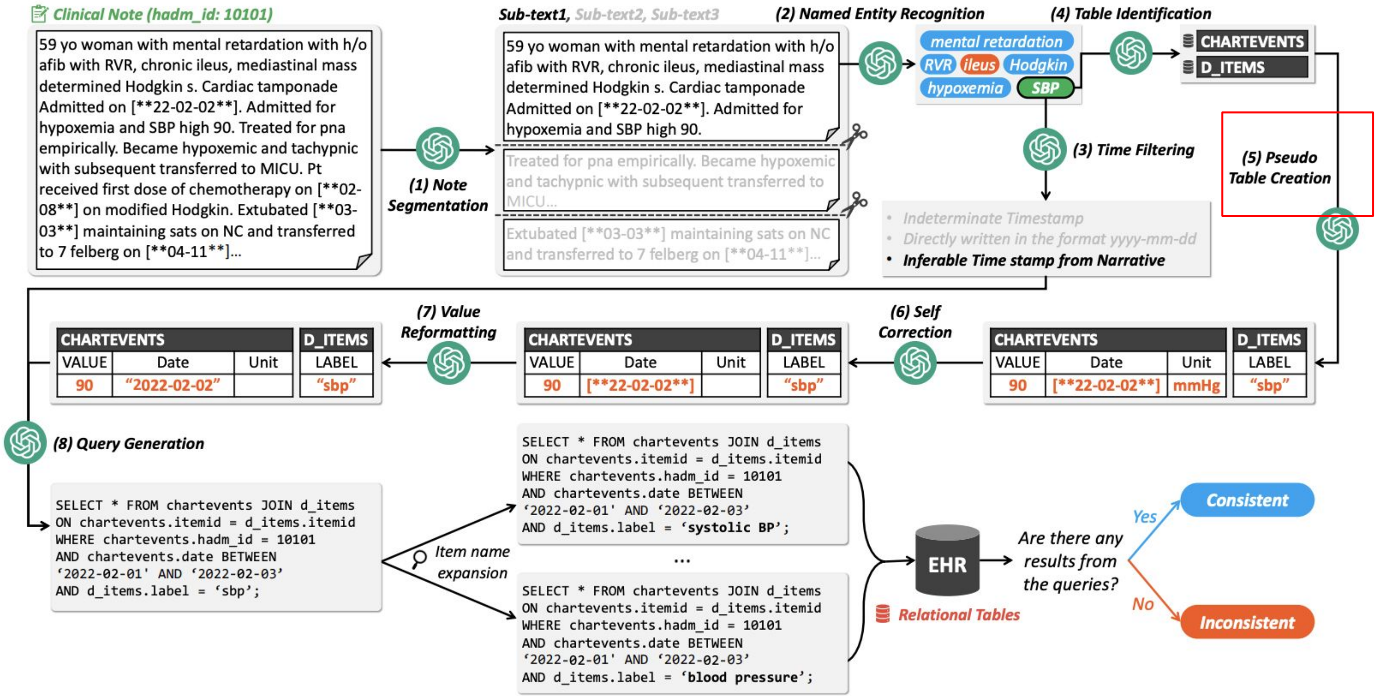
Novel Framework: CheckEHR

- Table Identification
  - Identify the relevant tables related to the entities



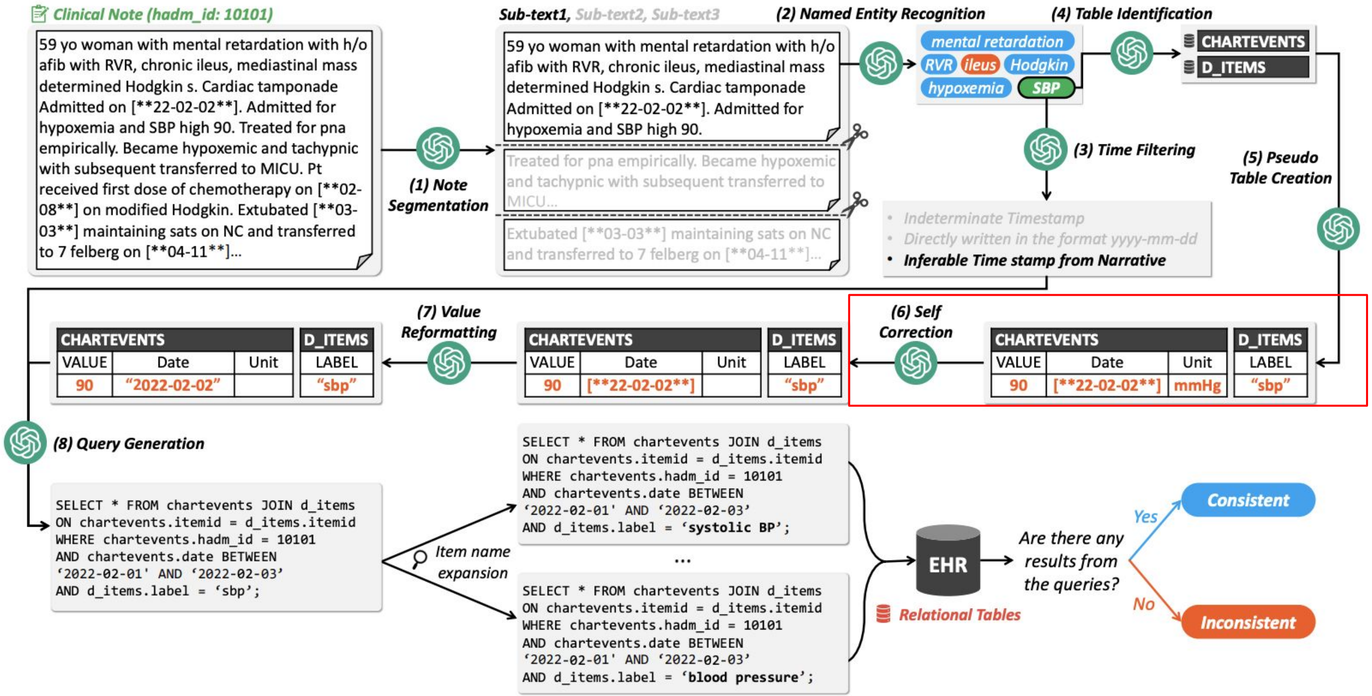
Novel Framework: CheckEHR

- Pseudo Table Creation
  - Creates a pseudo table to effectively extract table-related information



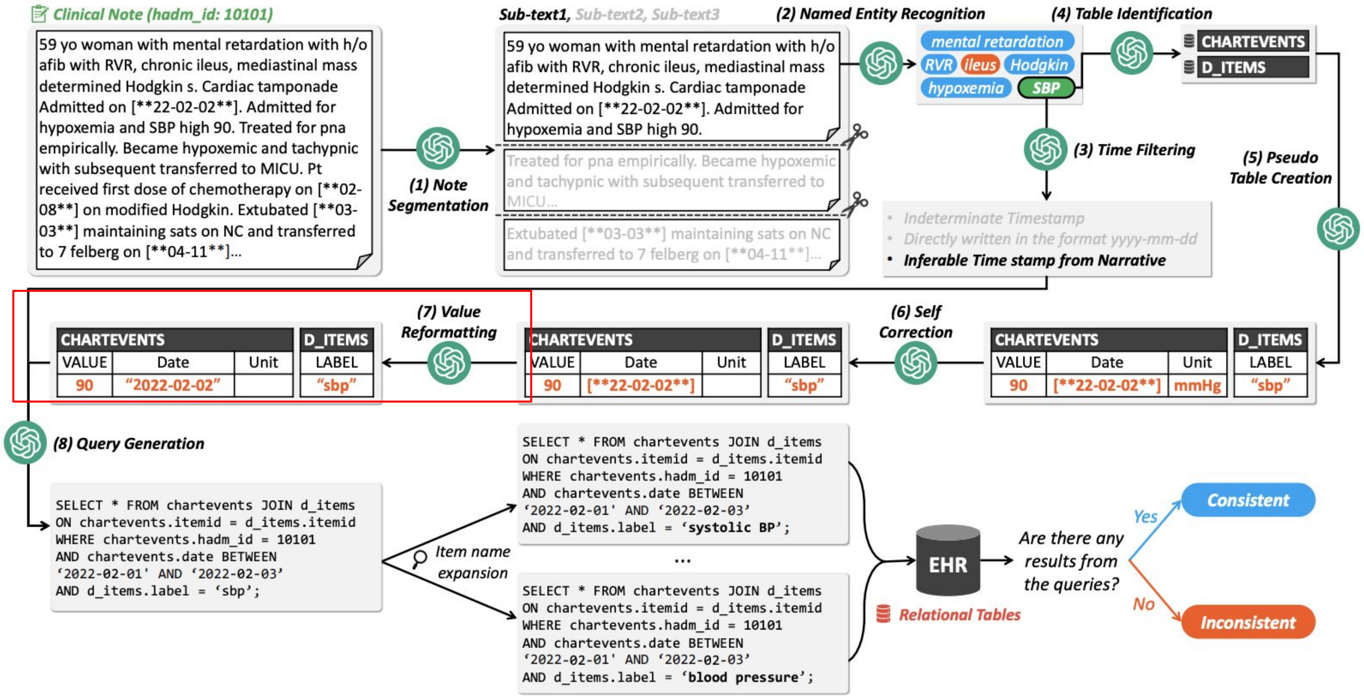
Novel Framework: CheckEHR

- Self-Correction
  - To address hallucination issue, the LLM re-evaluates whether the pseudo table created in the previous stage is directly aligned with the notes



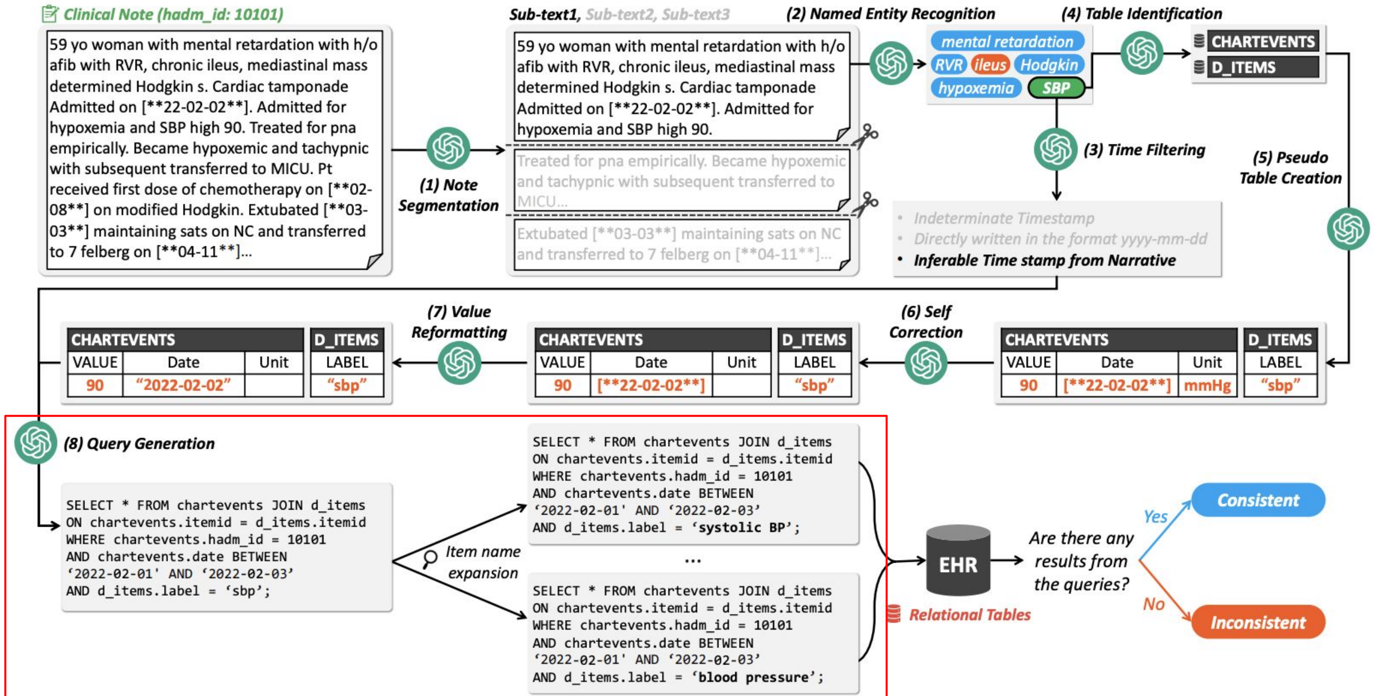
Novel Framework: CheckEHR

- Value Reformatting
  - Align the data types between the generated pseudo table and the actual table



Novel Framework: CheckEHR

- Query Generation
  - Using the results from the Time Filtering and Value Reformatting, the LLM creates an SQL query



Results

- The main results of CheckEHR on MIMIC-III

Shot	Models	Discharge Summary			Physician Note			Nursing Note			Total		
		Rec	Prec	Inters	Rec	Prec	Inters	Rec	Prec	Inters	Rec	Prec	Inters
Zero	Tulu2	11.82	27.48	46.92	9.1	20.15	40.83	15.32	23.23	30.37	12.08	23.62	38.37
	Mixtral	-	-	-	-	-	-	-	-	-	-	-	-
	Llama-3	50.82	35.54	69.70	52.92	33.89	72.71	53.45	44.61	81.48	**52.39**	38.01	**74.03**
	GPT-3.5 (0613)	45.04	46.71	74.58	40.14	37.32	70.07	43.30	44.53	70.81	42.83	**42.85**	71.82
Few	Tulu2	40.01	49.42	70.66	49.98	47.08	85.33	44.77	40.50	78.40	44.95	45.66	78.13
	Mixtral	54.70	49.76	71.21	53.71	37.97	83.48	69.86	49.65	85.01	54.70	45.79	79.90
	Llama-3	50.44	47.01	76.25	56.11	42.75	84.30	52.60	38.08	75.96	53.05	42.61	78.83
	GPT-3.5 (0613)	64.31	54.64	81.60	54.64	44.01	81.41	64.25	47.25	95.74	**61.06**	**48.63**	**86.25**

Summary

- We introduce EHRCon, a carefully crafted dataset designed to improve the accuracy and reliability of EHRs
- Alongside EHRCon, we present CheckEHR, an innovative framework that leverages LLMs to efficiently verify data consistency within EHRs.
- Our study lays the groundwork for future advancements in automated and dependable healthcare documentation systems, ultimately enhancing patient safety and streamlining healthcare processes.

Thank you!

Dataset : https://github.com/dustn1259/EHRCon
Contact : yeonsu.k@kaist.ac.kr , jiho.kim@kaist.ac.kr