

# Visualizing and Understanding the Errors of Information Extraction Results

System Demonstrations: Data Science  
S39

**Huan He, PhD**

Mayo Clinic

Twitter: #IS23



# Disclosure

---



I and my spouse/partner have no relevant relationships with commercial interests to disclose.

# Learning Objectives

---

After participating in this session, the learner should be better able to:

- Learn a serverless architecture for designing web-based tools
- Learn the visual designs to support error analysis of information extraction results

# Outline

---

1. Background - Error analysis for information extraction
2. System Design - Architecture and Visualization
3. Demo

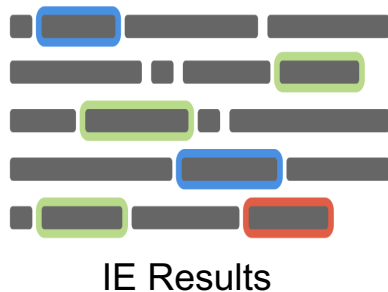
# Background - Information Extraction (IE)

Information extraction (IE) is the task of automatically extracting structured information from unstructured and/or semi-structured machine-readable documents and other electronically represented sources.

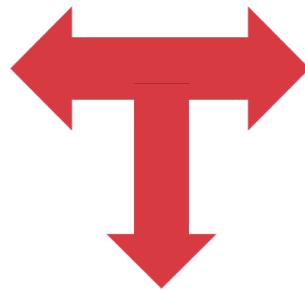
In our cases, the IE task is mainly processed by the natural language processing (NLP) techniques:



# Background - IE Errors



Comparison



Gold Standard  
Corpus (GSC)

Most of the IE results may be correct ...  
but ...



**Errors!?**

# Background - Error Types



- **False Positives:** IE systems can sometimes identify information that is not actually present in the GSC, leading to false positives.
- **False Negatives:** IE systems can also miss relevant information in the GSC, leading to false negatives.

# Background - Error Categories



- **Linguistic errors.** Synonymy, anaphora resolution errors, etc.
- **Contextual errors.** Negation, temporal, section, certainty, subject, exclusion, etc.
- **Annotation errors.** Missing annotation, incomplete annotation, inaccurate annotation, terminology, bias, etc.
- **Definition errors.** Ambiguity, inconsistency of definition, change of status, etc.
- **Other.**



# Background - Error Analysis



=H4/SUM(H3:H4)						A	B	A	B	C
C	D	E	F	G	H	item	result	concept	FP	FN
1				Predicted Classification		2	TP	60	2	2
2				Negative	Positive	3	FP	16	1	1
3				9	7	4	FN	18	0	2
4		Actual classification	Negative	4	4	5	Error Rate	0.36170213	1	1
5			Positive			6	Accuracy	0.63829787	1	0
6		Recall			0.5	7	Precision	0.78947368	1	0
7		Accuracy			0.541666667	8	Recall	0.76923077	1	0
8		Precision			0.363636364	9	F1-Score	0.77922078	4	10
9		F1 Score				10			5	2
10										

Confusion matrix, categorize the errors by type and concepts

A	B	C	D	E	F	G	H
spans	tag	text	error	file_name	sentence_sp	sentence	label_1
206~212	VAX	PFIZER	FP	sys_doc_1.tx	47~53	A 61-year-old male patient received bnt162b2	Exclusion
265~271	VAX	EL1284	FP	sys_doc_1.tx	26~32	second dose (lot number: EL1284, expiry date	Exclusion
589~594	OTHER_AE	Sulfa	FN	G_doc_1.tx	43~48	Medical history included known allergies: Sulf	Exclusion
596~607	OTHER_AE	Hypothyroid	FN	G_doc_1.tx	50~61	Medical history included known allergies: Sulf	Exclusion
609~615	OTHER_AE	Asthma	FN	G_doc_1.tx	63~69	Medical history included known allergies: Sulf	Exclusion
617~631	OTHER_AE	Hi Cholesterol	FN	G_doc_1.tx	71~85	Medical history included known allergies: Sulf	Exclusion
692~698	MEDICATION	ZYRTEC	FP	sys_doc_1.tx	60~66	Concomitant medication included cetirizine h	Exclusion
666~690	MEDICATION	cetirizine hydrochl	FN	G_doc_1.tx	34~58	Concomitant medication included cetirizine h	Absence of C
168~176	VAX	BNT162B2	FN	G_doc_2.tx	75~83	A female patient of an unspecified age (age:	Incorrect Ann
251~257	VAX	expiry	FN	G_doc_2.tx	158~164	A female patient of an unspecified age (age:	Incorrect Ann

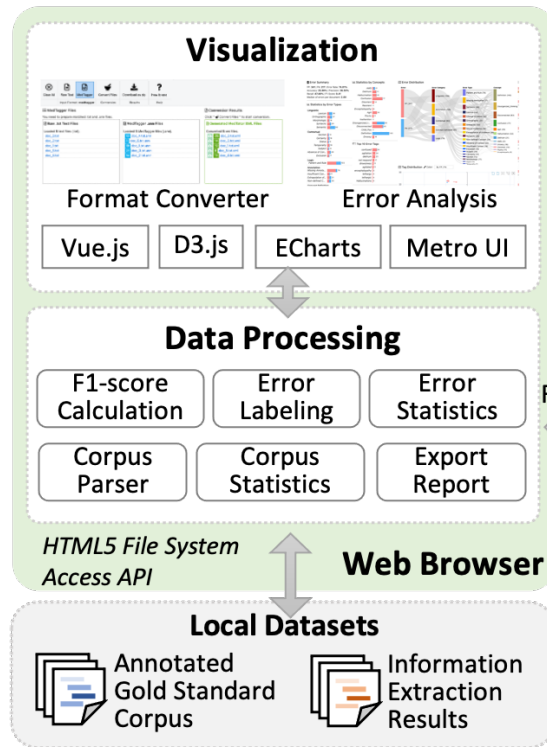
Evaluate the errors and analyze the causes of errors

# Challenges in error analysis

1. Due to the complexity of the error types, it's hard to gain insights into errors
  1. FP / FN
  2. Linguistic, contextual, annotation, definition, others, etc.
  3. Annotation concepts
  4. Project related types
2. Manually labeling errors is time-consuming without assistance

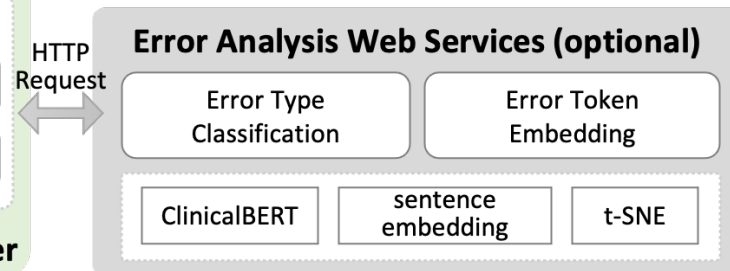
	A	B	C	D	E	F	G	H
1	<b>spans</b>	<b>tag</b>	<b>text</b>	<b>error</b>	<b>file_name</b>	<b>sentence_sp</b>	<b>sentence</b>	<b>label_1</b>
2	206~212	VAX	PFIZER	FP	sys_doc_1.tx	47~53	A 61-year-old male patient received bnt162b2	Exclusion
3	265~271	VAX	EL1284	FP	sys_doc_1.tx	26~32	second dose (lot number: EL1284, expiry date	Exclusion
4	589~594	OTHER_AE	Sulfa	FN	G_doc_1.tx	43~48	Medical history included known allergies: Sulf	Exclusion
5	596~607	OTHER_AE	Hypothyroid	FN	G_doc_1.tx	50~61	Medical history included known allergies: Sulf	Exclusion
6	609~615	OTHER_AE	Asthma	FN	G_doc_1.tx	63~69	Medical history included known allergies: Sulf	Exclusion
7	617~631	OTHER_AE	Hi Cholesterol	FN	G_doc_1.tx	71~85	Medical history included known allergies: Sulf	Exclusion
8	692~698	MEDICATION	ZYRTEC	FP	sys_doc_1.tx	60~66	Concomitant medication included cetirizine h	Exclusion
9	666~690	MEDICATION	cetirizine hydrochl	FN	G_doc_1.tx	34~58	Concomitant medication included cetirizine h	Absence of C
10	168~176	VAX	BNT162B2	FN	G_doc_2.tx	75~83	A female patient of an unspecified age (age:	Incorrect Ann
11	251~257	VAX	expiry	FN	G_doc_2.tx	158~164	A female patient of an unspecified age (age:	Incorrect Ann

# System Design - Visual Error Analysis



A web-based **serverless** visual error analysis tool that facilitates the exploration of IE errors

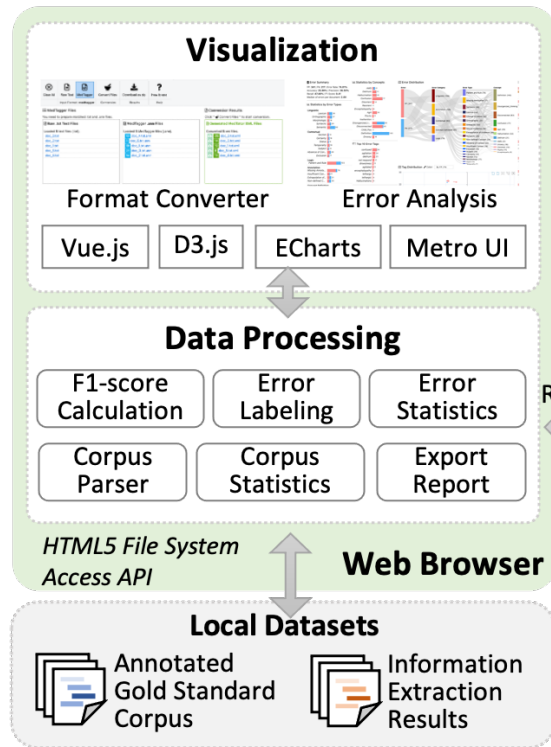
- Completely run in the user's web browser
- All features are implemented within the browser
- A server is **NOT** required



**Optional** server support for advanced features:

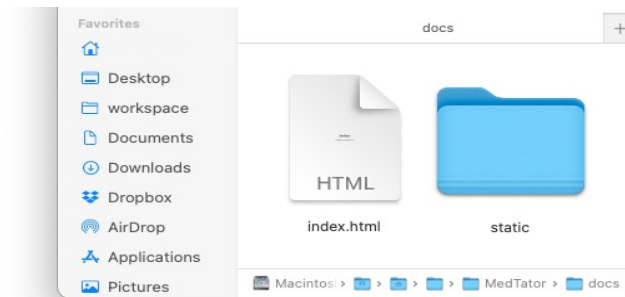
- Error type classification based on ML
- Error token embedding and dimension reduction

# System Design - Visual Error Analysis



## The advantages of “Serverless” architecture

1. No data is sent out:
  - Run locally in user’s web browser
  - Operate text files on locally
2. No “install”:
  - No need to install any runtimes (e.g., Java, Python, Nginx, etc.)



# Demonstration - Sample Dataset

## Vaccine Adverse Event Reporting System (VAERS) Dataset

The VAERS database contains information on unverified reports of adverse events (illnesses, health problems and/or symptoms) following immunization with US-licensed vaccines.

raw\_txt

doc\_1.txt

doc\_2.txt

doc\_3.txt

doc\_4.txt

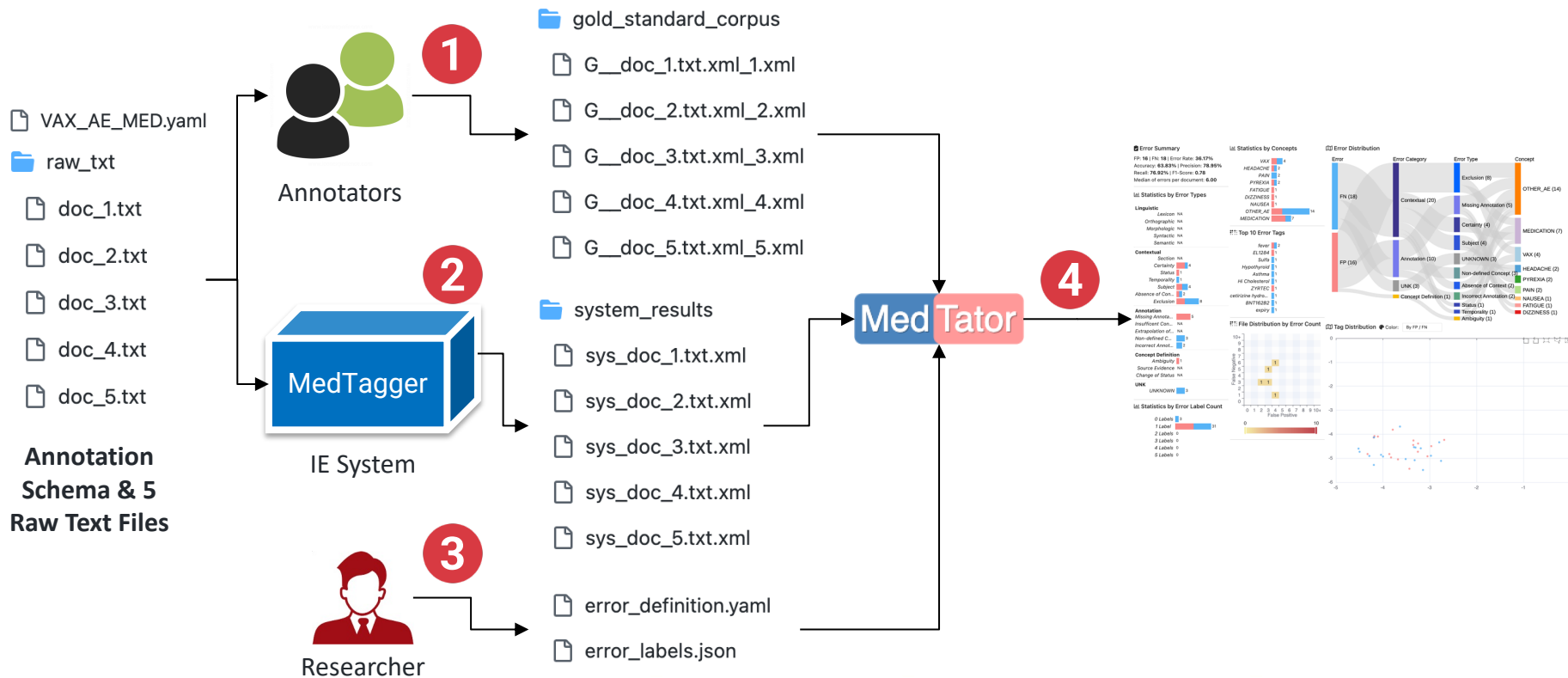
doc\_5.txt

5 text files

- 1 Mildly feverish evening after vaccine; Soreness and tenderness at left arm injection site; This is a spontaneous report from a contactable physician (patient). A 61-year-old male patient received bnt162b2 (PFIZER-BIONTECH COVID-19 VACCINE; second dose (lot number: EL1284, expiry date: unknown) via an unspecified route of administration on 07Jan2021 07:00 at single dose (Left arm) for COVID-19 immunization. Patient took bnt162b2 first dose (lot number: EK5730, expiry date: unknown) on 17Dec2020 07:00 am (Left arm) for COVID-19 immunization. Medical history included known allergies: Sulfa, Hypothyroid, Asthma, Hi Cholesterol. Concomitant medication included cetirizine hydrochloride (ZYRTEC), atorvastatin, levothyroxine. The patient previously took tamiflu and experienced allergies. Facility type vaccine was hospital. The patient did not receive any other vaccines within 4 weeks prior to the COVID vaccine. Prior to vaccination, the patient was not diagnosed with COVID-19. Since the vaccination, the patient has not been tested for COVID-19. The patient experienced Mildly feverish evening after vaccine, Soreness and tenderness at left arm injection site on 07Jan2021. The patient took Motrin as treatment. Therapeutic measures were taken as a result of mildly feverish evening after vaccine, soreness and tenderness at left arm injection site. The outcome of the events was recovered on an unspecified date. The events were reported as non-serious.

Raw text files

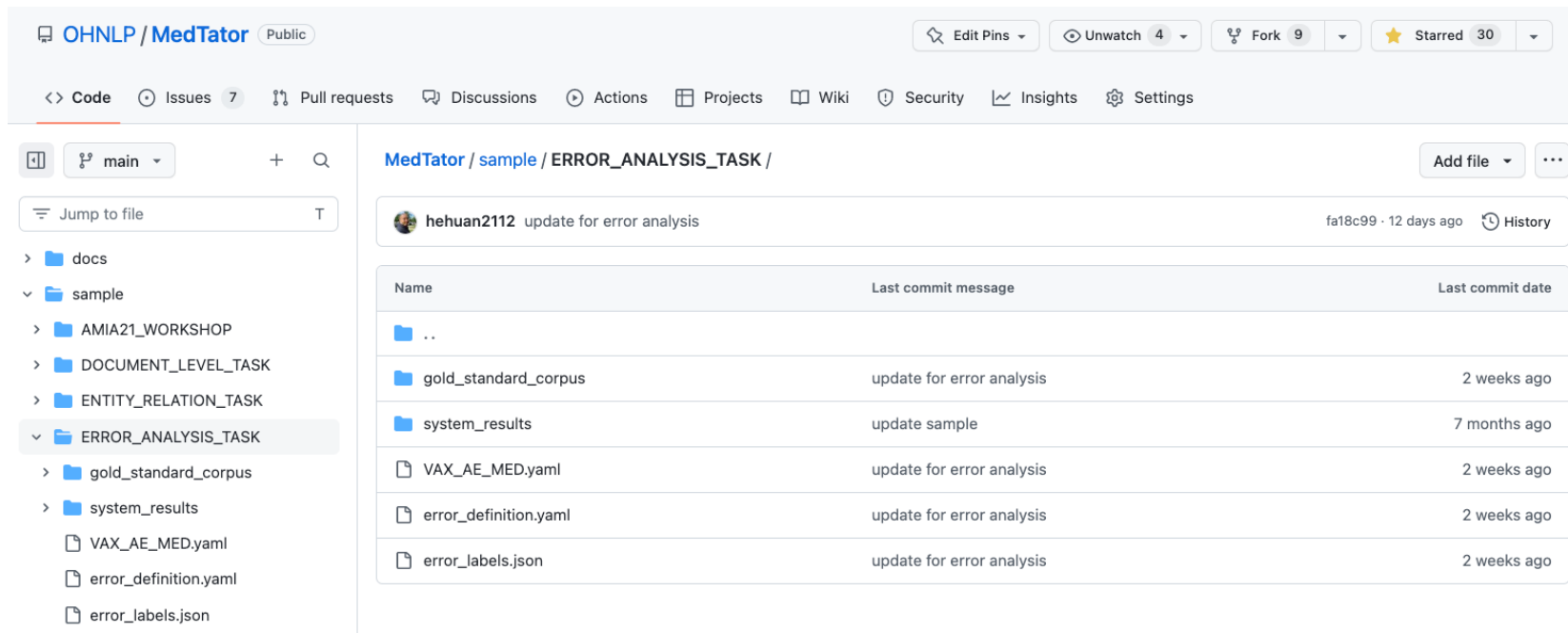
# Demonstration - Sample Dataset



# Demonstration

Live demo: <https://ohnlp.github.io/MedTator/>

A sample dataset: [https://github.com/OHNLP/MedTator/tree/main/sample/ERROR\\_ANALYSIS\\_TASK](https://github.com/OHNLP/MedTator/tree/main/sample/ERROR_ANALYSIS_TASK)



OHNLP / MedTator (Public)

Edit Pins Unwatch 4 Fork 9 Starred 30

Code Issues 7 Pull requests Discussions Actions Projects Wiki Security Insights Settings

main

Jump to file

- docs
- sample
  - AMIA21\_WORKSHOP
  - DOCUMENT\_LEVEL\_TASK
  - ENTITY\_RELATION\_TASK
  - ERROR\_ANALYSIS\_TASK**
    - gold\_standard\_corpus
    - system\_results
      - VAX\_AE\_MED.yaml
      - error\_definition.yaml
      - error\_labels.json

MedTator / sample / ERROR\_ANALYSIS\_TASK /

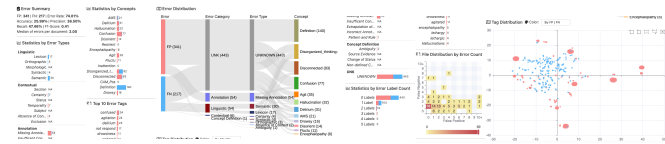
hehuan2112 update for error analysis fa18c99 · 12 days ago History

Name	Last commit message	Last commit date
..		
gold_standard_corpus	update for error analysis	2 weeks ago
system_results	update sample	7 months ago
VAX_AE_MED.yaml	update for error analysis	2 weeks ago
error_definition.yaml	update for error analysis	2 weeks ago
error_labels.json	update for error analysis	2 weeks ago

# Takeaways

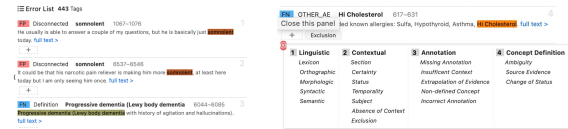
- Use multiple coordinated charts to visualize the errors from different perspectives to gain insights into the errors

- Stacked bar charts
- Sankey diagram
- Scatter plot for texts



- Using a well-designed user interface to facilitate the manual labeling

- Reduce cognitive load
- Improve efficiency



- A serverless architecture that removes the need for an always-on server

- Easy to deploy and distribute (by free static web servers)
- No data will be sent out



# Acknowledgments

## Co-authors and team members



Sunyang Fu



Liwei Wang



Andrew Wen



Sijia Liu



Sungrim Moon



Kurt Miller



Hongfang Liu



Donna Ihrke



Heling Jia



Taylor Hawkins

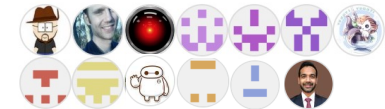


Katelyn Cordie



Samuel McKinven

## Community Participants



This work was supported by the National Center for Advancing Translational Sciences of the National Institutes of Health under award number U01TR002062.



National  
COVID  
Cohort  
Collaborative

Collaborative Analytics  
Workstream  
Subgroup of  
Natural Language  
Processing

# Thank you!

Email: [Liu.Hongfang@mayo.edu](mailto:Liu.Hongfang@mayo.edu)  
[He.Huan@mayo.edu](mailto:He.Huan@mayo.edu)

