

---

# Detecting Product Innovation in News Articles

**Sponsor: Dr. Gizem Korkmaz, UVA Biocomplexity Institute & Initiative (BII)**

**Advisor: Gerard P. Learmonth**

— By: Keyu Chen, Ben Cosgro,  
Oretha Domfeh, Alex Stern —

---

# Problem Description and Methods

Sponsor: UVA Biocomplexity Institute & Initiative (BII)

## Traditional Innovation Measures

### Example BRDIS Survey Question:

During the three years 2014 to 2016, did your company introduce:

- a. New or significantly improved goods (Exclude the simple resale of new goods purchased from other companies and changes of a solely aesthetic nature)?  Yes  No
- b. New or significantly improved services? .....  Yes  No

## Questions we want to answer

- Whether the article mentions an innovation?
- What company launched the innovation?
- What are the product names that the company launched?

## Our Methods

### Data Pre-processing

- 1200 labeled data for training and testing
- 360K news articles

### Innovation Classification

- Detect innovation articles

### Name Entity Recognition

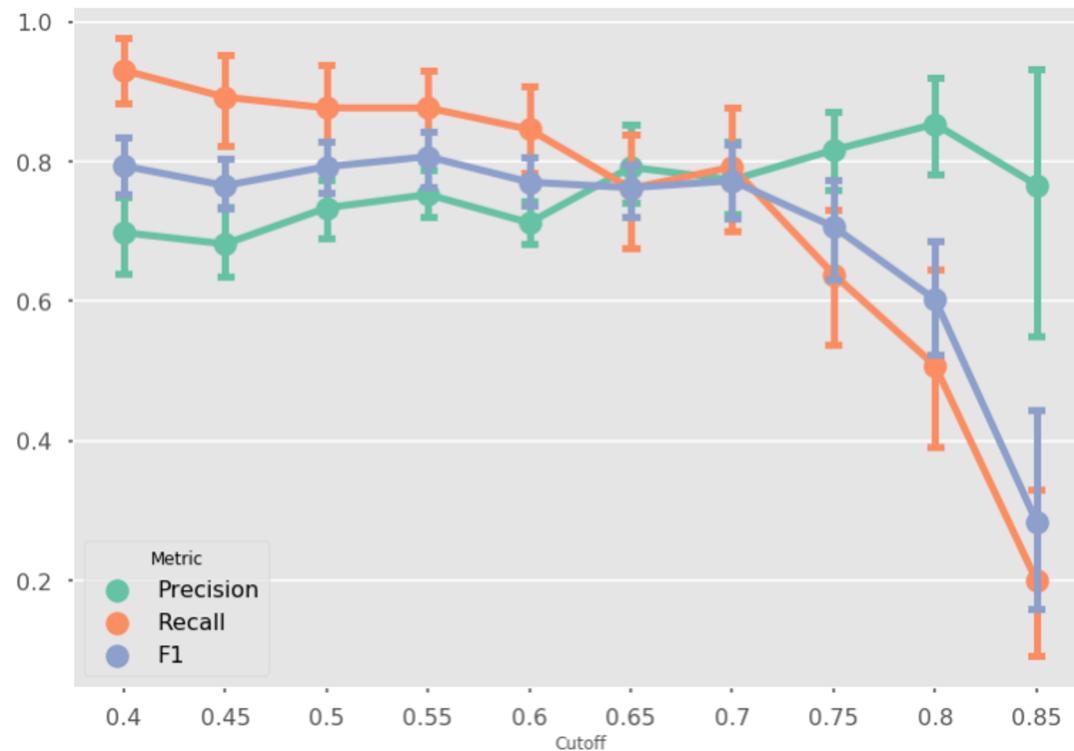
- Extract company names

### Question Answering

- Extract company and product names

### Compare Detected Innovators to Companies in FDA

# Classification



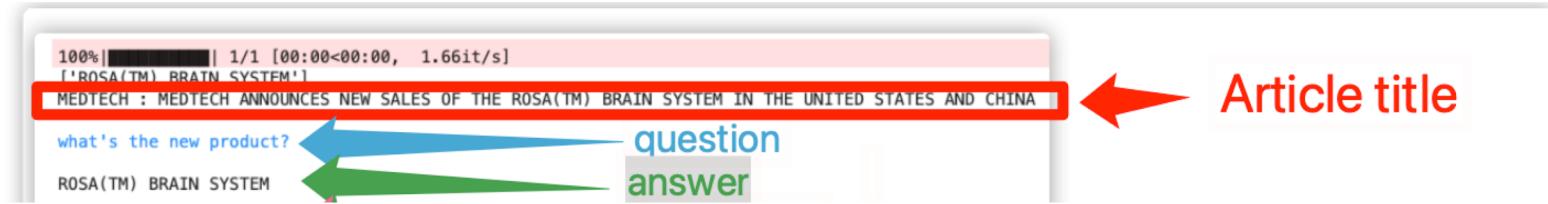
# Named Entity Recognition

- Identifies and categorizes entities within an article into 4 categories
  - Location, Person, Organization, Miscellaneous

Subset of Article Processed	Hand-Labeled Innovator	Companies from DNA
<i>Title Only</i>	62.5%	53.1%
<i>Paragraph Only</i>	87.5%	65.6%
<i>Title and Paragraph</i>	90.6%	75%

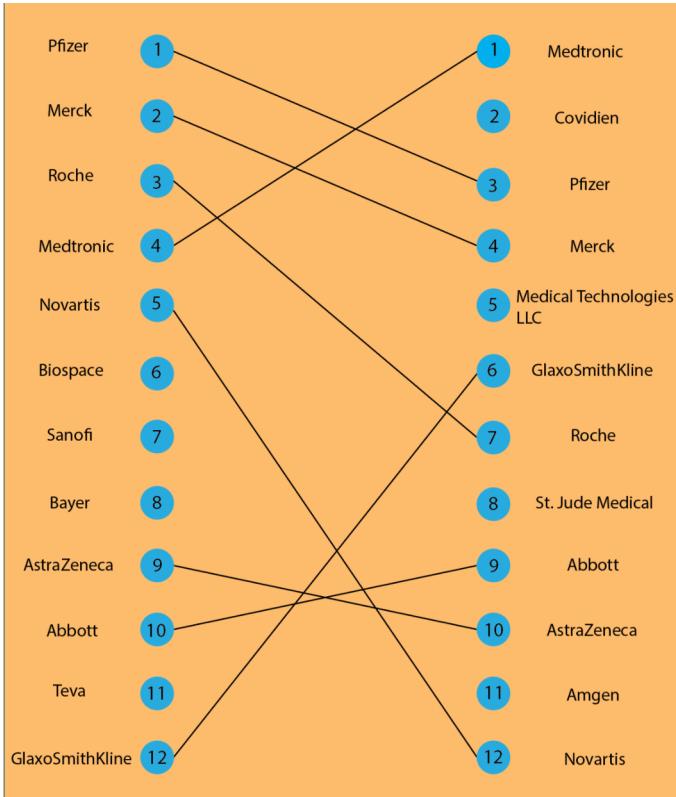
# Question Answering

- Retrieves key information (company and product names)



Subset of Article Processed	Company Names	Product Names
Individual answers	78%	81%
Combined Answers (title and title+lead)	97%	91%
Combined answers (title+lead)	94%	91%

# Results



Comparison of QA and NER top 12 and top 50 produces **67%** matching rate

## Comparison of Top 50 Unique Companies in QA and NER results to other Databases

<i>National Drug Codes</i>	8%
<i>FDA Drug Approvals</i>	44%